Reaction Schemes that are Easily Confused with a Reversible First Order Reaction

Detailed Kinetic Analysis of Two Reaction Schemes that are Easily Confused with the Reversible First Order Reaction

Ágnes Balogh[a], Gábor Lente[a]*, József Kalmár[b], István Fábián[a]

[a]Department of Inorganic and Analytical Chemistry, University of Debrecen, P.O.B. 21, Debrecen, H-4010 Hungary. Fax: + 36 52 489667; Tel: + 36 52 512900/22373; E-mail: lenteg@science.unideb.hu

[b] MTA-DE Homogeneous Catalysis and Reaction Mechanisms Research Group, P.O.B. 21, Debrecen, H-4010 Hungary

Abstract

A detailed kinetic analysis of two schemes, one involving coupled consecutive processes and another featuring parallel reactions and decay of ???, is presented here using Taylor series expansion. It is shown that both of these schemes are easily confused with the reversible second order reaction in a routine kinetic study. The kinetic traces predicted by both schemes are sufficiently close to pseudo-first order curves so that it is practically impossible to identify the deviations based on data with the usual experimental errors, which was also demonstrated by fitting simulated theoretical curves to exponential functions. The dependence of the pseudo-first order rate constants on the concentration of the excess reagent features the same trend as in the case of a reversible reaction: a straight line with an intercept is observed. This analysis emphasizes that the reversible nature of reactions should be demonstrated by direct equilibrium studies, kinetic observations alone might be misleading.

Keywords

statistical kinetics, kinetic modelling, reversible reactions, kinetic analysis, pseudo-first order kinetics

Introduction

Statistical kinetics is a phenomenon when certain coincidences in the values of rate constants and other parameters make the kinetic behavior of a multi-step process indistinguishable from a single-step scheme.[1,2] A notable and well understood case is based on two consecutive first order reactions:

$$A \xrightarrow{k_a} B \xrightarrow{k_b} C \qquad (1)$$

If this reaction is monitored by absorbance measurements, the time dependence of the absorbance signal is given by the following equation:[1,2]

$$\mathrm{Abs} = \frac{\varepsilon_A (k_a - k_b) - \varepsilon_B k_a + \varepsilon_C k_b}{k_a - k_b} c_0 e^{-k_a t} + \frac{(\varepsilon_B - \varepsilon_C) k_a}{k_a - k_b} c_0 e^{-k_b t} \qquad (2)$$

In this formula, c_0 is the initial concentration of A, whereas ε_A, ε_B and ε_C are the molar absorptivities of species A, B, and C at the wavelength of the detection. Equation 2 describes a double exponential function, but certain parameter coincidences may transform it to a single exponential behavior. One of such coincidences is $\varepsilon_B = \varepsilon_C$, which makes the second term undetectable in equation 2. Another coincidence, $(\varepsilon_A - \varepsilon_B)k_a = (\varepsilon_A - \varepsilon_C)k_b$, results in 0 as the multiplying factor before the first term. These coincidences may seem rare, but they are quite common (and also valid at the entire useful wavelength range) in practice when multiple equivalent reaction centers and absorbing moieties are involved. Experimental examples of this phenomenon can be found both in the recent and older chemical literature.[3-11]

It should be emphasized that in the classic cases of statistical kinetics,[1-11] there are no approximations or experimental limitations involved, it is the exact mathematical solution of the scheme that may be misleading in experimental studies because of the coincidences. However, there is another and arguably much more common reason why a kinetic curve could be misidentified as a simpler one: the usual experimental errors in measuring the absorbance or other features of the reactive system may make two mathematically distinct types of

kinetic curves in distinguishable. This effect is sometimes deliberately taken advantage of: the classical example is to use a high excess of each reagent except one to create pseudo-first order conditions.[2]

This paper will give a detailed mathematical and statistical analysis of two reaction schemes that give kinetically, at least seemingly, indistinguishable results from the well-known reversible second order reaction[1,2]:

$$A + B \underset{k_-}{\overset{k_+}{\rightleftharpoons}} C \qquad (3)$$

The usual approach is that one of the reagents (e.g. B) is used in large excess in order to detect pseudo-first order curves, whose observed rate constants (k_{obs}) are determined as a function of the concentration of the excess reagent B. The expected dependence is given as follows:

$$k_{obs} = k_+[B] + k_- \qquad (4)$$

If the plot of k_{obs} vs. [B] gives a straight line with an intercept, this is understood as a validation of the scheme. Rate constant k_+ is calculated from the slope of the plot and k_- is the intercept.

In the recent literature, two cases have been identified when the procedure described in the previous paragraph was successful, yet the possibility of a reversible second order reaction was ruled out based on independent experimental observations or plausible chemical considerations.[12-14] In both cases, numerical integration of alternative kinetic schemes was used to show that other interpretations of the experimental data are also in agreement with the kinetic observations. The present study identifies the sources of the coincidences leading to this interesting phenomenon and also analyzes what parameter values (rate constants and initial concentrations) are typically required for distinguishing the kinetic schemes on the basis of kinetic data.

Results and Discussion

Theoretical Background. The calculations presented here will use Taylor series expansion for the mathematical analysis of kinetic curves.[15] This is also a possible method for numerical integration of kinetic schemes.[2] According to Taylor's theorem, the value of a differentiable function $f(t)$ at any value of t can be calculated as an infinite sum (called the Taylor series) using the derivatives of the function at $t = 0$. A possible form of this theorem is given as follows:

$$f(t) = \sum_{i=0}^{\infty} \frac{d^i f}{dt^i}(0)\frac{t^i}{i!} =$$

$$= f(0)\frac{1}{0!} + \frac{df}{dt}(0)\frac{t}{1!} + \frac{d^2 f}{dt^2}(0)\frac{t^2}{2!} + \frac{d^3 f}{dt^3}(0)\frac{t^3}{3!} + \frac{d^4 f}{dt^4}(0)\frac{t^4}{4!} + \frac{d^5 f}{dt^5}(0)\frac{t^5}{5!} + \cdots \tag{5}$$

Taylor's theorem has very diverse applications in science. In chemical kinetics, the rate equation gives the first derivate of a concentration-time trace as the function of the concentrations. This fact offers an advantageous way to derive the Taylor series of a concentration time trace even if the trace itself cannot be given by a closed-form function.

Coupled Consecutive Reactions. The first of the two kinetic schemes analyzed here features an initial first order reaction, the product of which reacts with the initial reagent in a consecutive, second order step.

$$\begin{aligned} &\mathrm{A} \xrightarrow{k_1} \mathrm{B} \\ &\mathrm{B} + \mathrm{A} \xrightarrow{k_2} \mathrm{P} \end{aligned} \tag{6}$$

The rate equations defined for this scheme are given as follows:

$$\begin{aligned} \frac{d[\mathrm{A}]}{dt} &= -k_1[\mathrm{A}] - k_2[\mathrm{A}][\mathrm{B}] \\ \frac{d[\mathrm{B}]}{dt} &= k_1[\mathrm{A}] - k_2[\mathrm{A}][\mathrm{B}] \end{aligned} \tag{7}$$

In practice, this scheme was proposed to interpret kinetic findings in the redox reaction between thiocyanate ion and peroxomonosulfate ion under conditions when thiocyante ion is used in large excess.[14] Therefore, in that particular example, A was HSO_5^-, B was HOSCN, whereas C meant decomposition products in general. In fact, some earlier observations in the same process were proposed to be interpreted by the reversible scheme given in Equation 3.[16-18] However, this scheme is clearly ruled out by both thermodynamic considerations and experimental evidence about the non-equilibrium nature of the process.[14]

It seems that the rate equations given in Eq. 7 do not have in easily accessible closed-form solution, although analytical solutions for a number of similar two-step schemes have been published recently.[19] In cases like this, occasionally it is useful to seek a solution that gives one of the concentrations as a function of the other rather than time.[2] In the present case, A is only consumed in reactions and never produced, which means that its concentrations is a monotonously decreasing function of time. Therefore, it seems useful to seek the concentration of B as a function of the concentration of A. A change in the independent variable in Eq. 7 leads to the following single differential equation:

$$\frac{d[\mathrm{B}]}{d[\mathrm{A}]} = -\frac{k_1 - k_2[\mathrm{B}]}{k_1 + k_2[\mathrm{B}]} \tag{8}$$

Notably, this is still an autonomous equation (because [A] does not appear on the right side), so it is also a separable and can be solved readily by simple integration. Before presenting the full solution, it needs to be remarked that Eq. 8 clearly implies that [B] increases monotonously as a function of time (it cannot have a maximum) if its initial concentration is 0 and the concentration $[\mathrm{B}] = k_1/k_2$ is a stationary point. When $[\mathrm{B}]_0 = 0$ (note that subscript will keep refer to time, as the final goal of the calculations is obtaining the time dependence), the concentration of B can never exceed k_1/k_2. The actual solution of Eq. 8 for the concentration of species A is easily stated:

$$[\mathrm{A}] = [\mathrm{A}]_0 + [\mathrm{B}] + \frac{2k_1}{k_2} \ln\left(1 - \frac{k_2[\mathrm{B}]}{k_1}\right) \tag{9}$$

Eq. 9 connects the concentrations of [A] and [B] in a way that features only a single parameter, which is the rate constant ratio k_1/k_2. This rate constant ratio has the dimension of concentration, so it can be scaled to the initial concentration of [A]. Therefore, Figure 1 shows typical plots of [A] vs. [B] for a number of different $k_1/(k_2[\mathrm{A}]_0)$ ratios (this expression can also be thought of as a dimensionless version of the ratio of the two rate constants). The concentrations on the x and y axes are also scaled with the initial concentration of species A ($[\mathrm{A}]_0$) and $[\mathrm{B}]_0 = 0$ is assumed. Figure 1 clearly shows that the final concentration of B (i.e. the one at [A] = 0) increases as the $k_1/(k_2[\mathrm{A}]_0)$ increases. This final concentration can in fact be determined analytically as well by inverting the formula Eq. 9. For this operation, the use of the Lambert W function[19] (abbreviated simply as W here), the inverse of the xe^x function, is necessary. The final formula obtained for $[\mathrm{B}]_\infty$ is as follows:

$$[\mathrm{B}]_\infty = \frac{k_1}{k_2} + \frac{2k_1}{k_2} \mathrm{W}\left(-\frac{1}{2} e^{-k_2[\mathrm{A}]_0/(2k_1) - 1/2}\right) \tag{10}$$

The Taylor series expansion for the time dependence of the concentration of B is given in detail in the Supporting information as Derivation I. The most important results are summarized in Table 1, which shows the value and the first four derivatives of [B] at $t = 0$, which are the first five terms in the Taylor series expansion as shown in Eq. 5. Based on the data in Table 1, it can be shown that [B] is quite well approximated by the following function, which describes a pseudo-first order kinetic curve:

$$[\mathrm{B}]_{\mathrm{app}} = \frac{k_1[\mathrm{A}]_0}{k_1 + k_2[\mathrm{A}]_0}\left(1 - e^{-(k_1 + k_2[\mathrm{A}]_0)t}\right) \tag{11}$$

Table 1 also shows the first five terms of the Taylor series expansion of this exponential function. It is clear that the first four terms are identical to the first four terms of the Taylor series of the exact solution, and the difference between the fifth terms is also quite minor.

This indicates a high degree of similarity between the functions. Indeed, in this sense, Eq. 11, which describes a pseudo-first order kinetic curve with $k_{obs} = k_1 + k_2[A]_0$, is very close to the exact solution of rate equation 7. More proof of the similarity will be given in the next paragraphs.

The end point ($[B]_\infty$) of the approximation curve described by Eq. 11 is simply the multiplication term $k_1[A]_0/(k_1 + k_2[A]_0)$. Figure 2 presents a comparison between the exact end point (as given Eq. 10) and the approximated point as a function of the scaled ratio of the rate constants, $k_1/(k_2[A]_0)$. The maximum difference between the two quantities is –4.1% when $k_1/(k_2[A]_0)$ is about 0.6. 4% is not much larger than the average error of concentration determination methods, so it could be argued that the approximate formula (at least for the final concentration) works reasonably well in the entire region of possible parameter values.

Further tests of the approximations were carried out by numerically integrating the rate equations given in Eq. 7. The process of scaling was used in these calculations:[2] time was measured in units of $1/k_1$ and concentration was measured in the units of $[A]_0$. This process is analogous to using dimensionless time and concentrations, and the benefit is that the number of parameters that need to be considered in the calculations is decreased without sacrificing the general nature of the considerations. In fact, when scaled kinetic traces are calculated, the only parameter remaining in the scheme described in Eq. 6 is the dimensionless parameter $k_1/(k_2[A]_0)$, which is usually called a shape parameter.[2] During the numerical integration, 7001 points were calculated between $t = 0$ and $t = 7/k_1$ using the software Scientist.[21] This selection ensured that the concentration practically reached its final value and also gave sufficient time resolution for the simulated curves. Examples of simulated kinetic traces are given in Figure S1 of the Supporting Information. These simulated curves were then fitted to an exponential function using the non-linear least squares algorithm also in the software Scientist:

$$[B]_{fit} = Ae^{-k_{obs}t} + E \qquad (12)$$

The typical residuals as a function of time obtained in this exponential fit are shown in Figure 3 for a number of different values of the shape parameter $k_1/(k_2[A]_0)$. In agreement with the relative errors of endpoint given in Figure 2, the largest residuals occur around $k_1/(k_2[A]_0) = 0.5$, but even these relatively high residuals are quite small in absolute terms, so even in the worst case, the fit would most probably be accepted for all practical purposes. The goodness of the exponential fit was further characterized by the square root of the average of the squared residuals:

$$S = \sqrt{\frac{1}{7001}\sum_{i=1}^{7001}\left(\frac{[B]_{fit} - [B]}{[A]_0}\right)^2} \tag{13}$$

The dependence of this average deviation of the value of the shape parameter $k_1/(k_2[A]_0)$ is shown in Figure 4. For comparison, the quantity $[B]_\infty/[A]_0 \times S$ is also displayed in this graph, which gives the average deviations relative to $[B]_\infty$ rather than $[A]_0$. In both curves, a maximum is seen between 0.1 and 1, which represents the conditions under which the fit to an exponential curve is the worst. However, the maximum values of S or $[B]_\infty/[A]_0 \times S$ on the y axis are lower than 10^{-2}, which clearly show that the average deviation of the exponential fits are below 1% under all conditions.

Figure 4 also displays k_{obs}/k_1 as a function of $k_1/(k_2[A]_0)$. The points in this series show the fitted k_{obs} values, whereas the straight line corresponds to $k_{obs} = k_1 + k_2[A]_0$, which is expected based on the approximation given in Eq. 11. In this case, the agreement of the individual points and the straight lines show how good the approximation $k_1 + k_2[A]_0$ for the fitted pseudo-first order rate constant (k_{obs}) is. Remarkably, there is a noticeable deviation in the region $0.01 < k_1/(k_2[A]_0) < 1$, which may occasionally exceed 10% of the k_{obs} value. This observation, considered together with the previous conclusions, demonstrates that the kinetic scheme given in Eq. 6 always gives rise to curves that cannot be distinguished experimentally

from an exponential curve, but the approximation of the pseudo-first order rate constant given in Eq. 11 may be more than 10% off under certain conditions.

Some further indicators of the acceptability of the pseudo-first order fit were also monitored. These are the fitted value of the concentration of B at the beginning and end of the curve.

$$[\mathrm{B}]_{\mathrm{fit}}(t=0) = A + E \qquad [\mathrm{B}]_{\mathrm{fit}}(t=\infty) = E \tag{14}$$

Figures S2 and S3 in the Supporting Information shows the difference between these fitted values from the exactly known values of $[\mathrm{B}]_0$ (= 0) and $[\mathrm{B}]_\infty$ as a function of $k_1/(k_2[\mathrm{A}]_0)$. It is seen that $[\mathrm{B}]_{\mathrm{fit}}(t = \infty)$ is always in excellent agreement with $[\mathrm{B}]_\infty$ and the maximum difference between $[\mathrm{B}]_{\mathrm{fit}}(t = \infty)$ and is about 5% of the value of $[\mathrm{A}]_0$. Again, these values and graphs imply that describing the concentration change of B as an exponential curve is a highly acceptable assumption given the usual experimental errors of concentration determination. Finally, Figure S4 in the Supporting Information compares the curves calculated by numerical integration, by the approximation formula in Eq. 11 and by the exponential fitting under on particular set of conditions. As expected Eq. 11 gives a very good approximation of the beginning of the actual concentration of B, but estimating the end point is less reliable by this formula.

It is also notable that the co-incidence of the Taylor series expansion only occurs for the concentration of B. For the concentration of A and P, the deviation from a pseudo-first order curve is much larger and would probably be detectable in experiments as well. Indeed, in the cited experimental example, the spectrophotometric detection were selective to species B (HOSCN).[14,16-18]

Now turning to the possibility of misidentifying scheme: suppose that the process A $\rightarrow$ B in the scheme of Eq. 6 is in fact a pseudo-first order process, which involves the reaction of A with an excess reagent R. In this case, the first order rate constant of the first process is

given as $k_1 = k_{2nd}[R]$. Therefore, a coupled consecutive reaction would be expected to give an experimental curve that is very close to exponential with $k_{obs} = k_{2nd}[R] + k_2[A]_0$. Unless the experimenter is careful enough to design experiments with varied initial concentrations of A (which is the deficiency reagent!), it is very easy to reach the erroneous conclusion that the reaction is reversible with a forward rate constant of k_{2nd} and a reverse rate constant of $k_2[A]_0$. In fact, an analysis of the final concentrations of B could further strengthen this conclusion, as the approximation formula in Eq. 11 shows that such an evaluation would yield practically the same equilibrium constant for the supposed equilibrium process ($k_{2nd}/k_2[A]_0$) as the ratio of the forward and reverse rate constants.

Parallel Reaction and Decay. The second model we consider in this article involves an adduct formation and parallel decay of one of the reactants. The earlier published experimental example is the ligand substitution reaction between the vitamin B_{12} derivative aquacobalamin/hydroxycobalamin (H_2OCbl^+/HOCbl) and secondary amine NONOates, R2N-NONOates.[12-13] The latter reagents (the "ligand") was confirmed to undergo a decomposition, a process that was termed "silent killer" later[13] because it does not change the observed absorbance values. The scheme itself is as follows:

$$\begin{aligned} &C \xrightarrow{k_3} E \\ &D + C \xrightarrow{k_4} F \end{aligned} \tag{15}$$

In the experimental example,[12] C was the ligand, whereas D was the metal-containing species H_2OCbl^+/HOCbl. The rate equations based on this scheme are given as follows:

$$\begin{aligned} &\frac{d[C]}{dt} = -k_3[C] - k_4[C][D] \\ &\frac{d[D]}{dt} = -k_4[C][D] \end{aligned} \tag{16}$$

Again, this is a kinetic model for which a closed form analytical solution has not been found yet.[13] However, an analysis similar to the one presented for the previous scheme is possible here as well. It is clear that both the concentrations of C and D are monotonous functions of time, so it may be meaningful to seek the concentration of D as a function of the concentration of C. The differential equation obtained is as follows:

$$\frac{d[\mathrm{D}]}{d[\mathrm{C}]} = \frac{k_4[\mathrm{D}]}{k_3 + k_4[\mathrm{D}]} \tag{17}$$

Again, this is an autonomous and separable differential equation, whose solution is particularly easily sated for [C]:

$$[\mathrm{C}] = [\mathrm{C}]_0 + [\mathrm{D}] - [\mathrm{D}]_0 + \frac{k_3}{k_4} \ln \frac{[\mathrm{D}]}{[\mathrm{D}]_0} \tag{18}$$

Scaling of the kinetic curves is possible similarly to the previous case. A convenient time unit is $1/(k_4[\mathrm{C}]_0)$, whereas a convenient concentration scale is provided by $[\mathrm{D}]_0$. Another possible time unit would be $1/k_3$, but this is not particularly useful in this case as the concentration of D will be of the main interest. This scaling now leaves two shape parameters: $[\mathrm{C}]_0/[\mathrm{D}]_0$ and $k_4[\mathrm{D}]_0/k_3$. Because the original experimental example used a large excess of the ligand (C here), and indeed the most interesting features of this system are demonstrated at an initial excess of C, most analysis in this paper was done for $[\mathrm{C}]_0/[\mathrm{D}]_0 \geq 1$. It is notable that the concentration of C always goes down to zero by the end of the process because of its first-order decay, but it is possible for some D to remain intact by the end. This can also be clearly observed in Figure 5, where a few typical curves for [C] vs. [D] are given under a number of different parameter values.

Again, the final concentration of D is of interest in this system. The exact analytical formula for this can be given similarly to Eq. 10, using the Lambert W function:

$$[\mathrm{D}]_\infty = \frac{k_3}{k_4} \mathrm{W}\left(\frac{k_4[\mathrm{D}]_0}{k_3} e^{k_4([\mathrm{D}]_0 - [\mathrm{C}]_0)/k_3} \right) \tag{19}$$

Similarly to the previous scheme, the rate equations shown in Eq. 16 facilitate the calculation of the derivatives of the concentration of D at $t = 0$. This is detailed in the Supporting information as Derivation 2, and the most important final results, which are the first four terms of the Taylor series expansion, are displayed in Table 2. At this time, Table 2 also shows the first four terms of the Taylor series expansion of the following exponential function:

$$[\mathrm{D}]_{\mathrm{app}} = \frac{k_4[\mathrm{C}]_0[\mathrm{D}]_0}{k_3 + k_4[\mathrm{C}]_0 + k_4[\mathrm{D}]_0} e^{-(k_3+k_4[\mathrm{C}]_0+k_4[\mathrm{D}]_0)t} + \frac{k_3[\mathrm{D}]_0 + k_4[\mathrm{D}]_0^{\,2}}{k_3 + k_4[\mathrm{C}]_0 + k_4[\mathrm{D}]_0} \tag{20}$$

It can be seen that the first three terms in the Taylor series expansion of the two functions are identical, and the difference between the fourth terms is also minor. Therefore, analogously to the previous case, this implies a considerably degree of similarity of these functions. This is in agreement with the conclusions of a previous paper that was based on the numerical integration of the rate equation in Eq. 16. REF!!! However, the fact that the first difference is seen in the fourth term rather than the fifth indicates that the similarity may be more limited than in the case of coupled consecutive reactions.

As in the previous case, it is instructive to compare the actual exact final concentrations of D (as shown in Eq. 19) with the approximation values provided by Eq. 20. This is done in Figure 6. Because the system has two shape parameters, the dependence with respect to both of these parameters should be studied. Figure 6 displays what is in effect a contour plot: the two shape parameters are given on the two axes, and four regions are identified in this graph where the difference between the exact and $[\mathrm{D}]_\infty$ values is smaller than 0.1%; between 0.1%, and 1%; between 1% and 10%; and greater than 10%. The region where the deviation exceeds 10% represents conditions where Eq. 20 does not approximate the exact solution of rate equation 16 well. These conditions are characterized by $1 < k_4[\mathrm{C}]_0/k_3 < 10$ or the simultaneous inequities $[\mathrm{C}]_0/[\mathrm{D}]_0 < 10$ and $k_3/(k_4[\mathrm{D}]_0) < 1$.

To obtain more information about the usefulness of the approximation formula in Eq. 20, detailed analysis similar to the one described for the previous case was carried out. The software Scientist[21] was first used to solve Eq. 16 numerically. Again 7001 points were generated in the time domain between $t = 0$ and $t = 10/(k_4[C]_0)$. Some simulated kinetic traces are show in Figure S5 of the Supporting Information. The kinetic curves generated by numerical integration were then fitted to an exponential function using the non-linear least squares algorithm also in the software Scientist:

$$[D]_{fit} = Ae^{-k_{obs}t} + E \tag{21}$$

Figure 7 gives the residuals of this fit for a number of selected cases. These residuals are in general somewhat larger than those in the previous case, which is in agreement with the fact that the Taylor series expansions compare better in Table 1 than in Table 2. However, the maximum residuals are still about 5%, which is not very large. The tendency of the residuals also resembles the theoretical residual shapes of second order processes, which are studied at large excess at one of the reagents and evaluated by exponential fitting.[2] The average deviation, S_D, of the fit was defined in a way that is somewhat different from the previous case:

$$S_D = \sqrt{\frac{1}{7001}\sum_{i=1}^{7001}\left(\frac{[D]_{fit} - [D]}{[D]_0}\right)^2} \tag{22}$$

The dependence of S_D on the two shape parameters is shown in Figure 8. This plot gives S_D as a function of $k_4[D]_0/k_3$ for eight different values of $[C]_0/[D]_0$. All eight curves show a maximum, which represent about 1% average deviation for the entire fit, roughly at $k_4[D]_0/k_3 = 3[D]_0/[C]_0$, i.e. $k_4[C]_0/k_3 = 3$. The region of maximum average fit error is exactly the same as the center of the area featuring the highest error in Figure 6. It is to be noted that even though Figure 6 displays an estimation error for the final concentration of D in excess of 10% (which is probably unacceptable), Figure 8 clearly shows that the average deviation of the

exponential fit are about 1% even in the worst cases. The very similar shape of the curves in Figure 8 and the symmetry of Figure 6 also imply that the goodness of the approximation given in Eq. 20 can be assessed based solely on the value of the combination parameter $k_4[C]_0/k_3$. The approximation works worst when $k_4[C]_0/k_3$ is about 3, and gets gradually better as the parameters move from this condition.

The analysis shows that this scheme of parallel reaction and decay also has a high potential to be confused with the scheme shown in Eq. 3. When C is used in high excess over D, curves very close to pseudo-first order are detected with $k_{obs} = k_3 + k_4[C]_0 + k_4[D]_0$. Because of the large excess of reagent C, the dependence of k_{obs} on $[D]_0$ (again, the deficiency reagent) is very difficult to pick up experimentally. Therefore, k_4 would be interpreted as the rate constant of the reaction between C and D, which is correct, but the k_3 term would be misidentified as the first order rate constant of the reverse reaction.

Conclusion

The results in this paper shows that the two kinetic schemes considered here, coupled consecutive processes and parallel reaction and decay, are indeed very difficult to distinguish from the reversible second order process based solely on kinetic data. The analysis here shows that designing experiments with varied concentrations of the deficiency reagent might serve as a clue in the first case, whereas the deviation from the pseudo-first order behavior at $k_4[C]_0/k_3 = 3$ in the second case might be substantial enough for experimental detection. In both cases, a direct equilibrium test of the reversible nature of the process, i.e. trying to shift the equilibrium back toward the initial substances, would clearly challenge the validity of the scheme given in Eq. 3. Therefore, it is concluded that such equilibrium tests should be carried out in every case to prove the reversible nature of the process.

References

1. Espenson, J. H. in *Chemical Kinetics and Reaction Mechanisms*. 2nd ed., McGraw-Hill, New York, **1995**.

2. Lente, G. in *Deterministic Kinetics in Chemistry and Systems Biology*, Springer, New York, Heidelberg, Dordrecht, London (2015)

3. Armstrong, F. A.; Henderson, R. A.; Sykes, A. G. *J. Am. Chem. Soc.* **1973**, *102*, 6545-6551.

4. Vanderheiden, D. B.; King, E. L. *J. Am. Chem. Soc.* **1973**, *95*, 3860-3866.

5. Marty, W.; Espenson, J. H. *Inorg. Chem.* **1979**, *5*, 1246-1250.

6. Buckingham, D. A.; Francis, D. J.; Sargeson, A. M. *Inorg. Chem.* **1974**, *13,* 2630-2639.

7. Kathirgamanathan, P.; Soares, A. B.; Richens, D. T.; Sykes, A. G. *Inorg. Chem.* **1985**, *24*, 2950-2954.

8. Algarra, A. G.; Fernandez-Trujillo, M. J.; Basallote, M. G. *Chem. Eur. J.* **2012**, *18*, 5036-5046.

9. Basallote, M. G.; Fernández-Trujillo, M. J.; Pino-Chamorro, J. Á.; Beltrán, T. F.; Corao, C.; Llusar, R.; Sokolov, M.; Vicent, C. *Inorg. Chem.* **2012**, *51*, 6794-6802.

10. Castillo, C. E.; González-García, J.; Llinares, J. M.; Máñez, M. A.; Jimenez, H. R.; García-España, E.; Basallote, M. G. *Dalton Trans.* **2013**, *42*, 6131-6141.

11. Pino-Chamorro, J. Á.; Algarra, A. G.; Fernández-Trujillo, M. J.; Hernández-Molina, R.; Basallote, M. G. *Inorg. Chem.* **2013**, *52*, 14334-14342.

12. Hassanin, A. A.; Hannibal, L.; Jacobsen, D. W.; M. F. El-Shahat, M. F.; Hamza, M. S. A.; Brasch, N. E. *Angew. Chem., Int. Ed.*, **2009**, *48*, 8909-8913.

13. Conrad, A. R.; Hassanin, H. A.; Tubergen, M. J.; Fábián, I.; Brasch, N. E. *New. J. Chem.* **2012**, *36*, 1408-1412.

14. Kalmár, J.; Lente, G.; Fábián, I. *Inorg. Chem.* **2013**, *52*, 2150-2156.

15. http://mathworld.wolfram.com/TaylorSeries.html and

http://mathworld.wolfram.com/TaylorsTheorem.html (last accessed: 19 March 2015)

16. Smith, R. H.; Wilson, I. R. *Aust. J. Chem.* **1966**, *19*, 1357-1363.

17. Smith, R. H.; Wilson, I. R. *Aust. J. Chem.* **1966**, *19*, 1365-1375.

18. Smith, R. H.; Wilson, I. R. *Aust. J. Chem.* **1967**, *20*, 1353-1366.

19. Lente, G. *J. Math. Chem.* **2015**, *53*, 1172-1183.

20. http://mathworld.wolfram.com/LambertW-Function.html (last accessed: 19 March 2015)

21. SCIENTIST, version 2.0, Micromath Software, Salt Lake City, UT, USA, 1995.

.

Table 1 Terms in the Taylor series expansion for the case of coupled consecutive reaction shown in Eq. 6.

at $t = 0$	From rate equation (Eq. 7)	From approximation (Eq. 11)	Difference
value	0	0	0
1st derivative	$k_1[\mathrm{A}]_0$	$k_1[\mathrm{A}]_0$	0
2nd derivative	$-k_1^2[\mathrm{A}]_0 - k_1k_2[\mathrm{A}]_0^2$	$-k_1[\mathrm{A}]_0(k_1 + k_2[\mathrm{A}]_0)$	0
3rd derivative	$k_1^3[\mathrm{A}]_0 + 2k_1^2k_2[\mathrm{A}]_0^2 + k_1k_2^2[\mathrm{A}]_0^3$	$k_1[\mathrm{A}]_0(k_1 + k_2[\mathrm{A}]_0)^2$	0
4th derivative	$-k_1^4[\mathrm{A}]_0 - 3k_1^3k_2[\mathrm{A}]_0^2 - k_1^2k_2^2[\mathrm{A}]_0^3 - k_1k_2^3[\mathrm{A}]_0^4$	$-k_1[\mathrm{A}]_0(k_1 + k_2[\mathrm{A}]_0)^3$	$2k_1^2k_2^2[\mathrm{A}]_0^3$

Table 2 Terms in the Taylor series expansion for the case of parallel reaction and decay shown in Eq. 15.

at $t = 0$	From rate equation (Eq. 16)	From approximation (Eq. 20)	Difference
value	$[D]_0$	$[D]_0$	0
1st derivative	$-k_4[C]_0[D]_0$	$-k_4[C]_0[D]_0$	0
2nd derivative	$k_3k_4[C]_0[D]_0 + k_4^2[C]_0[D]_0^2 + k_4^2[C]_0^2[D]_0$	$k_4[C]_0[D]_0(k_3 + k_4[C]_0 + k_4[D]_0)$	0
3rd derivative	$-k_3^2k_4[C]_0[D]_0 - 2k_3k_4^2[C]_0[D]_0^2 - 3k_3k_4^2[C]_0^2[D]_0$ $-k_4^3[C]_0[D]_0^3 - 4k_4^3[C]_0^2[D]_0^2 - k_4^3[C]_0^3[D]_0$	$-k_4[C]_0[D]_0(k_3 + k_4[C]_0 + k_4[D]_0)^2$	$k_3k_4^2[C]_0^2[D]_0 +$ $2k_4^3[C]_0^2[D]_0^2$

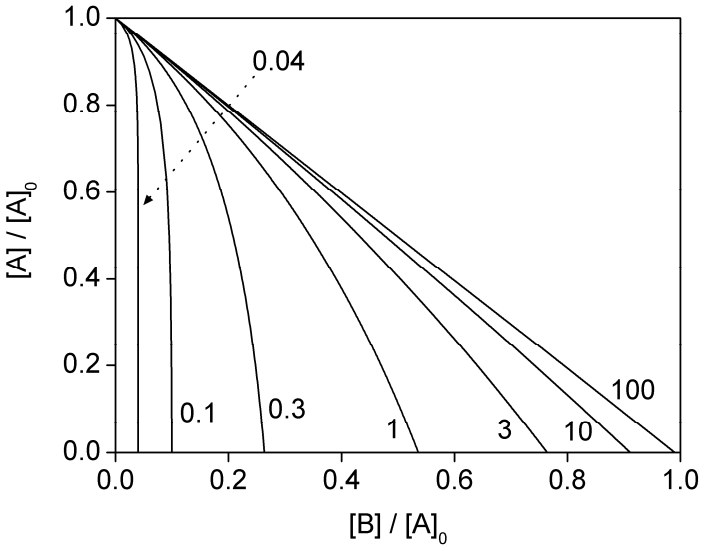

Figure 1 Scaled concentration of species A as a function of the scaled concentration of species B in the scheme shown in Eq. 6. The numbers in the graph display the $k_1/(k_2[A]_0)$ parameter values for the particular curves.

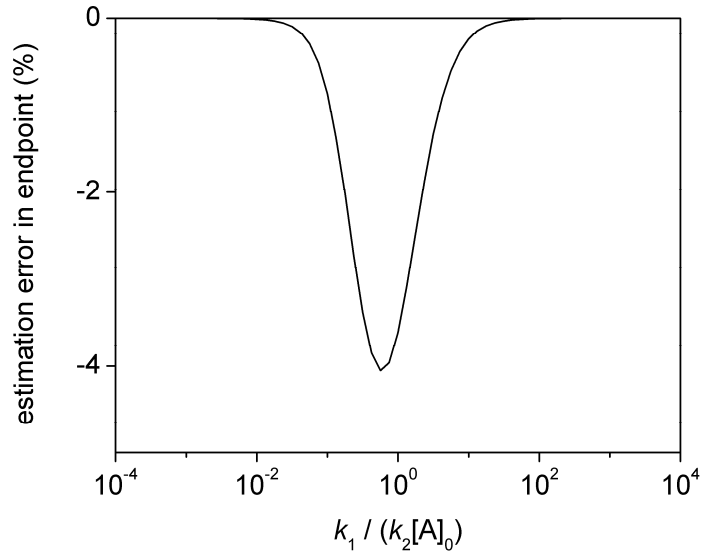

Figure 2 Relative estimation error of the end point by the approximate formula in Eq. 11 for the scheme shown in Eq. 6

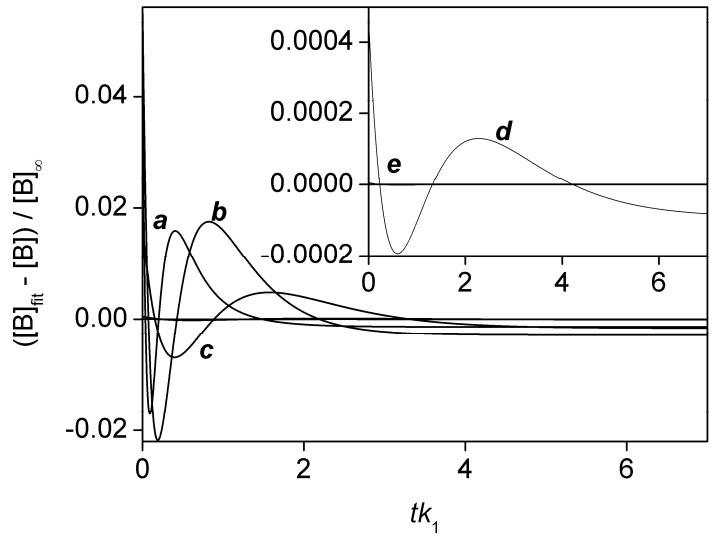

Figure 3 Residuals as a function of scaled time from the fitting of Eq. 12 for a few sample curves simulated based on the scheme in Eq. 6. Curve ***e*** is indistinguishable from the $y = 0$ line in the inset on this scale. Parameter values: $k_1/(k_2[A]_0) =$ 0.1 (***a***), 0.222 (***b***), 1 (***c***), 10 (***d***), 100 (***e***).

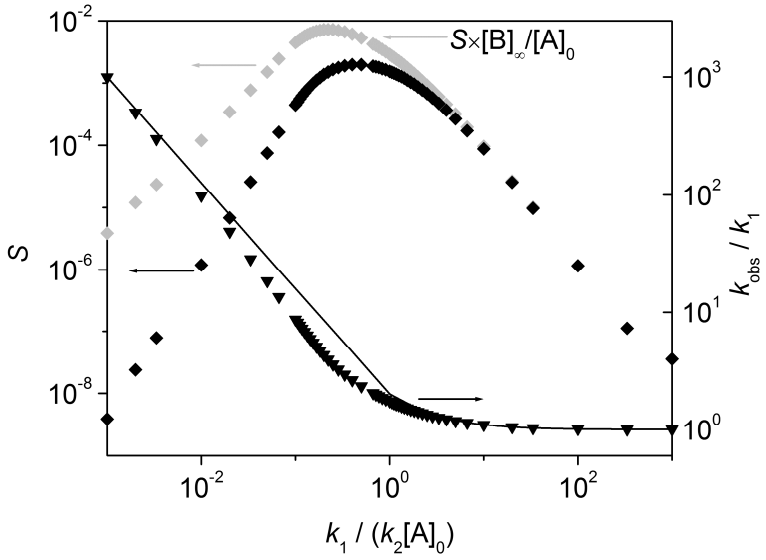

Figure 4 Average relative deviations from the fitting of Eq. 12 and the estimated k_{obs} values as a function of shape parameter $k_1/(k_2[A]_0)$ for the scheme shown in Eq. 6.

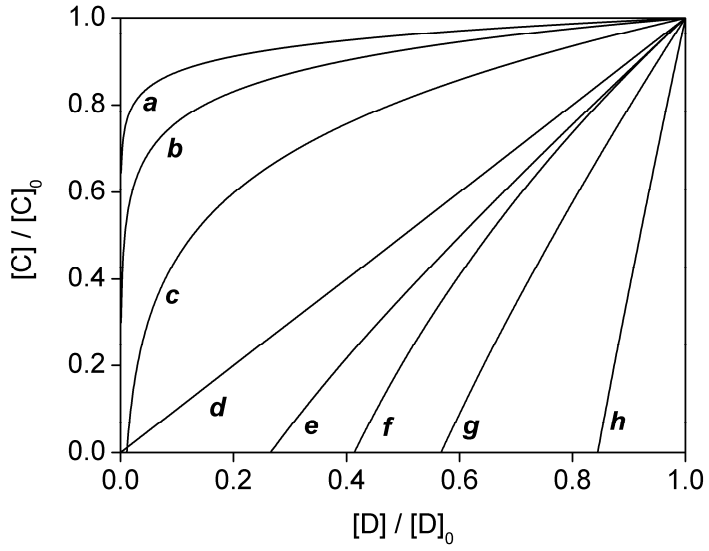

Figure 5 Scaled concentration of species A as a function of the scaled concentration of species B in the scheme shown in Eq. 15. Particular values for the two shape parameters of the curves drawn: $[C]_0/[D]_0 =$ 100 (***a***), 100 (***b***), 10 (***c***), 1 (***d***), 1 (***e***), 5 (***f***), 1 (***g***), 1 (***h***); $k_4[D]_0/k_3 =$ 0.2 (***a***), 0.1 (***b***), 0.5 (***c***), 1000 (***d***), 5 (***e***), 0.2 (***f***), 1 (***g***), 0.2 (***h***).

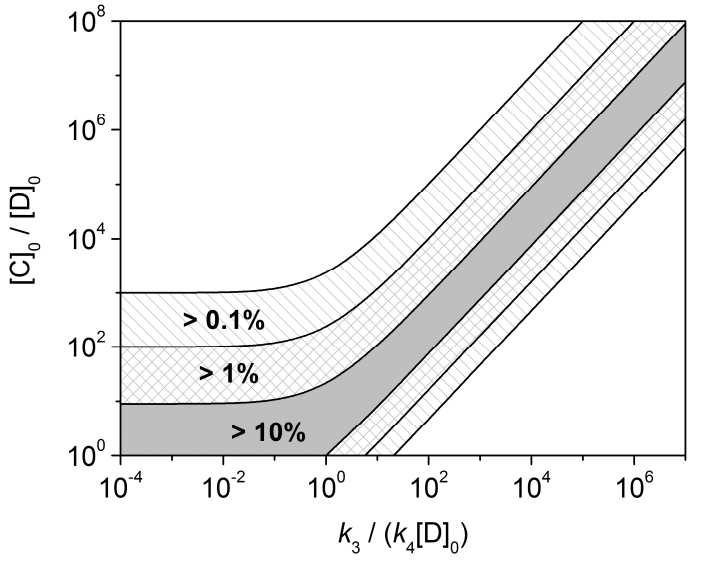

Figure 6 Relative estimation error of the end point by the approximate formula in Eq. 20 as a function of shape parameters $[C]_0/[D]_0$ and $k_3/(k_4[D]_0)$ for the scheme shown in Eq. 15.

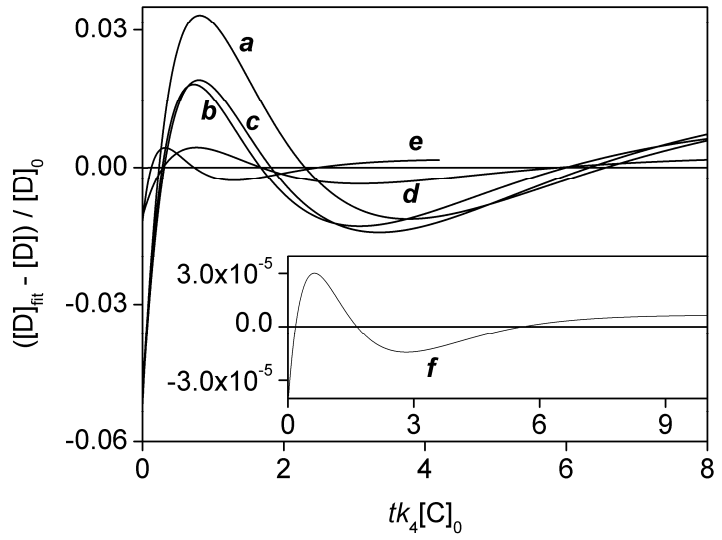

Figure 7 Residuals as a function of scaled time from the fitting of Eq. 21 for a few sample curves simulated based on the scheme in Eq. 15. Parameter values: $[C]_0/[D]_0 =$ 10 (***a***), 100 (***b***), 10^4 (***c***), 1000 (***d***), 300 (***e***), 3000 (***f***); $k_4[D]_0/k_3 =$ 0.3 (***a***), 0.02 (***b***), 3×10^{-4} (***c***), 0.02 (***d***), 0.002 (***e***), 3 (***f***).

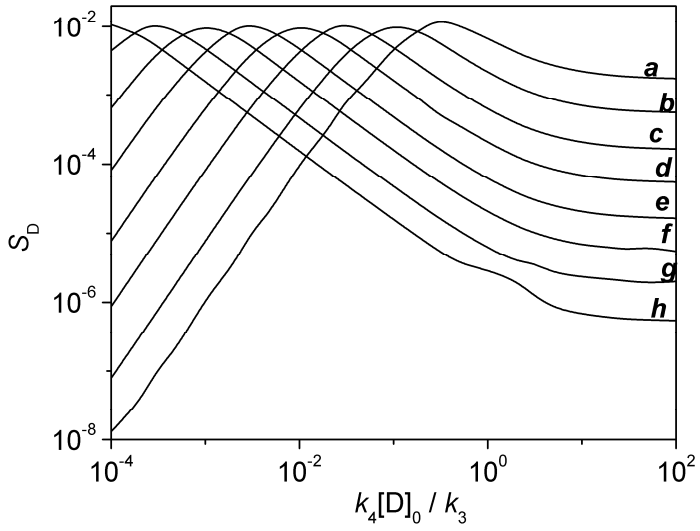

Figure 8 Average relative deviations from the fit of Eq. 21 as a function of shape parameter $k_1/(k_2[A]_0)$ for the scheme shown in Eq. 15. Parameter values: $[C]_0/[D]_0 =$ 10 (***a***), 30 (***b***), 100 (***c***), 300 (***d***), 1000 (***e***), 3000 (***f***), 1×10^4 (***g***), 3×10^4 (***h***).

Supporting Information

for

Detailed Kinetic Analysis of Two Reaction Schemes that are Easily Confused with the Reversible First Order Reaction

Ágnes Balogh[a], Gábor Lente[a]*, József Kalmár[b], István Fábián[a]

[a] Department of Inorganic and Analytical Chemistry, University of Debrecen, P.O.B. 21, Debrecen, H-4010 Hungary. Fax: + 36 52 489667; Tel: + 36 52 512900/22373; E-mail: lenteg@science.unideb.hu

[b] MTA-DE Homogeneous Catalysis and Reaction Mechanisms Research Group, P.O.B. 21, Debrecen, H-4010 Hungary

Derivation I.

The mechanism defines the following set of differential equations (using the simple notations, x = [A], y = [B], $a = k_1$ and $b = k_2$ with the number of dots above the variable indicating the number of derivations as a function of time):

$$\dot{x} = -ax - bxy$$

$$\dot{y} = ax - bxy$$

Further differentials can be given successively:

$$\ddot{x} = -a\dot{x} - b\dot{x}y - bx\dot{y}$$

$$\ddot{y} = a\dot{x} - b\dot{x}y - bx\dot{y}$$

$$\dddot{x} = -a\ddot{x} - b\ddot{x}y - 2b\dot{x}\dot{y} - bx\ddot{y}$$

$$\dddot{y} = a\ddot{x} - b\ddot{x}y - 2b\dot{x}\dot{y} - bx\ddot{y}$$

$$\ddddot{x} = -a\dddot{x} - b\dddot{x}y - 3b\ddot{x}\dot{y} - 3b\dot{x}\ddot{y} - bx\dddot{y}$$

$$\ddddot{y} = a\dddot{x} - b\dddot{x}y - 3b\ddot{x}\dot{y} - 3b\dot{x}\ddot{y} - bx\dddot{y}$$

The derivatives at $t = 0$ can consequently be given:

$$\dot{x}_0 = -ax_0 - bx_0 0 = -ax_0$$

$$\dot{y}_0 = ax_0 - bx_0 0 = ax_0$$

$$\ddot{x}_0 = -a(-ax_0) - bx_0(ax_0) = a^2x_0 - abx_0^2$$

$$\ddot{y}_0 = a(-ax_0) - bx_0(ax_0) = -a^2x_0 - abx_0^2$$

$$\dddot{x}_0 = -a(a^2x_0 - abx_0^2) - 2b(-ax_0)(ax_0) - bx_0(-a^2x_0 - abx_0^2) = -a^3x_0 + 4a^2bx_0^2 + ab^2x_0^3$$

$$\dddot{y}_0 = a(a^2x_0 - abx_0^2) - 2b(-ax_0)(ax_0) - bx_0(-a^2x_0 - abx_0^2) = a^3x_0 + 2a^2bx_0^2 + ab^2x_0^3$$

$$\overset{\bullet\bullet\bullet\bullet}{x}_0 = -a(-a^3x_0 + 4a^2bx_0^2 + ab^2x_0^3) - 3b(a^2x_0 - abx_0^2)(ax_0) - 3b(-ax_0)(-a^2x_0 - abx_0^2) -$$

$$-bx_0(a^3x_0 + 2a^2bx_0^2 + ab^2x_0^3) = a^4x_0 - 11a^3bx_0^2 - 3a^2b^2x_0^3 - ab^3x_0^4$$

$$\overset{\bullet\bullet\bullet\bullet}{y}_0 = a(-a^3x_0 + 4a^2bx_0^2 + ab^2x_0^3) - 3b(a^2x_0 - abx_0^2)(ax_0) - 3b(-ax_0)(-a^2x_0 - abx_0^2) -$$

$$-bx_0(a^3x_0 + 2a^2bx_0^2 + ab^2x_0^3) = -a^4x_0 - 3a^3bx_0^2 - a^2b^2x_0^3 - ab^3x_0^4$$

The values of the derivatives at time zero were directly inserted into Table 1.

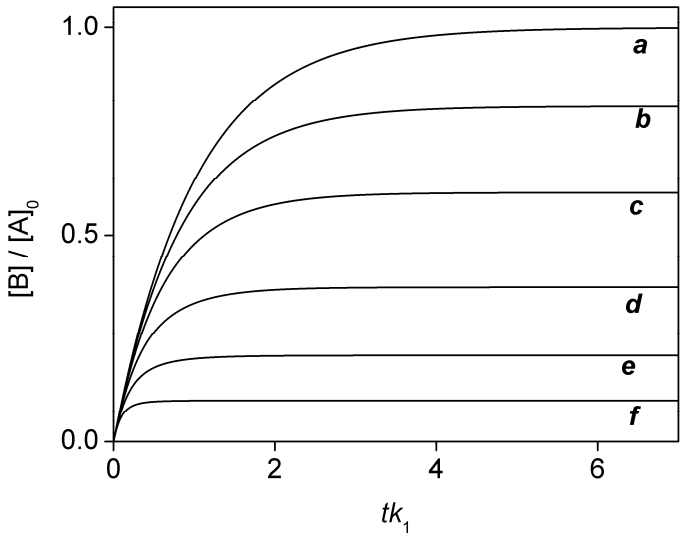

Figure S1 Simulated scaled kinetic traces based on the scheme shown in Eq. 6. Parameter values: $k_1/(k_2[A]_0) = 10^4$ (***a***), 4500 (***b***), 2000 (***c***), 750 (***d***), 250 (***e***), 1 (***f***).

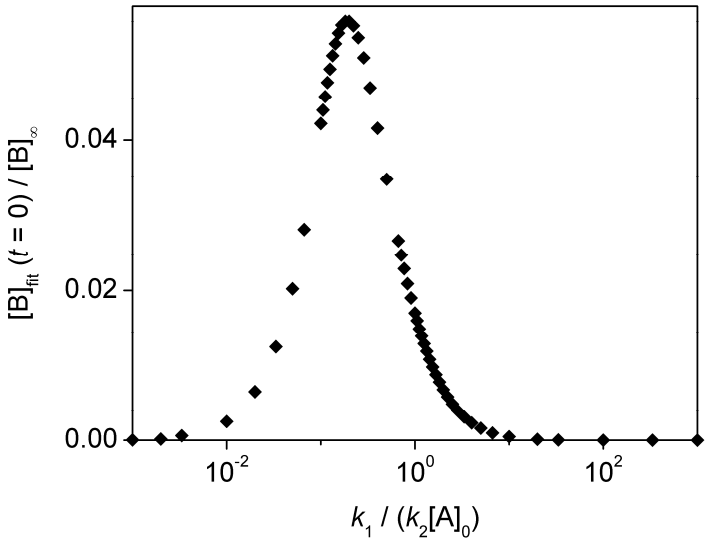

Figure S2 Relative error of the estimation of the initial value of [B] by Eq. 12 as a function of $k_1/(k_2[A]_0)$ for the scheme shown in Eq. 6.

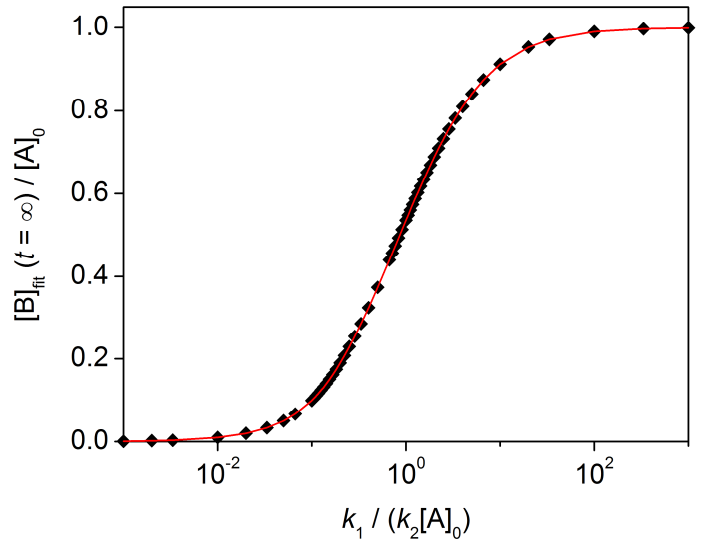

Figure S3 Relative error of the estimation of $[B]_\infty$ by Eq. 12 as a function of $k_1/(k_2[A]_0)$ for the scheme shown in Eq. 6.

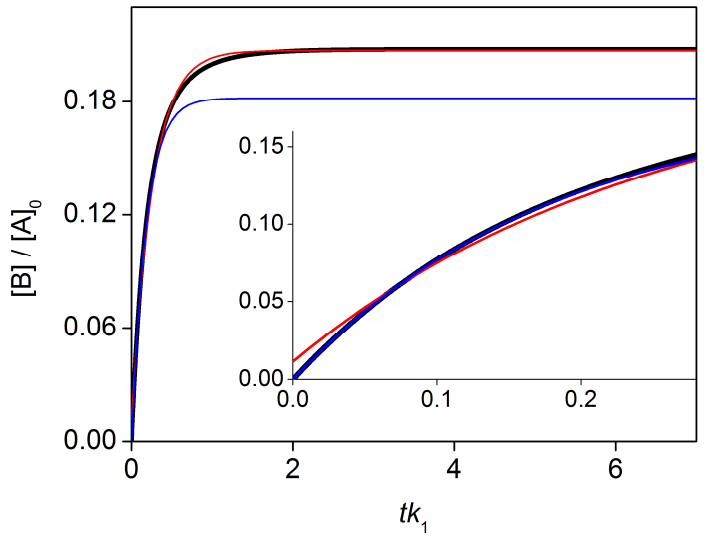

Figure S4 Comparison of the time dependence of the concentration of B calculated by the direct simulation based on the scheme shown in Eq. 6 (black curve), the approximation equation in Eq. 11 (blue line) and the exponential fit of Eq. 12 (red line).

Derivation II.

This time, the mechanism defines the following set of differential equations (using the simple notations, x = [C], y = [D], $a = k_3$ and $b = k_4$ with the number of dots above the variable indicating the number of derivations as a function of time):

$$\dot{x} = -ax - bxy$$

$$\dot{y} = -bxy$$

Further differentials can be given successively:

$$\ddot{x} = -a\dot{x} - b\dot{x}y - bx\dot{y}$$

$$\ddot{y} = -b\dot{x}y - bx\dot{y}$$

$$\dddot{x} = -a\ddot{x} - b\ddot{x}y - 2b\dot{x}\dot{y} - bx\ddot{y}$$

$$\dddot{y} = -b\ddot{x}y - 2b\dot{x}\dot{y} - bx\ddot{y}$$

The derivatives at $t = 0$ can consequently be given:

$$\dot{x}_0 = -ax_0 - bx_0y_0 = -ax_0 - bx_0y_0$$

$$\dot{y}_0 = -bx_0y_0 = -bx_0y_0$$

$$\ddot{x}_0 = -a(-ax_0 - bx_0y_0) - b(-ax_0 - bx_0y_0)y_0 - bx_0(-bx_0y_0) =$$
$$= a^2x_0 + 2abx_0y_0 + b^2x_0{y_0}^2 + b^2{x_0}^2y_0$$

$$\ddot{y}_0 = -b(-ax_0 - bx_0y_0)y_0 - bx_0(-bx_0y) = abx_0y_0 + b^2x_0{y_0}^2 + b^2{x_0}^2y_0$$

$$\dddot{x}_0 = -a(a^2 x_0 + 2abx_0 y_0 + b^2 x_0 {y_0}^2 + b^2 {x_0}^2 y_0) - b(a^2 x_0 + 2abx_0 y_0 + b^2 x_0 {y_0}^2 + b^2 {x_0}^2 y_0) y_0$$
$$-2b(-ax_0 - bx_0 y_0)(-bx_0 y_0) - bx_0 (abx_0 y_0 + b^2 x_0 {y_0}^2 + b^2 {x_0}^2 y_0) =$$
$$= -a^3 x_0 - 3a^2 bx_0 y_0 - 3ab^2 x_0 {y_0}^2 - 4ab^2 {x_0}^2 y_0 - b^3 x_0 {y_0}^3 - 4b^3 {x_0}^2 {y_0}^2 - b^3 {x_0}^3 y_0$$

$$\dddot{y}_0 = -b(a^2 x_0 + 2abx_0 y_0 + b^2 x_0 {y_0}^2 + b^2 {x_0}^2 y_0) y_0 - 2b(-ax_0 - bx_0 y_0)(-bx_0 y_0)$$
$$-bx_0 (abx_0 y_0 + b^2 x_0 {y_0}^2 + b^2 {x_0}^2 y_0) =$$
$$= -a^2 bx_0 y_0 - 2ab^2 x_0 {y_0}^2 - 3ab^2 {x_0}^2 y_0 - b^3 x_0 {y_0}^3 - 4b^3 {x_0}^2 {y_0}^2 - b^3 {x_0}^3 y_0$$

The values of the derivatives at time zero were directly inserted into Table 2.

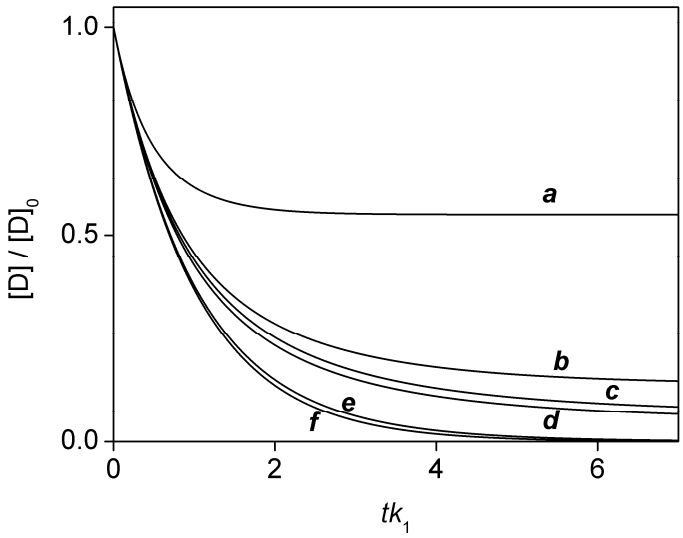

Figure S5 Simulated scaled kinetic traces based on the scheme shown in Eq. 15. Parameter values: $[C]_0/[D]_0 =$ 300 (***a***), 100 (***b***), 10 (***c***), 10^4 (***d***), 1000 (***e***), 3000 (***f***); $k_4[D]_0/k_3 =$ 0.002 (***a***), 0.02 (***b***), 0.3 (***c***), 3×10^{-4} (***d***), 0.02 (***e***), 3 (***f***).