concvar: A program solving time-dependent concentration variation for complex reactions

Tian Lu

Beijing Kein Research Center for Natural Sciences, Beijing 100024, P. R. China (http://www.keinsci.com)

E-mail: sobereva@sina.com

Abstract: In this article we introduce concvar program, which aims at studying variation of concentrations of various chemical substances in complex reactions over reaction time by numerically solving chemical kinetics equations. Linking relationship and free energies of all minima and transition states in energy profile of the studied reaction are needed as input. Concvar is easy to use, flexible, and significant for both practical research and physical chemistry teaching. Executable files and a detailed manual of concvar can be freely downloaded at http://sobereva.com/soft/concvar.

1. Introduction

A complex chemical reaction is composed of multiple elementary reactions, and there are multiple minima on the overall energy profile, which correspond to various chemical substances (reactant, intermediates and product). It is not only theoretical interesting, but also of great practical significance, to investigate how concentration of various substances changes over reaction time. To obtain such information, reaction rate constant (k) for each elementary reaction needs to be known, which can be readily obtained based on free energies of minima and the transition states connecting them in combination with transition state theory (TST). At present, there are mature and reliable quantum chemical methods to calculate free energy accurately.[1] On the other hand, a set of chemical kinetic equations need to be solved. For complex reactions involving multiple elementary reactions, in order to derive time-dependent concentration variations, the kinetic equations have to be solved numerically; in other words, numerical simulation of reaction progress is needed.

As far as we know, currently there is no easy-to-use, flexible and freely available code to simulating concentration variation of complex reactions by simply inputting free energies and simulation conditions. To fill this evident gap, we developed a code named concvar, which will be described in this article. Theoretical chemists can easily use this program to study many problems such as reaction half-life, product distribution, and rate-determining states. Physical chemistry teachers may use this program to make students better understand the principles of chemical kinetics.

In the rest of this article, we will first outline the theory involved in the simulation of concvar, then characters of concvar will be briefly described. After that, in Section 4 a few examples are given to illustrate the capability and value of concvar. Finally, in Section 5 we conclude the article.

2. Theory

During the simulation of concvar program, concentration (c) of a minimum, for

example system i, changes according to the following equation at each simulation step:

$$c_i(t + \Delta t) = c_i(t) + \left[r_i^{\text{react}}(t) + r_i^{\text{gen}} - r_i^{\text{sink}} \right] \Delta t \tag{1}$$

where Δt is simulation step size, t and $t + \Delta t$ correspond to time of current and next step, respectively. The initial concentration $c(0)$ of each minimum is defined in input file.

The r_i in above equation is rate of concentration variation of substance i. Three types of sources are considered by concvar. The first one comes from occurrence of relevant elementary reactions

$$r_i^{\text{react}}(t) = \sum_j \left[k_{j \to i} c_j(t) - k_{i \to j} c_i(t) \right] \tag{2}$$

where j loops over all minima connected to minimum i. $k_{j \to i}$ corresponds to reaction rate constant from other adjacent substance j to i, which increases concentration of i, while $k_{i \to j}$ corresponds to reaction rate constant from i to other adjacent substance j, which decreases concentration of i.

The k in Eq. 2 is normally obtained via transition state theory (TST)[2] based on free energy barrier at standard state concentration ($\Delta G^{\text{o},\neq}$), and it is strongly dependent of reaction temperature (T). For reactions in liquid phase, the standard state concentration is 1 M, and expression of TST in this case is simply

$$k = \kappa \sigma \frac{k_{\text{B}} T}{h} \exp\left[\frac{-\Delta G^{\text{o},\neq}(T)}{RT} \right] \tag{3}$$

where h is Planck constant, k_{B} is Boltzmann constant, and R is ideal gas constant. σ denotes degeneracy of reaction path and κ corresponds to transmission factor due to tunneling effect, they are not explicitly considered by concvar (to take them into account, users should directly specify the k values containing the two factors in input file).

The r_i^{gen} and r_i^{sink} in Eq. 1 correspond to generation and consumption rates of minimum i due to external reasons, respectively, and they can be manually defined in input file. Introduction of r^{sink} term may lead to unphysical negative concentration according to Eq. 1, so, concentration is automatically set to zero by concvar when it is detected to be negative during simulation.

Concvar performs simulation according to Eq. 1, during which the concentration of each minimum changes continuously, until maximum number of steps is reached or one of predefined concentration criteria is satisfied. Concentrations can also be fixed at constant values upon requested by user.

The choice of step size is a crucial point. Maximal simulation time corresponds to product of step size and number of simulation steps, and simulation cost is positively proportional to simulation steps. Therefore, for a simulation of specific length of time, the larger the step size, the lower the cost. However, the step size should not be too large, otherwise simulation result will be inaccurate or fully meaningless. According to our experience, assume that all elementary reactions are unimolecular, or some of them are bimolecular but one of the reactants has a concentration of 1 M, then maximally acceptable step size can be estimated as $0.1/k_{max}$, where k_{max} is maximal k of all elementary steps. This condition is equivalent to

$$
\begin{aligned}
& \Delta t < 0.1 / k_{max} \\
& \Rightarrow \Delta t\, k_{max} < 0.1 \\
& \Rightarrow \Delta t\, c\, k_{max} < 0.1c
\end{aligned}
\tag{4}
$$

where c is the concentration of reactant of the elementary reaction with k_{max}. Since the left side of above equation corresponds to concentration variation of the reactant in a simulation step, the condition guarantees that concentration should vary less than 10% at each step.

Now it is clear that the fastest elementary step determines maximally acceptable step size. A contradictory problem often encountered in practical research is that the energy barriers between some minima on the entire energy profile is very low, and the k of the corresponding interconversion reactions between them is significantly lower than that of other elementary reactions. Therefore, if performing the simulation directly, a very small simulation step size must be used, but in this case it is difficult to observe the occurrence of other elementary reactions within an affordable number of simulation steps. Concvar provides a workaround to this difficulty. Given that interconversion between the aforementioned minima is quite rapid and they can easily reach thermodynamic equilibrium state, the interconversion reactions can be ignored and thus

a relatively large step size could be employed, and at the meantime Boltzmann redistribution of their total concentrations is conducted at each simulation step. Specifically, at each step the concentration of one of the minima that are easily interconverted is set to be

$$c_i = w_i^{\text{Boltz}} \sum_j c_j \tag{5}$$

where j loops over all the aforementioned minima, and w_i is occurrence probability of i among them given by Boltzmann distribution

$$w_i = \frac{e^{-\Delta G_i/RT}}{\sum_j e^{-\Delta G_j/RT}} \times 100\% \tag{6}$$

where ΔG_i is relative free energy of i with respect to the lowest free energy among the aforementioned minima.

3 Characters of concvar

Concvar is a command-line program and was developed based on Fortran 95, it can be run directly without installation and preparation of running environment. Executable files and a very detailed manual of concvar can be freely download at http://sobereva.com/soft/concvar.

An input file recording free energies and simulation setting is needed to run concvar. During simulation, concentrations of various substances will be continuously written to a text file. After the simulation has completed, user may use scientific data plotting tools such as Origin, Gnuplot and Grace to draw concentration variation curves.

Concvar has the following features:

• Input file is very convenient to manually create, and output information is very easy to understand

• Arbitrary number of minima and transition states can be defined

• Each minimum can connect arbitrary number of reaction paths

• Reaction rate constants can be manually set or automatically derived by transition state theory

• Rate of generation and consumption of any minimum can be defined

• Concentrations of minima can change spontaneously or be fixed at specific values

• Criterion of concentration for any minimum can be set to stop the simulation at the right time

• Boltzmann redistribution among specific minima can be conducted at every simulation step

• Both unimolecular and bimolecular elementary reactions can be included

4. Examples

Some examples of concvar are briefly given in the next few sections, more descriptions with complete input files can be found in concvar manual.

4.1 Example 1: Elementary reaction

The first example is very simple, namely an elementary reaction. As illustrated in Fig. 1, there are two minima and one transition state. The free energy of standard state concentration (G^o) of the three systems at 298.15 K are 0.0, 20.0 and -50.0 kcal/mol, respectively. Clearly, reverse reaction is negligible compared to forward reaction when temperature is not extremely high.
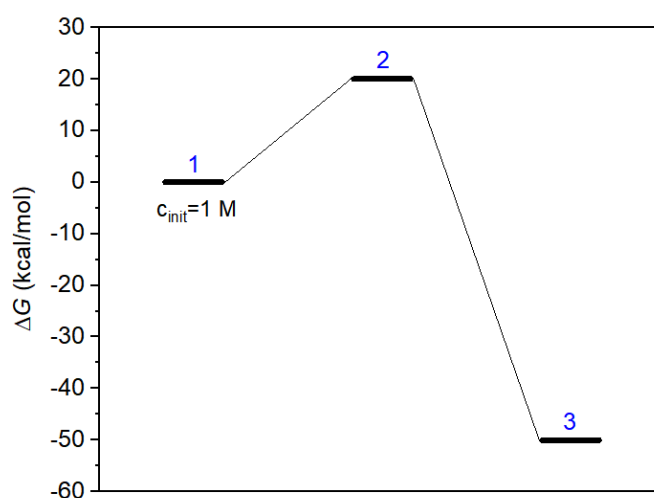

Fig. 1 Energy profile of an elementary reaction. Blue texts are system indices.

This example utilizes concvar to simulate concentration variations within 200 s at 298.15 K. The initial concentration of system 1 (reactant) is set to be 1.0 M, while that of system 3 (product) is default to be zero. The input file for this simulation is shown as follows, in which "stepsize" is step size of simulation in second, "nstep" is simulation steps, total simulation time length is 2000000*1E-4 = 200 s. "outfreq" is interval of steps of outputting concentrations to output file, in current simulation the concentrations are exported every 1000*1E-4 = 0.1 s. G^o of various systems in Fig. 1 are recorded in "$G" field, the linking relationship between transition states and minima is given in "$link" field, "$cinit" field defines initial concentrations. It is obvious that the input file of concvar is fairly concise.

```
temper= 298.15
stepsize= 1E-4
nstep= 2000000
outfreq= 1000
ioutfmt= 2

$G
0
20
-50

$link
1,2,3

$cinit
1 1.0
```

After booting up concvar and inputting path of the input file, the simulation starts. Note that prior to the simulation, concvar outputs detailed information about the simulation on screen, including free energy barriers and k of all involved elementary steps, as shown below. From which users can readily judge if input file has been correctly created and loaded. From the following information one can also see that the step size of current simulation is suggested to be smaller than 7.36 s, which is estimated

via the way mentioned in Section 2. The current simulation must have very satisfactory numerical accuracy since the employed step size (10^{-4} s) is smaller than 7.36 s by many orders of magnitude.

```
Temperature:  298.150 K
Simulation stepsize: 1.000E-04 s
Number of simulation steps:    2000000
Total simulation time:   200.000000 s (   3.333333 min)
Output frequency: per   1000 steps ( 0.100000 s)

Number of systems:    3
Number of minima:    2
   1 #  System index:    1,  G:    0.00 kcal/mol,  c(init):    1.000000 M
   2 #  System index:    3,  G:  -50.00 kcal/mol,  c(init):    0.000000 M

Number of reactions:    2
   1 #  System    1 to    3,  G barr:   20.00 kcal/mol,  k:  1.358867E-02 s^-1
   2 #  System    3 to    1,  G barr:   70.00 kcal/mol,  k:  3.040630E-39 s^-1
Maximum k is  1.358867E-02 s^-1
Stepsize is suggested to be smaller than 7.36E+00 s for present situation
```

During simulation, one can open the generated *conc.txt* file in current folder any time to monitor the concentrations at the latest step. After calculation has completed, from the *conc.txt* one can see the following information, which shows concentrations of all minima (systems 1 and 3, corresponding to reactant and product, respectively) during the simulation. It can be seen that at the end of the simulation (200 s), the concentration of reactant is 0.066 M and that of product is 0.934 M. The total concentration is always 1.0 M since generation and consumption terms are not defined in this simulation example.

```
 Step      Time(s)      1         3        Total
        0  0.0000E+00  1.0000000  0.0000000  1.0000000
     1000  1.0000E-01  0.9986421  0.0013579  1.0000000
     2000  2.0000E-01  0.9972860  0.0027140  1.0000000
...ignored
  1998000  1.9980E+02  0.0662037  0.9337963  1.0000000
  1999000  1.9990E+02  0.0661138  0.9338862  1.0000000
```

```
2000000  2.0000E+02  0.0660240  0.9339760  1.0000000
```

Half-life of a reaction can be easily obtained using concvar. By simply adding the following content to the input file, the simulation will be stopped if concentration of system 1 (reactant) is lower than 0.5 M (half of initial concentration).

```
$conccrit
1 < 0.5
```

During the new simulation, concvar stops at step 510092, which corresponds to 51.0092 s. Therefore, half-life of this reaction is 51.0092 s. For unimolecular elementary reaction, it is known that exact half-life is $\ln(2)/k$,[3] the result of concvar is in perfect agreement with the analytic solution.

4.2 Example 2: Complex reaction consisting of multiple elementary steps

This example employs concvar to simulate concentration variation of the complex reaction illustrated in Fig. 2. It can be seen that systems 1 and 7 are reactant and product, respectively. Systems 2, 4 and 6 are transition states (TSs), while systems 3 and 5 are intermediates (IMs). Forward and reverse barriers of all elementary steps are 15 kcal/mol and 20 kcal/mol, respectively. In this simulation, all substances have zero initial concentration, the reactant has a generation rate of 0.1 M/s, while product has a consumption rate of 0.08 M/s. The generation and consumption rates can be defined by adding "$gen" and "$sink" fields to the input file, respectively.
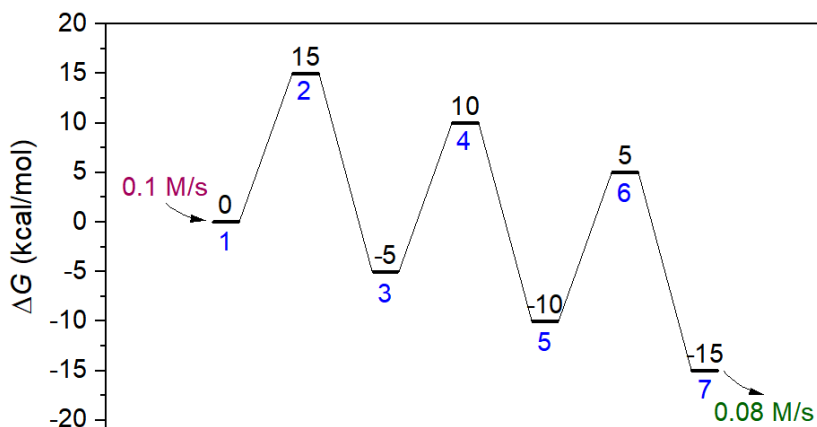
Fig. 2 Energy profile of a complex reaction consisting of multiple elementary steps. Blue texts are system indices.

The simulation is conducted at 250 K for 50 s, the concentration variations of all minima printed by concvar are plotted together in Fig. 3. Due to the continuous external input of reactant, it can be seen that concentration of system 1 (reactant), system 3 (IM1) and system 5 (IM2) accumulate gradually in sequence, then their concentrations reach the same maximum value (0.249 M) and keep unchanged since 20 s. Concentration of product is exactly zero until 10.7 s, because its consumption rate is set to be 0.08 M/s, while before 10.7 s its production rate due to reaction from system 5 (IM2) is lower than this value. After about simulation time of 20 s, concentration of product increases linearly with rate of 0.02 M/s, this is because reactant, IM1 and IM2 have reached steady state, namely their concentrations no longer change, while generation rate of reactant is higher than consumption rate of product by 0.02 M/s.
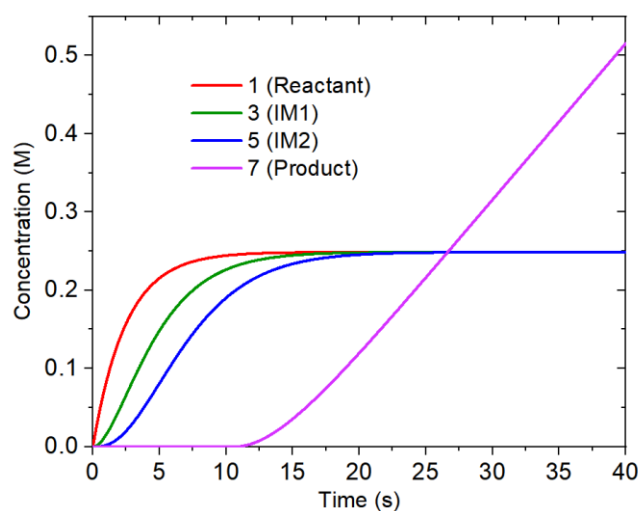
Fig. 3 Concentration variation of reactant, two intermediates and product during simulation of the model illustrated in Fig. 2.

4.3 Example 3: Complex reaction containing competitive paths

The energy profile of this example is shown as Fig. 4, it can be seen that there are two competitive reaction paths, respectively leading to two different products. Which one is major product? The answer is not quite clear according to common chemistry knowledge. However, by simulating the reaction via concvar, not only major product can be identified, but also ratio of the two products can be obtained.
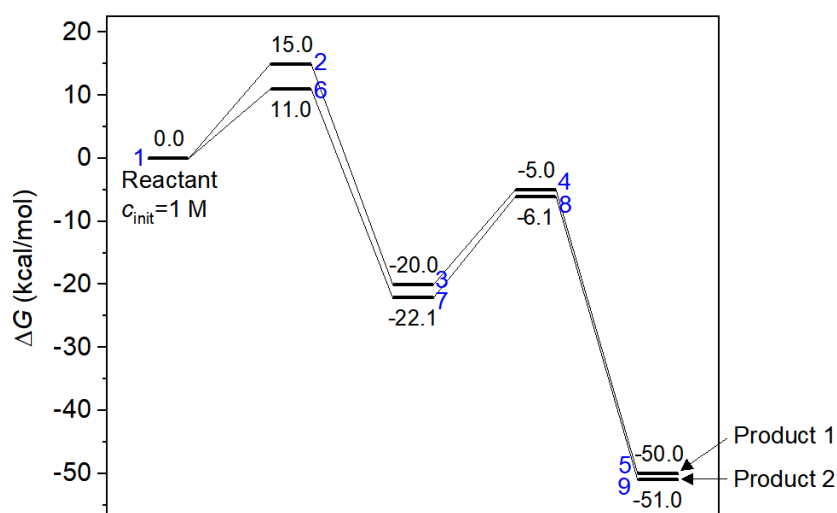

Fig. 4 Energy profile containing two competitive reaction paths. Blue texts are system indices.

In this simulation, initial concentration of reactant is set to be 1 M, and the simulation is conducted at 298.15 K for 1 s. The content of output file is shown below, systems 5 and 9 correspond to the two products. It can be seen that at the end of simulation, almost all reactant has transformed to the product of the second reaction path (0.9988 M), while product of the first reaction path is negligible (about 0.0012 M), see highlighted data in the output below. Also note that the reaction has proceeded sufficiently, because the concentrations of the reactant (system 1) and intermediates (systems 3 and 7) are basically zero at the end. Therefore, one can draw conclusion that major product is the product 1, and the reaction is extremely selective.

```
  Step     Time(s)      1          3          5          7          9        Total
       0  0.0000E+00  1.0000000  0.0000000  0.0000000  0.0000000  0.0000000  1.0000000
    1000  1.0000E-03  0.0000000  0.0010982  0.0000699  0.9875062  0.0113257  1.0000000
    2000  2.0000E-03  0.0000000  0.0010313  0.0001367  0.9760981  0.0227339  1.0000000
 ignored...
  998000  9.9800E-01  0.0000000  0.0000000  0.0011681  0.0000092  0.9988227  1.0000000
  999000  9.9900E-01  0.0000000  0.0000000  0.0011681  0.0000091  0.9988228  1.0000000
 1000000  1.0000E+00  0.0000000  0.0000000  0.0011681  0.0000090  0.9988229  1.0000000
```

4.4 Example 4: Verifying Curtin–Hammett principle

Curtin–Hammett principle is a well-known principle in chemical kinetics,[4] it states that for a reaction that has a pair of reactants that undergo rapid interconversion, and meantime each one can irreversibly transform to a product, the product distribution is determined by the energy difference between the two rate-determining transition states.

The energy profile shown in Fig. 5 is a representative case when Curtin–Hammett principle is well suited. The systems 3 and 5 are two reactants, since the barrier between them is very low, their interconversion is quite fast and thus thermodynamics equilibrium between them is always satisfied at ambient condition. Because both the two reverse barriers (15+22=37 and 20+40=60 kcal/mol) are relatively high, possibility

of reverse reaction is negligible. According to Curtin–Hammett principle, since TS 1 is notably lower than TS 2, after the reaction is fully completed, concentration of product 1 should be significantly higher than that of product 2, despite that reactant 2 has a lower energy and thus possesses higher concentration than reactant 1 during the reaction, and despite that product 2 has a much lower energy than product 1 and thus it is much more thermodynamically favourable.
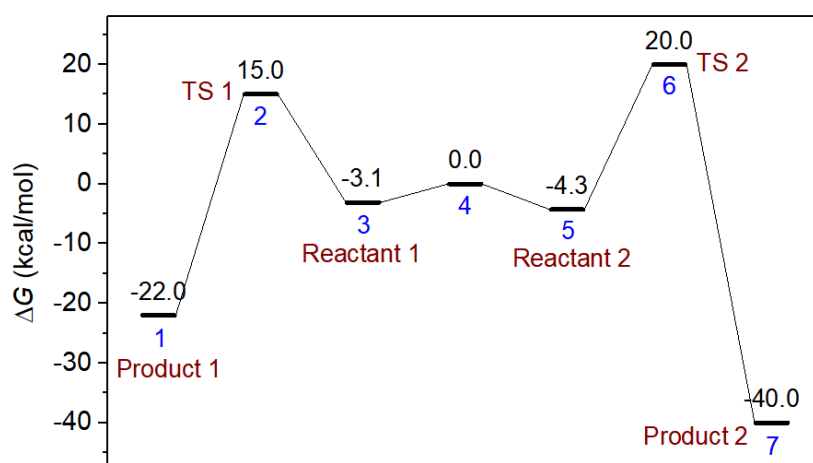

Fig. 5 A typical energy profile for which Curtin–Hammett principle works. Blue texts are system indices.

To verify the prediction by Curtin–Hammett principle, in this example concvar is used to numerically simulate the above reaction at 298.15 K. Because $k_{3 \to 5}$ and $k_{5 \to 3}$ are larger than other k in this simulation by many orders of magnitude, if they are explicitly considered like other elementary reactions, then extremely small step size (about 10^{-12} s) is needed, and in this case the number of steps needed for long enough simulation to study product distribution will be too large to be acceptable. Because interconversion between systems 3 and 5 is by far faster than any other elementary reaction, assumption of their thermodynamic equilibrium status is fully valid. So, in the present simulation, a special field *$Boltzmann* is added to the input file to force the concentration ratio between these two reactants to satisfy Boltzmann distribution at each step. Sum of concentrations of the two reactants in the simulation is set to be 1.0 M.
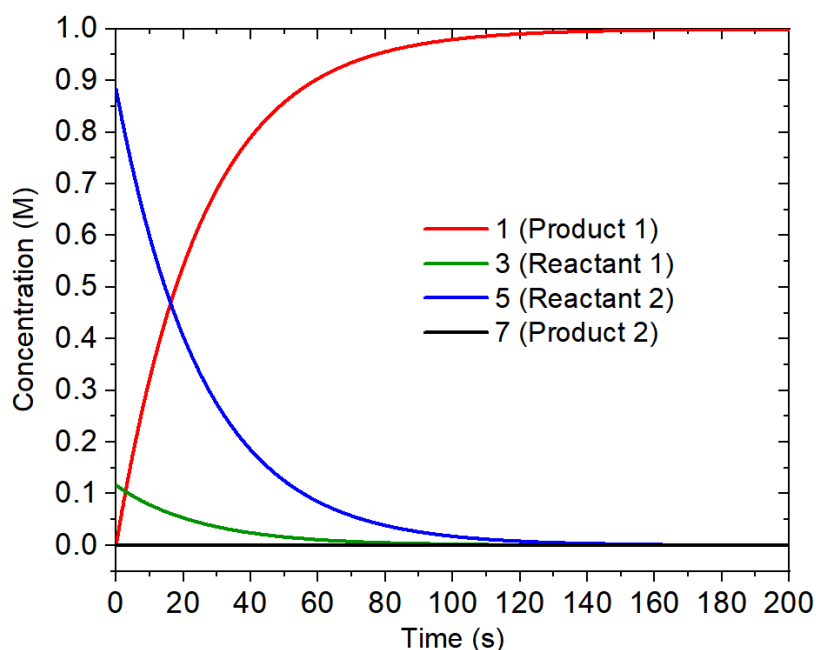
Fig. 6 Concentration variation of reactants and products during simulation of the model illustrated in Fig. 5.

The simulation result is plotted as Fig. 6. From the figure it can be seen that concentration of the two reactants gradually decreases, and their ratio is fixed throughout the simulation (0.131:1). Concentration of product 2 is always nearly zero, while that of product 1 accumulates rapidly, clearly product 1 is the dominant product. This observation is fully in line with the estimation due to the Curtin–Hammett principle.

4.5 Example 5: Finding rate-determining transition state

Rate-determining transition state (RDTS) refers to the transition state whose energy (more specifically, Gibbs free energy) plays most crucial role on determining overall reaction rate. Slight decrease of energy of RDTS makes reaction rate increase markedly. As illustrated in this example, concvar can also be used to determine RDTS by examining sensitivity of rate of product generation on energy of various transition states.
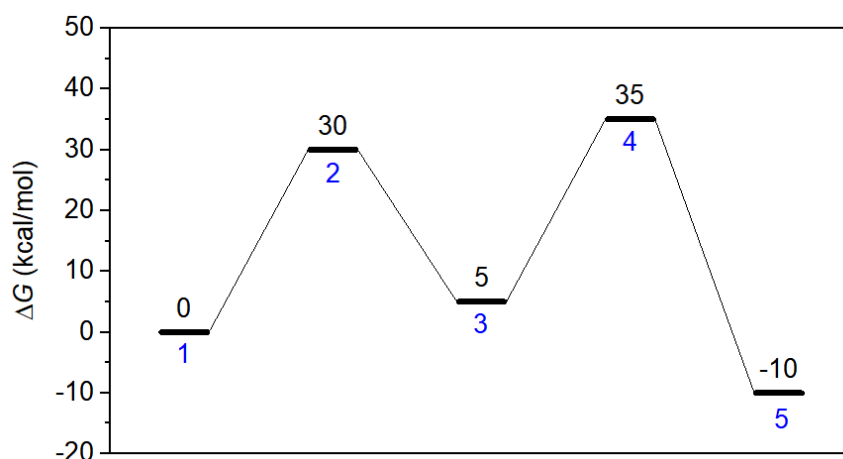
14

Fig. 7 A representative energy profile containing two elementary reactions. Blue texts are system indices.

Fig. 7 is a representative energy profile containing two elementary reactions, both of which have forward barrier of 30 kcal/mol. Systems 2 and 4 are the respective transition state, which one determines overall reaction rate most? To answer this question, a simulation is conducted by concvar at 350 K, initial concentration of reactant is set to be 1 M. Because the generation of the product is strongly hindered by the two relatively high barriers, in order to finally gain an obvious product concentration, the simulation time of this example is quite long (10^8 s). Because the maximal k in this simulation is small ($k_{3 \rightarrow 1}=1.79 \times 10^{-3}$ s^{-1}) and thus large step size can be used, we take step size of 10 s and only 10^7 simulation steps are needed in total. On a common notebook CPU (Intel i7-10870H), the simulation merely takes 16 s.

After complete of the simulation, product (system 5) has a concentration of 0.09678 M. If lowering free energy of the first transition state (system 2) by 1 kcal/mol and then repeating the simulation, finally the product will have a concentration of 0.09684 M. Obviously, the reaction rate has little sensitivity to the energy of the first transition. In contrast, if lowering free energy of the second transition state (system 4) by 1 kcal/mol and then repeating the simulation, finally the product has a concentration of 0.34796 M. Because the generation rate of product in this case becomes evidently

higher than before, one can unambiguously identify the second transition state as the RDTS.

The way of judging RDTS in this section could be viewed as a reliable and general method. We also note that there are manners of judging rate-determining state or rate-determining step without numerical simulation but by simply inspecting free energy profile, see Refs. [5] and [6].

4.6 Example 6: Complex reaction containing bimolecular elementary steps

Concvar also supports inclusion of bimolecular elementary steps in the studied complex reaction. To illustrate this point, the energy profile in Fig. 8 is taken as example. The first elementary step has bimolecular reactants, while the second elementary step has bimolecular products. Systems 2 is regarded as adding molecule, while system 7 is viewed as leaving molecule, they are treated normally as minima in concvar. Since only energy barriers affect result, it is customary to define free energy of adding and leaving molecules to zero. In the current simulation, initial concentration of system 1 is set to be 1.0 M, and concentration of the adding molecule keeps constant at 0.3 M via *$fix* field in the input file. In addition, this simulation assumes that the concentration of the leaving molecule is gradually lost at a maximal rate of 0.0001 M/s.
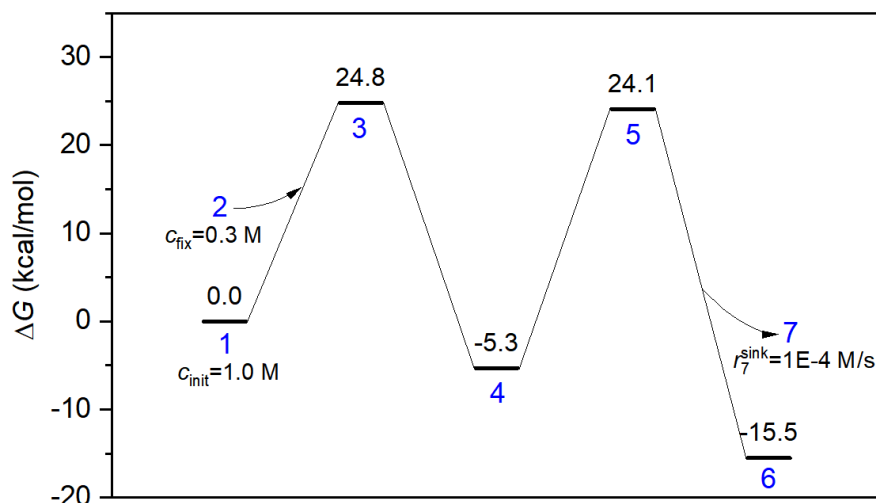
Fig. 8 A representative energy profile containing bimolecular elementary reactions. Blue texts are system indices.

The concentration variations of various substances during 10000 s simulation at 400 K are plotted as Fig. 9. As can be seen, after tens of seconds of the reaction, the reactant almost fully became intermediate. The concentration of adding molecules is always constant at 0.3 M due to fixing setting. With the further progress of the reaction, the concentration of the intermediate gradually decreased and the product was formed. After 10000 s, the original 1 M reactant all turned into the product. The concentration of leaving molecules peaked at about 2800 s, since its generation is slower than consumption after this time point, leaving molecule is fully consumed at the end of the entire simulation.
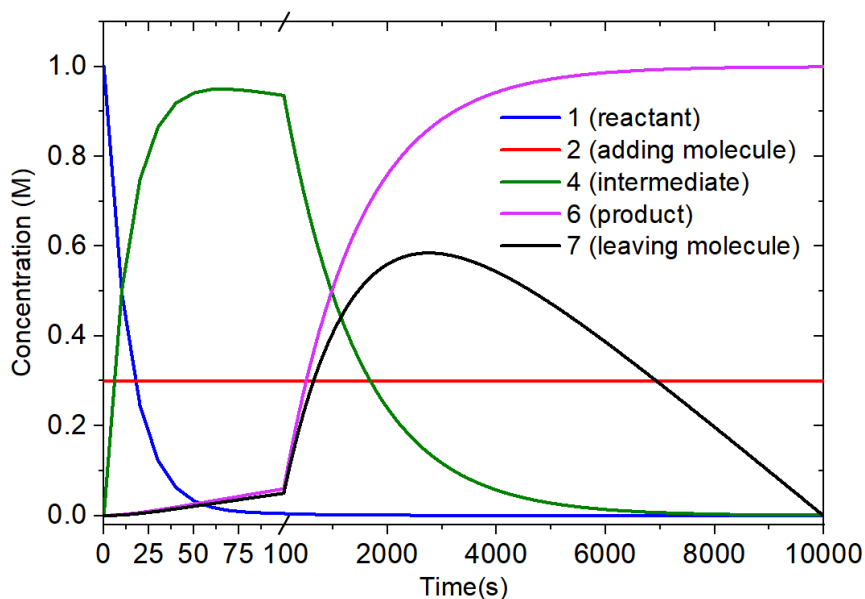
Fig. 9 Concentration variation of various substances during simulation of the model illustrated in Fig. 8. Since concentration of reactant and intermediate varies sharply at the beginning of the reaction, a break is set on the horizontal axis of the figure so that one can clearly examine their concentration variations.

5. Conclusion

In this article, our code named concvar is introduced, which is dedicated to solve time-dependent concentration variation of substances involved in complex reactions. As can be seen from the examples in this article, simulation model can be very flexibly defined, and thus concvar is applicable to almost all situations. Moreover, the input file of concvar is quite easy to create, and the outputted file can be conveniently plotted as curve map to intuitively examine how concentrations vary during the simulation. We have showed that concvar could be used to derive reaction half-life time, identify rate-determining transition state, and estimate ratio of products of competitive reaction paths, which are quite important in practical researches relating to chemical kinetics. In addition, concvar is also significant for physical chemistry education. By analyzing changes in substance concentrations in actual simulations, students can deepen their understanding of the principles of chemical kinetics and better comprehend concepts

18

such as transition state theory, transformation between substances, reaction rate constant, rate-determining states, reaction half-life, competitive reactions and Curtin–Hammett principle.

Acknowledgments

Conflicts of interest

The authors declare no conflict of interest.

References

[1]   T. Lu, Q. Chen, Shermo: A general code for calculating molecular thermochemistry properties. *Comput. Theor. Chem.* **2021**, *1200*, 113249.

[2]   B. C. Garrett, D. G. Truhlar, In Encyclopedia of Computational Chemistry; P. v. R. Schleyer, Ed.; John Wiley & Sons: West Sussex, 1998, p 3094.

[3]   P. Atkins, J. d. Paula, J. Keeler, Atkins' Physical Chemistry, Oxford University Press: New York, 2018.

[4]   Wikipedia contributors. Curtin–Hammett principle. in *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Curtin%E2%80%93Hammett_principle (accessed on Jan 12, 2022).

[5]   J. R. Murdoch, What is the rate-limiting step of a multistep reaction? *J. Chem. Educ.* **1981**, *58*, 32.

[6]   S. Kozuch, J. M. L. Martin, The Rate-Determining Step is Dead. Long Live the Rate-Determining State! *ChemPhysChem* **2011**, *12*, 1413.