eSPC, an Online Data Analysis Platform for Molecular Biophysics

KinGenie 1.0

User Documentation

November 2026

Contents

1. Surface-based binding **2**
1.1. Import data . . . 2
1.1.1. Input file . . . 2
A) Octet Biolayer interferometry (BLI) . . . 2
B) Gator Biolayer interferometry . . . 2
C) CSV file . . . 2
1.1.2. Processing . . . 2
1.2. Analysis . . . 3
1.2.1. Creating a Dataset . . . 3
1.2.2. Region of interest . . . 3
1.2.3. Models . . . 3
1.2.3.1. One-to-one (kinetics) . . . 3
1.2.3.2. One-to-one (steady-state) . . . 4
1.2.3.3. One-to-one (mass transport limitation) . . . 4
1.2.3.4. One-to-one (induced fit / two-step) . . . 5
1.2.3.5. One-to-one (conformational selection) . . . 6
1.2.3.6. Heterogeneous analyte . . . 7
1.2.3.7. Heterogeneous ligand . . . 7
1.2.3.8. Two-to-one (the ligand has two binding sites) . . . 8
1.2.3.9. Two-to-one (steady-state) . . . 8
1.2.3.10. Summary . . . 9
1.2.4. Signal and fitting . . . 9
1.3. Glossary (surface-based binding) . . . 10

2. Solution-based binding **10**
2.1. Import data . . . 10
2.1.1. Input file . . . 10
2.1.2. Processing . . . 10
2.2. Analysis . . . 10
2.2.1. Creating a dataset . . . 10
2.2.2. Models . . . 11
2.2.2.1. Single exponential . . . 11
2.2.2.2. Double exponential . . . 11
2.2.2.3. One binding site . . . 11
2.2.2.4. One binding site (induced fit) . . . 12
2.2.2.5. One binding site (conformational selection) . . . 12
2.2.3. Signal and fitting . . . 13
Contact details . . . 13

1. Surface-based binding

1.1. Import data

1.1.1. Input file

KinGenie can parse files generated by Octet (Sartorius) and Gator instruments, and common comma-separated-values (CSV) files.

A) Octet Biolayer interferometry (BLI)

Import the files with extension *.frd*. There should be one file per sensor. Data from multiple sensors can be imported simultaneously, if all of them share the same metadata.

B) Gator Biolayer interferometry

To import the results from a Gator BLI experiment, many files need to be imported together: the comma-separated-values (csv) file with the traces data, that contain the words 'Assay' and 'Channel'; the files describing the metadata, named *Setting.ini* and *ExperimentStep.ini*.

C) CSV file

For multi-cycle kinetics, the CSV file should have the following columns:

- *Time*: time data in seconds
- *Signal*: signal values at each time point
- *Smax*: Integer, identifies a certain measurement.
- *Analyte_concentration_micromolar_constant*: analyte concentration in micromolar units

For single-cycle kinetics, the CSV file should have one extra column:

- Cycle: cycle number for single cycle kinetics

Note: CSV example files can be exported from the Simulation panel. In the current surface CSV importer, only rows with the first distinct `Smax` value are loaded; use a single `Smax` value per input file.

1.1.2. Processing

Given a certain experiment, the sensor traces can be processed in the following ways:

- Alignment of the association phase: For each step, labelled in the metadata as a step where the association occurs, the mean of the last ten points from the previous step (should be the baseline) is calculated and subtracted from the association step and all subsequent steps.
- Baseline subtraction: Given a group of sensors and a reference sensor, the signal from the reference sensor is subtracted from the signal of the other sensors.
- Average: Given a group of sensors, a new sensor is created with the average signal.
- Inter-step correction (shift the **dissociation** phase): For each step, labelled in the metadata as a step where the dissociation occurs, the difference between the mean of the first n points and the mean of the last n points from the previous step (should be the association) is calculated and subtracted from the dissociation step. The value of n is selected by the user.
- Inter-step correction (shift the **association** phase): For each step, labelled in the metadata as a step where the association occurs, the difference between the mean of the last n points and the mean of the first n points from the next step (should be the dissociation) is calculated and subtracted from the association step. The value of n is selected by the user.
- Merge steps by index: Given the index of a certain reference step, the next or previous step inherits the step type. For example, if step 5 (dissociation) is selected as reference step, and step 6 (baseline) as "step to merge", both steps will be labelled as dissociation.
- Merge by step name: Given a step type, such as association or dissociation, all steps immediately following a step of that type will inherit the selected step type.

Given two experiments, the sensor traces from one experiment can be processed in the following way:

- Subtract experiment: Subtract compatible experiments sensor by sensor. The reference experiment may have the same number of sensors or a single reference sensor. The interface can also restrict subtraction to compatible interaction steps; the sampled time arrays must match the selected subtraction mode.

For surface-based experiments, double referencing combines reference-sensor subtraction with blank-injection

subtraction. First subtract the signal of a sensor without immobilized ligand, exposed to the same analyte concentration. Then subtract a blank injection that has itself undergone the same reference-sensor correction. This avoids subtracting the common background twice.

1.2. Analysis

1.2.1. Creating a Dataset

Once the traces are imported and preprocessed, they can be combined for global analysis. Traces with the same sample ID will be fitted with global kinetic parameters, such as *K*d and *k*off. Additionally, traces with the same sample ID and same *S*max ID will be fitted with a shared *S*max parameter, if the option to fit the data linked by *S*max is activated. The *S*max parameter is related to the theoretical total binding capacity of the sensor. Share *S*max only when the traces can reasonably be assumed to have the same binding capacity. This commonly applies to successive injections on the same sensor in single-cycle kinetics; use separate groups when loading, regeneration, or sensor activity changes the capacity.

1.2.2. Region of interest

Surface-based binding experiments contain two types of steps where (un)binding takes place: association and dissociation. During the association phase, the concentration of the analyte is kept constant, while during the dissociation the concentration of the analyte drops to zero. Both regions can be fitted with shared kinetic parameters. Alternatively, it is possible to fit the association phase, or only the dissociation phase. When fitting the dissociation phase, the *k*off is estimated, but not the *K*d.

If the association phase has reached equilibrium, the equilibrium signal versus analyte concentration can be fitted instead. This is useful when the kinetic transient is too fast to resolve, provided the concentration series contains enough information to estimate the affinity. Overall, the regions of interest for the fitting are:

- Association and dissociation (*K*d and *k*off is fitted)
- Association (*K*d and *k*off is fitted)
- Dissociation (*k*off is fitted)
- Steady-state (*K*d is fitted)

1.2.3. Models

In the models below, **L refers to the immobilized molecule (ligand)**, and **A to the analyte** in solution. Except in the explicit mass-transport description, surface association assumes a constant analyte concentration at the binding sites; the analyte concentration is zero during dissociation.

1.2.3.1. One-to-one (kinetics)

The system is described by the chemical equation:

$$A + L \rightleftharpoons AL \tag{1}$$

The differential equations for each species concentration are:

$$\frac{d[AL]}{dt} = [A][L]k_{on} - [AL]k_{off} \tag{2}$$

$$\frac{d[A]}{dt} = [AL]k_{off} - [A][L]k_{on} \tag{3}$$

$$\frac{d[L]}{dt} = [AL]k_{off} - [A][L]k_{on} \tag{4}$$

where *k*on is the association rate, and *k*off is the dissociation rate. The equilibrium dissociation constant is related to *k*on and *k*off as follows:

$$K_d = \frac{k_{off}}{k_{on}} \tag{5}$$

The signal S evolves according to the ordinary differential equation (ODE):

$$\frac{dS}{dt} = \frac{k_{off}}{K_d}[A](S_{max} - S) - k_{off}S \tag{6}$$

where $\frac{dS}{dt}$ is the rate of change of signal over time and *S*max is the total binding capacity. The analytical solution of the ODE is:

$$S(t) = S_{eq} + (S_0 - S_{eq})e^{-\nu t} \tag{7}$$

where

$$S_{eq}([A], K_d, S_{max}) = \frac{[A]}{[A] + K_d}S_{max} \tag{8}$$

$$\nu([A], K_d, k_{off}) = k_{off}\frac{[A] + K_d}{K_d} \tag{9}$$

For a single capacity group, the fitted parameters are *K*d, *k*off, and *S*max; *k*on is derived as *k*off / *K*d. Additional capacity groups introduce additional *S*max parameters. *S*0 is the initial signal, normally zero for the first association and carried forward between continuous single-cycle steps.

Note: If the dissociation phase is analysed alone, excluding the association phase, then we fit the initial signal. On the other hand, the *K*d parameter cannot be fitted. The analytical solution for the dissociation phase is:

$$S(t, S_0, k_{off}) = S_0 \cdot e^{-k_{off} \cdot t} \tag{10}$$

1.2.3.2. One-to-one (steady-state)

In this model, the association phase is assumed to have reached equilibrium, and the signal, at a given analyte concentration is described by the following equation:

$$S([A], K_d, S_{max}) = \frac{S_{max}[A]}{K_d + [A]} \tag{11}$$

where S is the signal at equilibrium, *S*max is the maximum theoretical response that could be observed (at infinite analyte concentration), $[A]$ is the analyte concentration, and *K*d is the equilibrium dissociation constant.

Note: Traces assigned to the same `Smax_ID` share one binding capacity and should have comparable active ligand loading. A dataset can contain several capacity groups with separate *S*max values and a shared *K*d; equal capacity is not required across all sensors.

1.2.3.3. One-to-one (mass transport limitation)

The system is described by the chemical equations:

$$A_{bulk} \rightleftharpoons A_{surface} \tag{12}$$

$$L + A_{surface} \rightleftharpoons LA \tag{13}$$

Using the reservoir approximation, the concentration of A_{bulk} remains constant.

The differential equations for each species concentration are:

$$\frac{d[A_{bulk}]}{dt} = 0 \tag{14}$$

$$\frac{d[L]}{dt} = [LA]k_{off} - [L][A_{surface}]k_{on} \tag{15}$$

$$\frac{d[A_{surface}]}{dt} = k_{off}[LA] + k_{tr,c}([A_{bulk}] - [A_{surface}]) - k_{on}[L][A_{surface}] \tag{16}$$

$$\frac{d[LA]}{dt} = [L][A_{surface}]k_{on} - [LA]k_{off} \tag{17}$$

Equations 15 to 17 express [L] and [LA] as equivalent concentrations within the same surface compartment. The physical transport rate $k_{tr,c}$ has units of s^{-1}. The association and dissociation rate constants are k_{on} and k_{off}.

Under the quasi steady-state approximation[1], a non-redundant ODE that provides a description of the system during the association phase is:

$$\frac{dS}{dt} = \frac{k_{on}[A]_0(S_{max} - S) - k_{off}S}{1 + k_{on}(S_{max} - S)/k_{tr,S}} \tag{18}$$

$$k_{tr,S} = \alpha k_{tr,c} \tag{19}$$

Here $[A]_0$ is the bulk analyte concentration, not depleted by binding. The conversion factor α relates bound concentration to measured signal, $S = \alpha[LA]$. The fitted transport coefficient $k_{tr,S}$ therefore has units of signal per μM per second and scales with the chosen signal units. The interface and pyKinGenie report this lumped coefficient as *k*tr; it must not be interpreted directly as the physical rate $k_{tr,c}$.

Overall, there are four variables to be fitted: *k*off, *K*d, *k*tr and *S*max.

Moreover, during the dissociation phase, when the concentration of A is zero, the differential equation required to describe the evolution of the signal is:

$$\frac{dS}{dt} = \frac{-k_{off}S}{1 + k_{on}(S_{max} - S)/k_{tr,S}} \tag{20}$$

Note: This model can only be used for fitting if the association and dissociation phases are selected.

1.2.3.4. One-to-one (induced fit / two-step)

The system is described by the chemical equations:

$$L + A \rightleftharpoons LA_{int} \tag{21}$$

$$LA_{int} \rightleftharpoons LA \tag{22}$$

*LA*int refers to an intermediate complex that changes to *LA* after an induced fit step. The differential equations for each species concentration are:

$$\frac{d[LA_{int}]}{dt} = [L][A]k_{on} + [LA]k_{rev} - [LA_{int}]k_c - [LA_{int}]k_{off} \tag{23}$$

$$\frac{d[L]}{dt} = [LA_{int}]k_{off} - [L][A]k_{on} \tag{24}$$

$$\frac{d[A]}{dt} = [LA_{int}]k_{off} - [L][A]k_{on} \tag{25}$$

$$\frac{d[LA]}{dt} = [LA_{int}]k_c - [LA]k_{rev} \tag{26}$$

Two non-redundant ODEs that provide a description of the system during the association phase are:

$$\frac{dS_1}{dt} = (-k_{off} - k_{on}[A])S_1 - k_{off}S_2 + k_{off}S_{max} \tag{27}$$

[1]Svitel, J., Boukari, H., Van Ryk, D., Willson, R. C., & Schuck, P. (2007). Probing the functional heterogeneity of surface binding sites by analysis of experimental binding traces and the effect of mass transport limitation. *Biophysical Journal*, *92*(5), 1742-1758. doi:10.1529/biophysj.106.094615.

$$\frac{dS_2}{dt} = -k_c S_1 + (-k_c - k_{rev})S_2 + k_c S_{max} \tag{28}$$

The total observed signal at any time is calculated as:

$$S_{total} = S_{max} - S_1 \tag{29}$$

In multi-cycle kinetics, the initial conditions are:

$$S_1 = S_{max} \tag{30}$$

$$S_2 = 0 \tag{31}$$

Overall, there are five variables to be fitted: *k*on, *k*off, *k*c, *k*rev and *S*max.

Note: This model can only be fitted if the association and dissociation phases are selected. One important assumption is that the signal generated by the intermediate and induced complex is the same.

1.2.3.5. One-to-one (conformational selection)

The system can be described by the chemical equations:

$$L_{inactive} \rightleftharpoons L_{active} \tag{32}$$

$$L_{active} + A \rightleftharpoons LA \tag{33}$$

*L*active refers to the ligand in an active conformation ready for binding, while *L*inactive is not capable of binding. The differential equations for each species concentration are:

$$\frac{d[L_{inactive}]}{dt} = [L_{active}]k_{rev} - [L_{inactive}]k_c \tag{34}$$

$$\frac{d[L_{active}]}{dt} = [L_{inactive}]k_c - [L_{active}]k_{rev} - [L_{active}][A]k_{on} + [LA]k_{off} \tag{35}$$

$$\frac{d[A]}{dt} = [LA]k_{off} - [L_{active}][A]k_{on} \tag{36}$$

$$\frac{d[LA]}{dt} = [L_{active}][A]k_{on} - [LA]k_{off} \tag{37}$$

Two non-redundant ordinary differential equations that provide a description of the system during the association phase are:

$$\frac{dS_1}{dt} = (-k_c - k_{rev})S_1 - k_{rev}S_2 + k_{rev}S_{max} \tag{38}$$

$$\frac{dS_2}{dt} = -k_{on}[A]S_1 + (-k_{on}[A] - k_{off})S_2 + k_{on}[A]S_{max} \tag{39}$$

In multi-cycle kinetics, the initial conditions are:

$$S_1 = S_{max} \,/\, (\frac{k_c}{k_{rev}} + 1) \tag{40}$$

$$S_2 = 0 \tag{41}$$

The total observed signal at any time is *S*2. Overall, the ODEs contain five variables to be fitted: *k*on, *k*off, *k*c, *k*rev and *S*max.

Note: This model can be used for simulating curves, but it is not available for fitting. One important assumption is that *L*active and *L*inactive are in equilibrium before the interaction with the analyte.

1.2.3.6. Heterogeneous analyte

In this model, the system can be described by the chemical equation:

$$A_i + L \rightleftharpoons A_iL \quad \forall i \in \{1, \dots, N\} \tag{42}$$

where Ai corresponds to the i-th analyte. Briefly, a Langmuir 1:1 interaction between each analyte Ai and the ligand (L) is assumed. The set of ODEs is:

$$\frac{dR_i}{dt} = k_{on,i} \cdot F_i \cdot [A]_0 \cdot R_{max,i} \cdot \delta - k_{off,i} R_i \tag{43}$$

$$R_{tot} = \sum_{i=1}^{N} R_i \tag{44}$$

where the subindex i refers to the i-th analyte. And δ is defined as:

$$\delta = 1 - \sum_{j=1}^{N} \frac{R_j}{R_{max,j}} \tag{45}$$

Note: This model can be used for simulating curves of a mixture of two analytes, but it is not available for fitting.

1.2.3.7. Heterogeneous ligand

In this model, the system can be described by the chemical equation:

$$A + L_i \rightleftharpoons AL_i \quad \forall i \in \{1, \dots, N\} \tag{46}$$

where Li corresponds to the ligand in the i-th state. Briefly, a Langmuir 1:1 interaction between each analyte A and the ligand (in the different states) is assumed. The set of ODEs is:

$$\frac{dS_i}{dt} = \frac{k_{off,i}}{K_{d,i}}[A](S_{max,i} - S_i) - k_{off,i} S_i \tag{47}$$

where Smax,i is the maximum signal for ligand site class i. For simplicity, we assume that there are only two classes of ligands. The total signal is then expressed as:

$$S = S1 + S2 \tag{48}$$

where $S1$ and $S2$ are respectively the signal produced by the ligand class/state "1" and ligand class/state "2". The values of Smax,1 and Smax,2 are determined by the fraction of sites of class "1" and global Smax:

$$S_{max,1} = S_{max,global} \cdot f_1 \tag{49}$$

$$S_{max,2} = S_{max,global} \cdot (1 - f_1) \tag{50}$$

During the dissociation phase, the analyte concentration drops to zero and the differential equation to describe the system is:

$$\frac{dS_i}{dt} = -k_{off,i} S_i \tag{51}$$

Note: The two classes have independent Kd,i and $k$off,i values, and do not interconvert. Their maximum responses are partitioned by $f1$ as in Eqs. 49-50. For the kinetic fit, the parameters are Kd,1, $k$off,1, Kd,2, $k$off,2, $f1$, and the total Smax per capacity group. A steady-state fit uses the sum of the two one-site equilibrium responses and estimates the two affinities, the response fraction, and the total capacity.

1.2.3.8. Two-to-one (the ligand has two binding sites)

The system can be described by the chemical equations:

$$L + A \rightleftharpoons LA \tag{52}$$

$$LA + A \rightleftharpoons LA_2 \tag{53}$$

where LA is the single bound complex and $LA2$ is the double-bound complex.

The differential equations for each species concentration are as follows:

$$\frac{d[L]}{dt} = -2k_{on}[L][A] + k_{off}[LA] \tag{54}$$

$$\frac{d[A]}{dt} = -2k_{on}[L][A] + k_{off}[LA] - \sqrt{\sigma}k_{on}[LA][A] + \frac{2k_{off}}{\sqrt{\sigma}}[LA_2] \tag{55}$$

$$\frac{d[LA]}{dt} = 2k_{on}[L][A] - k_{off}[LA] - \sqrt{\sigma}k_{on}[LA][A] + \frac{2k_{off}}{\sqrt{\sigma}}[LA_2] \tag{56}$$

$$\frac{d[LA_2]}{dt} = \sqrt{\sigma}k_{on}[LA][A] - \frac{2k_{off}}{\sqrt{\sigma}}[LA_2] \tag{57}$$

Here, σ is the cooperativity factor. If σ is lower than one, there is negative cooperativity, if σ is equal to one, there is no cooperativity, and if σ is higher than one, there is positive cooperativity.

Two non-redundant ordinary differential equations that provide a description of the system during the association phase are:

$$\frac{df_{LA}}{dt} = 2k_{on}[A] - (2k_{on}[A] + k_{off} + \sqrt{\sigma}k_{on}[A])f_{LA} + (\frac{2k_{off}}{\sqrt{\sigma}} - 2k_{on}[A])f_{LA_2} \tag{58}$$

$$\frac{df_{LA_2}}{dt} = \sqrt{\sigma}k_{on}[A]f_{LA} - \frac{2k_{off}}{\sqrt{\sigma}}f_{LA_2} \tag{59}$$

where f_{LA} is the fraction of protein in the single-bound state, and f_{LA_2} is the fraction of protein in the double-bound state.

The total signal S is given by the signal produced by f_{LA} and f_{LA_2}:

$$S = f_{LA} \cdot S_{max,LA} + f_{LA_2} \cdot S_{max,LA_2} \tag{60}$$

where $S_{max,LA}$ is the maximum signal produced by the single-bound complex and S_{max,LA_2} is the signal produced by the double-bound complex.

Note: The two sites have the same intrinsic binding rates before occupancy is considered, and the statistical factors account for the number of available or occupied sites. The cooperativity factor modifies the second binding step. KinGenie uses different signal amplitudes for the singly and doubly bound states.

1.2.3.9. Two-to-one (steady-state)

In this model, the association phase is assumed to have reached equilibrium, and the signal, at a given analyte concentration is described by the following equation:

$$S = S_{max,LA}f_{LA} + S_{max,LA_2}f_{LA_2} \tag{61}$$

with f_{LA} and f_{LA_2} being:

$$f_{LA} = \frac{2 \cdot (C/K_d)}{1 + 2(C/K_d) + \sigma\,(C/K_d)^2} \tag{62}$$

$$f_{LA_2} = \frac{\sigma(C/K_d)^2}{1 + 2(C/K_d) + \sigma(C/K_d)^2} \tag{63}$$

where C is the analyte concentration and KD is the intrinsic equilibrium dissociation constant, and σ is the cooperativity factor.

1.2.3.10. Summary

Model name	Type of analysis	Binding parameters (excluding response amplitudes)	Simulation	Fitting
One-to-one (steady-state)	Equilibrium	K_d	NO	YES
Two-to-one (steady-state)*	Equilibrium	K_d, σ	NO	YES
Heterogeneous ligand (steady-state)	Equilibrium	$K_{d,1}$, $K_{d,2}$, f_1	NO	YES
One-to-one (kinetics)	Kinetics	K_d, k_{off}	YES	YES
One-to-one (mass transport limitation)	Kinetics	K_d, k_{off}, $k_{tr,S}$	YES	YES
One-to-one (induced fit / two-step)	Kinetics	k_{on}, k_{off}, k_c, k_{rev}	YES	YES
Two-to-one*	Kinetics	K_d, k_{off}, σ	YES	YES
One-to-one (conformational selection)	Kinetics	k_{on}, k_{off}, k_c, k_{rev}	YES	NO
Heterogeneous analyte	Kinetics	$K_{d,1}$, $k_{off,1}$, $K_{d,2}$, $k_{off,2}$, analyte fractions	YES	NO
Heterogeneous ligand (kinetics)	Kinetics	$K_{d,1}$, $k_{off,1}$, $K_{d,2}$, $k_{off,2}$, f_1	YES	YES

* The two-site Kd refers to the microscopic equilibrium dissociation constant.

1.2.4. Signal and fitting

The one-to-one (steady-state) and one-to-one (kinetics) signal is calculated using an analytical equation, while for the mass-transport-limitation and heterogeneous analyte models we solve the ODE using the *solve_ivp* function from *scipy*, that numerically integrates a system of ordinary differential equations given a set of initial conditions. Last, for the induced-fit and conformational-selection models, the ODEs are solved using the matrix exponential method.

Models available for fitting use non-linear least squares through `scipy.optimize.curve_fit`, with bounds on the fitted parameters. Some routines expand the bounds and refit automatically. Inspect the fitted parameters bounds before trusting results.

The standard deviation of all fitted parameters is then computed using the square root of diagonal values from the fit parameter covariance matrix reported by *scipy.curve_fit* function. These are approximate standard errors based on a local linear approximation around the optimum, not guaranteed underestimates of the uncertainty. They can be unreliable when the model is poorly identifiable, parameters are near bounds, or the error assumptions are inappropriate.[2] Relative errors are calculated as $100 * (std(\hat{\theta}) \ / \ \hat{\theta})$ where $\hat{\theta}$ refers to the estimate of the parameter and $std(\hat{\theta})$ to the standard deviation.

In the case of the one-to-one (steady-state) and one-to-one (kinetics) models, we also provide the marginal asymmetric confidence interval. Briefly, the lower and upper bounds of the 95 % confidence interval are given by the values of Kd (or *koff*) satisfying

$$RSS(K_d) = RSS_0(1 + \frac{criticalValue}{n - p}) \tag{64}$$

where $RSS0$ is the residual sum of squares using the best estimates for all the parameters, $RSS(K\text{d})$ is the residual sum of squares using a fixed value of Kd (fitting again the other parameters), n is the number of data points, p is the number of parameters, and *criticalValue* is the 95th percentile of the Fisher-Snedecor distribution with **1 numerator degree of freedom and** $n - p$ **denominator degrees of freedom**, $F_{0.95}(1, n - p)$.

[2] SciPy, `scipy.optimize.curve_fit`, especially the covariance and parameter-uncertainty notes.

General note:

In the kinetics models, the *S*max parameter can be fitted globally within a capacity group or locally for each trace. Link it only when the selected traces are expected to have the same active binding capacity; single-cycle measurements on the same sensor are a common example.

1.3. Glossary (surface-based binding)

Term	Definition
Analyte	Biomolecule supplied in solution at a constant concentration during a surface-based association step
Association phase	Phase of a surface-based experiment where the concentration of the analyte is kept constant and flows over the surface
Baseline	Phase of a surface-based experiment where the biosensors measure the signal of the buffer. A baseline before the association step is needed for baseline-based alignment; already aligned or simulated traces can be imported without it.
Biosensor	Disposable reactive surface used to immobilize the ligand and detect binding
Dissociation phase	Phase of a surface-based experiment where the concentration of the analyte is dropped to zero and the signal diminishes due to the dissociation of the complex. It comes immediately after the association phase
Double-reference	Reference-surface subtraction followed by subtraction of a blank injection that has undergone the same reference-surface correction
*K*d	Equilibrium dissociation constant ($K_d = k_{off}/k_{on}$ for a one-step binding interaction)
*k*off	Rate of dissociation of the complex in the reaction: $PL \rightarrow P + L$
*k*on	Rate of formation of the complex in the reaction: $P + L \rightarrow PL$
Ligand	In a surface-based experiment, the ligand refers to the biomolecule that is fixed to the surface, and is usually a protein. On the other hand, in a solution-based experiment, the ligand moves freely in solution and could be a small molecule
Loading	Phase of a surface-based experiment where the ligand is flown at a fixed concentration to immobilize it
Reference sensor	Sensor without immobilized active ligand (or with a non-binding control), exposed to the same analyte concentration to correct for background and non-specific binding
*S*max	Maximum theoretical binding capacity of the sensor. Depends on the amount of loading and is usually unique for a given sensor and experiment
Steady-state	Condition of a chemical reaction where all state variables remain constant, such as the concentration of the biomolecules.

2. Solution-based binding

2.1. Import data

2.1.1. Input file

A) A CSV file with the following columns:

- Time: time data in seconds
- Signal: signal values at each time point
- Protein_concentration_micromolar: protein concentration in micromolar units
- Ligand_concentration_micromolar: ligand concentration in micromolar units

2.1.2. Processing

Given a certain experiment, a time cutoff can be selected such that all traces will be truncated at the specified time point.

2.2. Analysis

2.2.1. Creating a dataset

After the traces are imported, inside the '2. Analyse' tab, a dataset for analysis can be created. Curves with the same SampleID will be fitted with shared kinetic parameters.

2.2.2. Models

2.2.2.1. Single exponential

In this model, the signal Y of each curve is fitted using a single exponential function:

$$Y = a_0 + a_1 \cdot e^{-k_{obs} \cdot t} \tag{65}$$

where $a0$ is the constant offset (the long-time limit), $a1$ is the pre-exponential term, $k$obs is the observed rate constant, and t is the time in seconds.

Note: In this model, the $k$obs values are fitted individually.

2.2.2.2. Double exponential

In this model, the signal Y of each curve is fitted using a double exponential function:

$$Y = a_0 + a_1 \cdot e^{-k_{obs,1} \cdot t} + a_2 \cdot e^{-k_{obs,2} \cdot t} \tag{66}$$

where $a0$ is the constant offset (the long-time limit), $a1$ and $a2$ are the pre-exponential terms, $k$obs,1 and $k$obs,2 are the observed rate constants, and t is the time in seconds.

Note: In this model, the $k$obs,1 and $k$obs,2 values are fitted individually.

2.2.2.3. One binding site

The system is described by the chemical equation:

$$A + B \rightleftharpoons AB \tag{67}$$

The differential equations for each species concentration are:

$$\frac{d[AB]}{dt} = [A][B]k_{on} - [AB]k_{off} \tag{68}$$

$$\frac{d[A]}{dt} = [AB]k_{off} - [A][B]k_{on} \tag{69}$$

$$\frac{d[B]}{dt} = [AB]k_{off} - [A][B]k_{on} \tag{70}$$

where $k$on is the association rate, and $k$off is the dissociation rate. The equilibrium dissociation constant is related to $k$on and $k$off as follows:

$$K_d = \frac{k_{off}}{k_{on}} \tag{71}$$

One ODE is required to calculate the species concentrations:

$$\frac{d[AB]}{dt} = ([A_t] - [AB])([B_t] - [AB])\frac{k_{off}}{K_d} - [AB]k_{off} \tag{72}$$

where $[A\mathrm{t}]$ is the total concentration of A, and $[B\mathrm{t}]$ is the total concentration of B. The initial condition is that $[AB]$ equals zero at $t = t0$. The parameter $t0$ can be left at zero, or fitted. The total signal is obtained as follows:

$$Y(t) = Y_a([A_t] - [AB]) + Y_b([B_t] - [AB]) + Y_{ab}([AB]) \tag{73}$$

where Ya, Yb and Yab is the signal produced by free (unbound) A, free B and the complex AB.

Note: Ya, Yb and Yab can be fitted or left at zero.

2.2.2.4. One binding site (induced fit)

The system is described by the chemical equations:

$$A + B \rightleftharpoons AB_{int} \tag{74}$$

$$AB_{int} \rightleftharpoons AB \tag{75}$$

ABint refers to an intermediate complex that changes to AB after an induced fit step. The differential equations for each species concentration are:

$$\frac{d[A]}{dt} = [AB_{int}]k_{off} - [B][A]k_{on} \tag{76}$$

$$\frac{d[B]}{dt} = [AB_{int}]k_{off} - [B][A]k_{on} \tag{77}$$

$$\frac{d[AB_{int}]}{dt} = [B][A]k_{on} + [AB]k_{rev} - [AB_{int}]k_c - [AB_{int}]k_{off} \tag{78}$$

$$\frac{d[AB]}{dt} = [AB_{int}]k_c - [AB]k_{rev} \tag{79}$$

The initial condition is that $[AB\text{int}]$ and $[AB]$ equal zero at the start of the reaction. By using the law of mass-conservation, only the two last ODE are required to obtain the concentration of all species:

$$[A] = [A_t] - [AB_{int}] - [AB] \tag{80}$$

$$[B] = [B_t] - [AB_{int}] - [AB] \tag{81}$$

The total signal is obtained as follows:

$$Y(t) = Y_a([A_t] - [AB_{int}] - [AB]) + Y_b([B_t] - [AB_{int}] - [AB]) + Y_{ab,int}[AB_{int}] + Y_{ab}[AB] \tag{82}$$

where Ya, Yb, Yab,int and Yab is the signal produced by free (unbound) A, free B and the complex ABint and the induced-complex AB.

Note: Ya, Yb, Yab.int, Yab can be fitted or left at zero. Yab.int can be assumed to be equal to Yab.

2.2.2.5. One binding site (conformational selection)

The system can be described by the chemical equations:

$$A_{inactive} \rightleftharpoons A_{active} \tag{83}$$

$$A_{active} + B \rightleftharpoons AB \tag{84}$$

Aactive refers to the molecule A in an active conformation ready for binding, while Ainactive is not capable of binding. The differential equations for each species concentration are:

$$\frac{d[A_{inactive}]}{dt} = [A_{active}]k_{rev} - [A_{inactive}]k_c \tag{85}$$

$$\frac{d[B]}{dt} = [AB]k_{off} - [A_{active}][B]k_{on} \tag{86}$$

$$\frac{d[A_{active}]}{dt} = [A_{inactive}]k_c - [A_{active}]k_{rev} - [A_{active}][B]k_{on} + [AB]k_{off} \tag{87}$$

$$\frac{d[AB]}{dt} = [A_{active}][B]k_{on} - [AB]k_{off} \tag{88}$$

The initial condition is that at $t = 0$, $[A\text{active}]$ and $[Ainactive]$ are in equilibrium, and $[AB]$ equals zero. By using the law of mass-conservation, only the two last ODE are required to obtain the concentration of all species:

$$[A_{inactive}] = [A_t] - [A_{active}] - [AB] \tag{89}$$

$$[B] = [B_t] - [AB] \tag{90}$$

The total signal is obtained as follows:

$$Y(t) = Y_a([A_{inactive}] + [A_{active}]) + Y_b([B]) + Y_{ab}([AB]) \tag{91}$$

where Ya, Yb, and Yab are the signal coefficients for free A (active or inactive), free B, and the complex AB, respectively. This expression assumes the active and inactive conformations of free A have the same signal coefficient.

Note: This model is still only available for simulation, but not for fitting.

2.2.3. Signal and fitting

For solution-based one-site binding, pyKinGenie uses an analytical solution. The induced-fit and conformational-selection models use `scipy.integrate.odeint` to integrate their reduced ODE systems; the concentrations of the other species are recovered by mass conservation. The observed signal is the weighted sum of the species concentrations.

The solution models available for fitting use non-linear least squares through `scipy.optimize.curve_fit`, with bounds on the fitted parameters. These fitting routines can expand the bounds and refit when an estimate is close to a boundary.

Contact details

For further assistance, please contact us:

Email: spc@embl-hamburg.de

Address: EMBL (c/o DESY), Notkestrasse 85, Build. 25a, 22607 Hamburg, Germany