A two-point IC_{50} method for evaluating the biochemical potency of irreversible enzyme inhibitors

Petr Kuzmič

BioKin Ltd., Watertown, Massachusetts, USA
http://www.biokin.com

Abstract

Irreversible (covalent) enzyme inhibitors cannot be unambiguously ranked for biochemical potency by using IC_{50} values determined at a single point in time, because the same IC_{50} value could originate either from relatively low initial binding affinity accompanied by high chemical reactivity, or the other way around. To disambiguate the potency ranking of covalent inhibitors, here we describe a data-analytic procedure relying on two separate IC_{50} values, determined at two different reaction times. In the case of covalent inhibitors following the two-step kinetic mechanism $E + I \rightleftharpoons E{\cdot}I \rightarrow EI$, the two IC_{50} values alone can be used to estimate both the inhibition constant (K_i) as a measure of binding affinity and the inactivation rate constant (k_{inact}) as a measure of chemical reactivity. In the case of covalent inhibitors following the one-step kinetic mechanism $E + I \rightarrow EI$, a simple algebraic formula can be used to estimate the covalent efficiency constant (k_{inact}/K_i) from a single experimental value of IC_{50}. The two simplifying assumptions underlying the method are (1) zero inhibitor depletion, which implies that the inhibitor concentrations are always significantly higher than the enzyme concentration; and (2) constant reaction rate in the uninhibited control assay. The newly proposed method is validated by using a simulation study involving 64 irreversible inhibitors with covalent efficiency constants spanning seven orders of magnitude.

Key words: enzyme kinetics; inhibition; irreversible inhibition; covalent inhibition; mathematical model; algebraic model

1. Introduction

Many medicines currently in use to treat various human diseases and symptoms are enzyme inhibitors. Furthermore, many important drugs and drug candidates are irreversible covalent inhibitors [1–4], which express their pharmacological effect by forming a permanent chemical bond with the protein target. Probably the most well known representative of this class is acetylsalicylate, or Aspirin, an irreversible covalent inhibitor of cyclooxygenase. Evaluating the biochemical potency of irreversible inhibitors in the process of pre-clinical drug discovery is exceedingly challenging. Even the task of simply arranging a list of potential drug candidates in order of their biochemical potency presents a serious obstacle. The main challenge is that the overall biochemical potency of irreversible enzyme inhibitors has two distinct components, namely, binding affinity measured by the inhibition constant (K_i) and chemical reactivity measured by the inactivation rate constant (k_{inact}).

Medicinal chemists and pharmacologists involved in drug discovery are accustomed to expressing the biochemical potency of enzyme inhibitors primarily in terms of the I_{50} [5][1]. However, in the case of irreversible inhibitors, the I_{50} by definition decreases over time, until at asymptotically infinite time it approaches one half of the enzyme concentration. Therefore, reporting the I_{50} for an irreversible inhibitor without also specifying the corresponding reaction time is essentially meaningless. An even more serious conceptual problem is that the same value of I_{50} could originate either from relatively high initial binding affinity (low K_i) and relatively low chemical reactivity (low k_{inact}), or the other way around. Therefore, two or more inhibitors with disparate molecular properties could easily manifest as having "the same" biochemical potency if the I_{50} assay is conducted at a single point in time.

This report presents a data-analytic procedure that relies on two separate I_{50} determinations, conducted at two different reaction times. If a given inhibitor follows a stepwise mechanism of inhibition, involving a kinetically detectable noncovalent intermediate, we show that it is possible to estimate both K_i and k_{inact} from the two time-dependent I_{50} measurements alone. Many highly potent covalent inhibitors display a one-step kinetic mechanism, apparently without the involvement of a clearly detectable noncovalent intermediate [6]. In those cases it is in principle impossible to measure the K_i and k_{inact} separately, but the data-analytic method described here allows for the determination of the covalent efficiency constant k_{eff}, also known as k_{inact}/K_i, from any single measurement of I_{50}. The proposed method is validated by using a simulation study involving 64 computer-generated inhibitors with molecular properties (K_i, k_{inact}, and k_{inact}/K_i) spanning at least six orders of magnitude.

[1] Throughout this manuscript, the conventionally used notation IC_{50} is abbreviated as I_{50}.

2. Methods

This section describes the theoretical and mathematical methods that were used in heuristic simulations described in this report. All computations were performed by using the software package DynaFit [7, 8].

2.1. Kinetic mechanisms of irreversible inhibition

In this report we will consider in various contexts the kinetics mechanisms of substrate catalysis and irreversible inhibition depicted in *Figure 1*. For further details see ref. [9].

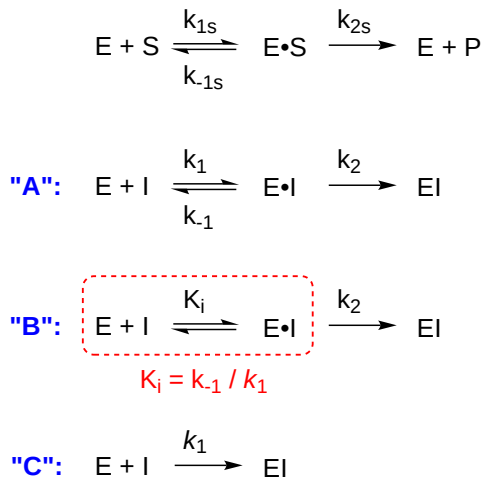

Figure 1: Kinetic mechanisms of substrate catalysis (top) and covalent inhibition (mechanisms **A** – **C**).

2.2. Mathematical models

2.2.1. Determination of k_1^* from a single measurement of I_{50}

On the assumption that the enzyme assay proceeds kinetically via the one-step inhibition mechanism **C**, the apparent second-order rate constant $k_1^* = k_1 / (1 + [\mathrm{S}]_0/K_\mathrm{M})$ can be computed directly from a single measurement of I_{50} by using Eqn (1), where t_{50} is the reaction time used to determine the experimental value of I_{50} and c is a constant depending on the dimensions used to express time and concentration. When time is expressed in seconds and concentrations in micromoles per liter, $c = 1.5936$. The requisite algebraic derivation is shown in the *Supporting Information*. Eqn (2) defines the second-order covalent efficiency constant in the context of the one-step mechanism **C**.

$$k_1^* = \frac{c}{I_{50}\, t_{50}} \tag{1}$$

$$k_\mathrm{eff} = k_1^* \left(1 + \frac{[\mathrm{S}]_0}{K_\mathrm{M}}\right) \tag{2}$$

2.2.2. Model discrimination analysis

On the assumption that the enzyme assay proceeds kinetically via the one-step inhibition mechanism **C**, the apparent second-order rate constant k_1^* is by definition invariant with respect to time, according to Eqn (1). Thus, in the idealized case of zero experimental error, two values $k_1^{*(1)}$ and $k_1^{*(2)}$ determined at two different stopping times $t^{(1)}$ and $t^{(2)}$ should be exactly identical. However, in the realistic case of non-zero experimental error, the two k_1^* values will always be ever so slightly different even if the one-step kinetic mechanism **C** is actually operating. In order to decide whether or not $k_1^{*(1)}$ and $k_1^{*(2)}$ are sufficiently similar to warrant the acceptance of the one-step kinetic model, we use the geometric standard deviation defined by Eqn (4), where μ_g is the geometric mean defined by Eqn (3).

$$\mu_g = \sqrt{k_1^{*(1)}\, k_1^{*(2)}} \tag{3}$$

$$\sigma_g = \exp \sqrt{\left(\ln \frac{k_1^{*(1)}}{\mu_g}\right)^2 + \left(\ln \frac{k_1^{*(1)}}{\mu_g}\right)^2} \tag{4}$$

The maximum acceptable value of σ_g will depend on the experimental situation. In the simulation study reported below, where the pseudo-random noise was equal to one percent of the maximum signal (e.g., fluorescence value), we found that $\sigma_g < 1.25$ was producing satisfactory results. Note that the geometric standard deviation σ_g is a dimensionless quantity measuring the "X-fold variation" associated with a group of numerical values. In our situation, $\sigma_g < 1.25$ means that the two values $k_1^{*(1)}$ and $k_1^{*(2)}$ in a statistical sense differ by less than a factor of 1.25, or roughly by 25 percent in either direction (lower or higher).

2.2.3. Determination of k_inact and K_i^* from two values of I_{50}

On the assumption that the enzyme assay proceeds kinetically via the two-step inhibition mechanism **B**, the apparent inhibition constant $K_\mathrm{i}^* = K_\mathrm{i}\,(1 + [\mathrm{S}]_0/K_\mathrm{M})$ can be computed directly from any two measurements of I_{50} by using Eqn (5), where $t_{50}^{(1)} < t_{50}^{(2)}$ are the two reaction times used to determine $I_{50}^{(1)}$ and $I_{50}^{(2)}$, respectively, and a is an empirical constant (see below).

$$K_\mathrm{i}^* = \frac{1 - \left(t_{50}^{(1)}/t_{50}^{(2)}\right)^{1/a}}{\dfrac{1}{I_{50}^{(1)}} - \dfrac{\left(t_{50}^{(1)}/t_{50}^{(2)}\right)^{1/a}}{I_{50}^{(2)}}} \tag{5}$$

$$k_\mathrm{inact} = \frac{1}{t_{50}^{(2)}} \exp\left[a \ln\left(\frac{K_\mathrm{i}^*}{I_{50}^{(2)}} - 1\right) + b\right] \tag{6}$$

$$k_\mathrm{eff} = \frac{k_\mathrm{inact}}{K_\mathrm{i}^*}\left(1 + \frac{[\mathrm{S}]_0}{K_\mathrm{M}}\right) \tag{7}$$

Once K_i^* is determined from Eqn (5), k_inact can be computed from it and from $I_{50}^{(2)}$ by using Eqn (6), where a and b are em-

pirical constants. The value of b depends on the units used to express time and concentration. When time is expressed in seconds and concentrations in micromoles per liter, $b = 0.558$. The value of $a = 0.9779$ irrespective of units. Eqns (5)–(6) are derived in the *Supporting Information*. Eqn (7) defines the second-order covalent efficiency constant in the context of the two-step mechanism **B**.

For routine calculations with real-world experimental data, inevitably affected by finite random noise, it is convenient to utilize a simplified version of Eqns (5)–(6) shown in Eqns (8)–(9). Here utilize the fact that the empirical constant $a = 0.9779$ is very nearly equal to unity, which means that $\left(t_{50}^{(1)}/t_{50}^{(2)}\right)^{1/a} \approx t_{50}^{(1)}/t_{50}^{(2)}$. Thus, after setting a to unity in Eqn (5) and multiplying both the numerator and denominator by $t_{50}^{(2)}$, we obtain Eqn (8).

$$K_{\mathrm{i}}^{*} \approx \frac{t_{50}^{(2)} - t_{50}^{(1)}}{\dfrac{t_{50}^{(2)}}{I_{50}^{(1)}} - \dfrac{t_{50}^{(1)}}{I_{50}^{(2)}}} \tag{8}$$

$$k_{\mathrm{inact}} \approx \frac{1}{t_{50}^{(2)}} \exp\left[\ln\left(\frac{K_{\mathrm{i}}^{*}}{I_{50}^{(2)}} - 1\right) + b\right] \tag{9}$$

2.2.4. *Implicit equation for I_{50} vs. time in mechanism* **B**

For validation purposes, the dependence of I_{50} on the reaction time was modeled by using the implicit algebraic Eqn (10), which is a minor variation of an equivalent implicit equation previously derived by Krippendorff *et al.* [10]. Given the values of k_{inact}, K_{i}^{*} and t_{50}, the iterative numerical solution to obtain the corresponding value of I_{50} was accomplished by using the Newton-Raphson method.[11]

$$0 = 1 - \exp\left(-\frac{I_{50}}{I_{50} + K_{\mathrm{i}}^{*}}\, k_{\mathrm{inact}}\, t_{50}\right) - \frac{I_{50}}{2\, K_{\mathrm{i}}^{*}}\, k_{\mathrm{inact}}\, t_{50} \tag{10}$$

2.2.5. *ODE model for covalent enzyme inhibition*

In the context of first-order ordinary differential-equation (ODE) modeling, the two-step inhibition mechanism **A** in *Figure 1* is mathematically represented by the ODE system defined by Eqns (11)–(17).

$$\begin{aligned} \frac{\mathrm{d[E]}}{\mathrm{d}t} &= -k_{1\mathrm{s}}[\mathrm{E}][\mathrm{S}] + (k_{-1\mathrm{s}} + k_{2\mathrm{s}})[\mathrm{E{\cdot}S}] \\ &\quad -k_{1}[\mathrm{E}][\mathrm{I}] + k_{-1}[\mathrm{E{\cdot}I}] \end{aligned} \tag{11}$$

$$\frac{\mathrm{d[S]}}{\mathrm{d}t} = -k_{1\mathrm{s}}[\mathrm{E}][\mathrm{S}] + k_{-1\mathrm{s}}[\mathrm{E{\cdot}S}] \tag{12}$$

$$\frac{\mathrm{d[E{\cdot}S]}}{\mathrm{d}t} = +k_{1\mathrm{s}}[\mathrm{E}][\mathrm{S}] - (k_{-1\mathrm{s}} + k_{2\mathrm{s}})[\mathrm{E{\cdot}S}] \tag{13}$$

$$\frac{\mathrm{d[P]}}{\mathrm{d}t} = +k_{2\mathrm{s}}[\mathrm{E{\cdot}S}] \tag{14}$$

$$\frac{\mathrm{d[I]}}{\mathrm{d}t} = -k_{1}[\mathrm{E}][\mathrm{I}] + k_{-1}[\mathrm{E{\cdot}I}] \tag{15}$$

$$\frac{\mathrm{d[E{\cdot}I]}}{\mathrm{d}t} = +k_{1}[\mathrm{E}][\mathrm{I}] - (k_{-1} + k_{2})[\mathrm{E{\cdot}I}] \tag{16}$$

$$\frac{\mathrm{d[EI]}}{\mathrm{d}t} = +k_{2}[\mathrm{E{\cdot}I}] \tag{17}$$

The ODE system defined by Eqns (11)–(17) was automatically generated by the software package DynaFit [8] from symbolic input. See the *Supporting Information* for details. The experimental signal was simulated according to Eqn (18), where F is the signal value, for example fluorescence intensity at time t, F_0 is the baseline offset (a property of the instrument), r_{P} is the molar response coefficient of the product P, and [P] is the product concentration at time t computed by solving the initial value problem defined by Eqns (11)–(17).

$$F = F_0 + r_{\mathrm{P}}\,[\mathrm{P}] \tag{18}$$

The microscopic rate constants k_1, k_{-1} and k_2 that we used in the simulation study described below can be related to the macroscopic kinetic constants as is shown in Eqns (19)–(21).

$$k_{\mathrm{inact}}^{(\mathrm{true})} = k_2 \tag{19}$$

$$K_{\mathrm{i}}^{(\mathrm{true})} = \frac{k_{-1} + k_2}{k_1} \tag{20}$$

$$k_{\mathrm{eff}}^{(\mathrm{true})} = \frac{k_1\, k_2}{k_{-1} + k_2} \tag{21}$$

2.2.6. *Determination of I_{50} from simulated signal*

The I_{50} values were determined by a fit of simulated florescence values to Eqn (22). The three adjustable model parameters were the control fluorescence intensity F_{c}, corresponding to zero inhibitor concentration; the I_{50}; and the Hill constant n. It was assumed that at asymptotically infinite inhibitor concentration the fluorescence signal is by definition equal to zero.

$$F = \frac{F_{\mathrm{c}}}{1 + \left(\dfrac{[\mathrm{I}]_0}{I_{50}}\right)^{n}} \tag{22}$$

3. Results

This section describes the results of a simulation study that was designed to validate the determination of the kinetic constants k_{inact}, K_{i}, and $k_{\text{cat}}/K_{\text{i}}$ from two measurements of I_{50}. First we present an illustrative example of an irreversible inhibitor following the one-step mechanism **C**. Next we demonstrate the newly proposed method on an inhibitor following the two-step rapid-equilibrium kinetic mechanism **B**. Finally, a summary of results is given for all simulated compounds.

3.1. Simulation study design

The simulation study was designed such that each of the three microscopic rate constants appearing in the kinetic mechanism **A** varied by three orders of magnitude, stepping by a factor of 10. The association rate constant k_1 varied from 10^4 to 10^7 $\text{M}^{-1}\text{s}^{-1}$; the dissociation rate constant k_{-1} varied from 0.001 to 1 s^{-1}; and the inactivation rate constant varied from 0.0001 to 0.1 s^{-1}. The corresponding covalent inhibition constants $K_{\text{i}} \equiv (k_{-1} + k_2)/k_1$ varied by six orders of magnitude from 110 pM to 110 μM; the second-order covalent efficiency constants $k_{\text{eff}} \equiv k_1\, k_2/(k_2 + k_{-1})$ varied by seven orders of magnitude from 0.99 $\text{M}^{-1}\text{s}^{-1}$ to 9.9×10^6 $\text{M}^{-1}\text{s}^{-1}$; the partition ratio k_2/k_{-1} varied by six orders of magnitude from $k_2/k_{-1} = 0.0001$ to $k_2/k_{-1} = 100$. Note that the partition ratio determines the extent to which the conventionally invoked rapid equilibrium approximation ($k_2/k_{-1} << 1$) is satisfied for any given compound. The corresponding "compound numbers" for the $4 \times 4 \times 4 = 64$ simulated inhibitors are summarized in *Table 1*.

k_1 $\mu\text{M}^{-1}\text{s}^{-1}$	k_{-1} s^{-1}	k_2 s^{-1} 0.1	0.01	0.001	0.0001
10	1	**1**	**17**	**33**	**49**
1	1	**2**	**18**	**34**	**50**
0.1	1	**3**	**19**	**35**	**51**
0.01	1	**4**	**20**	**36**	**52**
10	0.1	**5**	**21**	**37**	**53**
1	0.1	**6**	**22**	**38**	**54**
0.1	0.1	**7**	**23**	**39**	**55**
0.01	0.1	**8**	**24**	**40**	**56**
10	0.01	**9**	**25**	**41**	**57**
1	0.01	**10**	**26**	**42**	**58**
0.1	0.01	**11**	**27**	**43**	**59**
0.01	0.01	**12**	**28**	**44**	**60**
10	0.001	**13**	**29**	**45**	**61**
1	0.001	**14**	**30**	**46**	**62**
0.1	0.001	**15**	**31**	**47**	**63**
0.01	0.001	**16**	**32**	**48**	**64**

Table 1: "Compound numbers" attached to each of the 64 simulated inhibitors.

The assumed values of substrate rate constants appearing in *Figure 1* were $k_{1\text{s}} = 1\ \mu\text{M}^{-1}\text{s}^{-1}$, $k_{-1\text{s}} = 1\ \text{s}^{-1}$, and $k_{2\text{s}} = 1\ \text{s}^{-1}$. Thus, the corresponding Michaelis constant had the value $K_{\text{M}} \equiv (k_{-1\text{s}} + k_{2\text{s}})/k_{1\text{s}} = 2\ \mu\text{M}$. The simulated substrate concentration was $[\text{S}]_0 = 2\ \mu\text{M}$, such that the adjustment factor for competitive inhibition was $1 + [\text{S}]_0/K_{\text{M}} = 2$. Each dose-response data set consisted of 12 progress curves corresponding to 11 nonzero inhibitor concentrations, plus the positive control progress curve at $[\text{I}]_0 = 0$. The maximum inhibitor concentration was set to one fifth of the "true" covalent inhibition constant K_{i}. For example the maximum concentration for compound **5** ($k_1 = 10\ \mu\text{M}^{-1}\text{s}^{-1}$, $k_{-1} = 0.1\ \text{s}^{-1}$, and $k_2 = 0.1\ \text{s}^{-1}$) was set to $[\text{I}]_0 = 0.2 \times (0.1 + 0.1)/10 = 0.004\ \mu\text{M}$. The remaining 10 inhibitor concentrations represented a 1:2 dilution series; the resulting inhibitor concentration range spanned three orders of magnitude. The maximum inhibitor concentrations ranged from 22 μM to 2 nM. The simulated enzyme concentration was $[\text{E}]_0 = 1$ pM, which is lower by at least a factor of 2 than the minimum nonzero inhibitor concentration generated for any compound.

The simulated experimental signal was assumed to be directly proportional to the concentration of the reaction product, P. The assumed molar response coefficient of the enzymatic product was $r_{\text{P}} = 10000$ instrument units (for example, relative fluorescence units) per μM. Each simulated fluorescence value was perturbed by normally distributed pseudo-random error equal to 1% of the maximum simulated signal value. Experimental signal values were simulated at five different stopping point, t = 15, 30, 60, 120, and 240 min. Importantly, *only two of the simulated signal values (generated t = 30 min and t = 2 hr) were used for the kinetic analysis*. The remaining time points were used merely to verify qualitative systematic trends in the simulated data, but were otherwise ignored for the purpose of determining the kinetic constants k_{inact}, K_{i}, and/or $k_{\text{inact}}/K_{\text{i}}$.

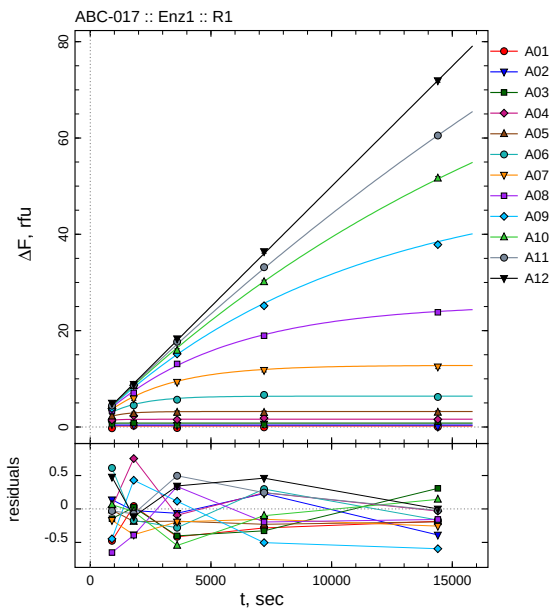

Figure 2: Raw experimental signal simulated for compound **17**. For details see text.

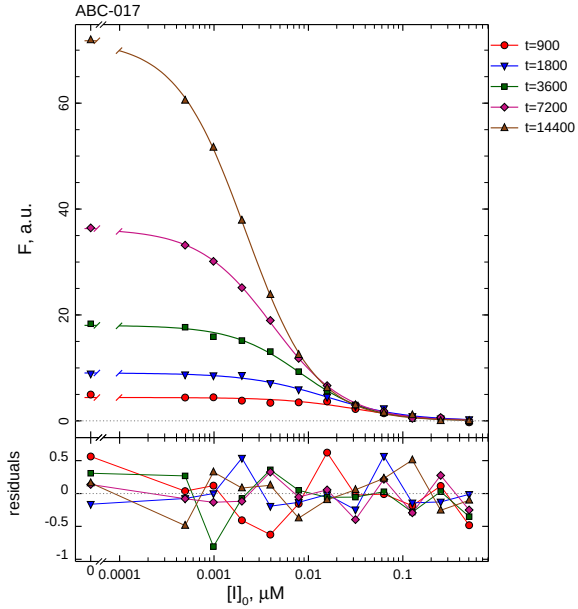

Figure 3: Results of fit of simulated experimental data for compound **17** (see *Figure 2*) to Eqn (22) to determine I_{50} values.

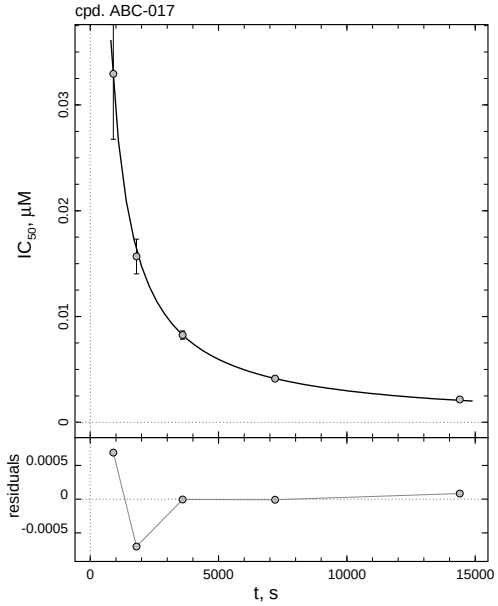

Figure 4: Results of fit of I_{50} results for compound **17** (see *Figure 3*) to the implicit Eqn (10).

3.2. *Example 1: One-step kinetic mechanism* **C**

Compound **17** (k_1 = 10 μM^{-1}s^{-1}; k_{-1} = 1 s^{-1}; k_2 = 0.01 s^{-1}) represents a typical example of a simulated inhibitor conforming to the one-step kinetic mechanism **C**, according to the present data-analytic procedure. The simulated reaction progress curves are shown in *Figure 2*, where the smooth theoretical model curves correspond to the numerical solution of the differential equation system Eqns (11)–(17), as described in section 2.2.5. The labels "A01" through "A11" in *Figure 2* correspond to inhibitor concentrations ranging from 505 nM to 0.49 nM (1:2 serial dilution); progress curve "A12" represents the positive control.

Each individual progress curve shown in *Figure 2* was fit separately to the three-parameter logistic Eqn (22). The best-fit values of the Hill constant n ranged from 1.01 to 1.17 for all 11 analyzed dose–response curves. Importantly, the best-fit values of I_{50} were *inversely proportional to the reaction time*, which immediate alerts us to the involvement of the one-step kinetic mechanism, as predicted by Eqn (1). The details are summarized in the *Supporting Information*. Briefly, at stopping times equal to 15, 30, 60, 120 and 240 minutes (increasing systematically by a factor of 2) the best-fit values of I_{50} were 33, 17, 8.2, 4.1, and 2.1 nM, again stepping approximately by a factor of 2, but in the opposite direction. Accordingly, as is predicted by the theoretical model specified by Eqn (1) for the one-step kinetic mechanism **C**, the calculated k_1^* value is largely invariant with respect to time. In particular, the five calculated k_1^* values were 53.8, 56.4, 53.7, 53.5, and 51.3 mM^{-1}s^{-1}, respectively (see *Supporting Information* for details).

For the purpose of the kinetic analysis, in this report we are purposely considering only two of the five stopping points, namely 30 min and 120 min. The geometric standard deviation for the two relevant values of k_1^* (in this case 56.4 and 53.5 mM^{-1}s^{-1}, respectively) was 1.04, which is lower than our acceptance criterion $\sigma_g < 1.25$. Thus, we accept as the final result the apparent k_1^* value corresponding to 120 min, in this case k_1^* = 53.5 mM^{-1}s^{-1}. This corresponds to $k_1 = k_1^*(1 + [S]_0/K_M)$ = 107 mM^{-1}s^{-1}. Similarly from the I_{50} value determined at t = 30 min, $k_1 = 56.3 \times (1 + 2/2)$ = 113 mM^{-1}s^{-1}. The "true" i.e. simulated value of the second-order covalent efficiency constant for compound **17** is $k_{\text{inact}}/K_{\text{i}} = k_{\text{eff}} = k_1\, k_2/(k_{-1} + k_2) = 10 \times 0.01/(1 + 0.01) = 0.099$ μM^{-1}s^{-1} = 99 mM^{-1}s^{-1}. Thus, the "true" value (99 mM^{-1}s^{-1}) and the two calculated values each based on a single determination of I_{50} (107 and 113 mM^{-1}s^{-1}, respectively) agree to within approximately ten to fifteen percent.

In conclusion, the covalent efficiency constant $k_{\text{inact}}/K_{\text{i}}$ for compound **17** could be determined from either of two I_{50} determinations, either at 30 minutes or two hours, using the simple formula represented by Eqn (1). Importantly, the fact that the two efficiency constant values are in good agreement ($\sigma_g <$ 1.25) provides an internal check on the underlying kinetic mechanism, in this case the one-step kinetic mechanism **C**.

3.3. *Example 2: Two-step kinetic mechanism* **B**

Compound **33** (k_1 = 10 μM^{-1}s^{-1}; k_{-1} = 1 s^{-1}; k_2 = 0.001 s^{-1}) represents a typical example of a simulated inhibitor conforming to the two-step kinetic mechanism **B**. Note that the only difference between compound **33** and compound **17** analyzed as Example 1 above is a ten-fold difference in the inactivation rate constant k_2. In the case of compound **17**, the inactivation rate constant was ten times higher (k_2 = 0.01 s^{-1}) compared to compound **33** (k_2 = 0.001 s^{-1}) . The association rate constant k_1 and the dissociation rate constant k_{-1} have identical values

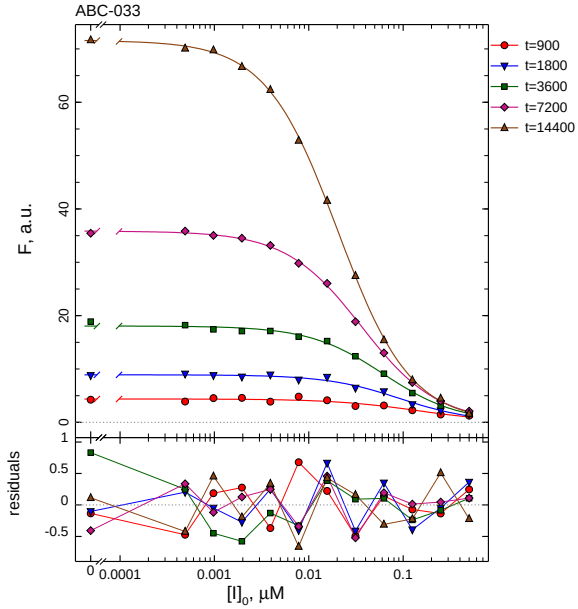

Figure 5: Results of fit of simulated experimental data for compound **33** to Eqn (22) to determine I_{50} values.

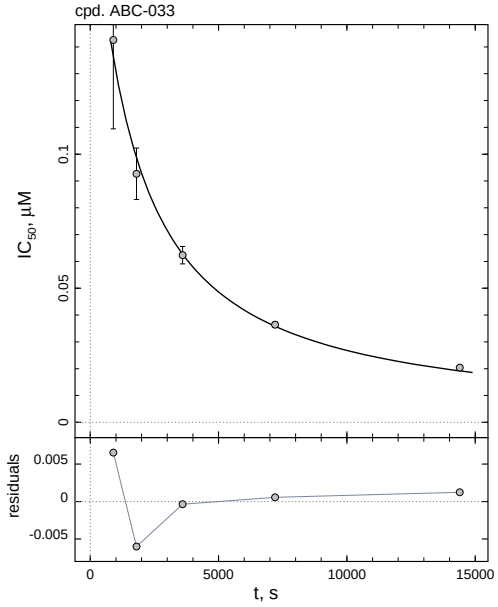

Figure 6: Results of fit of I_{50} results for compound **33** (see *Figure 5*) to the implicit Eqn (10).

for both compounds.

The simulated experimental signal was fit to the three-parameter logistic Eqn (22). The best-fit values of the Hill constant n ranged from 0.97 to 1.13 for all 11 analyzed progress curves. The best-fit values of I_{50} corresponding to each of the five different stopping times (15, 30, ..., 240 min) are displayed in *Figure 6*. Full details of the kinetic analysis are shown in the *Supporting Information*. Briefly, the best-fit values of I_{50} at stopping times 30 and 120 min were 92.7 and 36.4 nM, respectively, which corresponds to k_1^* values 9.6 and 6.1 $\mathrm{mM^{-1}s^{-1}}$ according to Eqn (1). The geometric standard deviation associated with these two numerical values is 1.38, which is higher than the cut-off acceptance criterion $\sigma_g < 1.25$. Therefore compound **33** is assigned the two-step kinetic mechanism **B**. The computation of the efficiency constant then proceeds in three consecutive steps. In the first step, we compute an estimate of the apparent inhibition constant using the simplified empirical Eqn (8), as follows:

$$K_\mathrm{i}^* = \frac{t_{50}^{(2)} - t_{50}^{(1)}}{\dfrac{t_{50}^{(2)}}{I_{50}^{(1)}} - \dfrac{t_{50}^{(1)}}{I_{50}^{(2)}}} = \frac{120 - 30}{\dfrac{120}{97.7} - \dfrac{30}{36.4}} = 191 \text{ nM}$$

After converting to micromoles per liter, $K_\mathrm{i}^* = 0.191\ \mu\mathrm{M}$. In the second step, we use the Eqn (6) to compute k_inact from K_i^* and one of the I_{50} value obtained at the later stopping time, in this case 7200 sec:[2]

$$\begin{aligned} k_\mathrm{inact} &= \frac{1}{t_{50}^{(2)}} \exp\left[a \ln\left(\frac{K_\mathrm{i}^*}{I_{50}^{(2)}} - 1\right) + b\right] \\ &= \frac{1}{7200} \exp\left[0.9779 \ln\left(\frac{0.191}{36.4} - 1\right) + 0.558\right] \\ &= 0.0009993\ \mathrm{s^{-1}} \end{aligned}$$

In the third and final step, we compute the covalent efficiency constant k_eff as a ratio of k_inact over K_i and simultaneously adjust both k_eff and K_i for the assumed kinetically *competitive* initial binding. Thus, $K_\mathrm{i} = K_\mathrm{i}^*/(1+[\mathrm{S}]_0/K_\mathrm{M}) = 0.0191/(1+2/2) = 0.0955\ \mu\mathrm{M}$ and $k_\mathrm{eff} = k_\mathrm{inact}/K_\mathrm{i} = 0.0105\ \mu\mathrm{M^{-1}s^{-1}}$. The "true" i.e. simulated value of the second-order covalent efficiency constant for compound **33** is $k_\mathrm{eff} = k_1\,k_2/(k_{-1} + k_2) = 10 \times 0.001/(1 + 0.001) = 0.0999\ \mu\mathrm{M^{-1}s^{-1}}$, which agrees within less than 5% with the calculated value $k_\mathrm{eff} = 0.0105\ \mu\mathrm{M^{-1}s^{-1}}$. The "true" i.e. simulated value of the covalent inhibition constant is $K_\mathrm{i} = (k_{-1} + k_2)/k_1 = (1 + 0.001)/10 = 0.1001\ \mu\mathrm{M}$, which agrees within less than 5% with the calculated value $K_\mathrm{i}' = 0.0955\ \mu\mathrm{M}$.

In conclusion, the covalent efficiency constant $k_\mathrm{inact}/K_\mathrm{i}$ for compound **33**, as well as the covalent inhibition constant K_i and the inactivation rate constant k_inact, could be determined from only two I_{50} determinations conducted at 30 minutes and two hours, using simple algebraic formulas that can be implemented in a common spreadsheet or a calculator. The theoretically expected and calculated values for all three kinetic constants differ by less than 10%.

[2] Recall that the empirical constants a and b in Eqn (6) strictly require that all concentrations be expressed in micromoles per liter and the reaction time is in seconds.

3.4. Summary of results for all 64 inhibitors

3.4.1. Assignment of kinetic mechanism

Preliminary investigations revealed that the assignment of the optimal kinetic mechanism (either one-step or two-step) to each inhibitor depends on the choice of stopping times $t_{50}^{(1)}$ and $t_{50}^{(2)}$, as well as on the choice of the empirical model-acceptance criterion σ_g. Using $t_{50}^{(1)} = 30$ min, $t_{50}^{(2)} = 120$ min, and $\sigma_g < 1.25$, the results are summarized in *Table 2*.

k_1 $\mu M^{-1}s^{-1}$	k_{-1} s^{-1}	k_2 s^{-1} 0.1	0.01	0.001	0.0001
10	1	C	C	B	B
1	1	C	C	B	B
0.1	1	C	C	B	B
0.01	1	C	C	B	B
10	0.1	C	C	C	B
1	0.1	C	C	B	B
0.1	0.1	C	C	B	B
0.01	0.1	C	C	B	B
10	0.01	C	C	B	B
1	0.01	C	C	B	B
0.1	0.01	C	C	B	B
0.01	0.01	C	C	B	B
10	0.001	C	C	C	B
1	0.001	C	C	C	B
0.1	0.001	C	C	C	B
0.01	0.001	C	C	C	B

Table 2: Kinetic mechanisms assigned to the simulated compounds using $t_{50}^{(1)} = 30$ min, $t_{50}^{(2)} = 120$ min, and $\sigma_g < 1.25$.

The results shown in *Table 2* can be summarized as follows. (i) Compounds **1** – **32**, characterized by a relatively *fast* chemical step with $k_2 \equiv k_{\text{inact}} \geq 0.01$ s^{-1}, were judged by the model selection algorithm to be following the *one-step* kinetic mechanism **C**, irrespective of the particular value of the partition ratio k_2/k_{-1}. (ii) Compounds **49** – **64**, characterized by a relatively *slow* chemical step with $k_2 \equiv k_{\text{inact}} \leq 0.0001$ s^{-1}, were judged by the model selection algorithm to be following the *two-step* kinetic mechanism **B**, again irrespective of the particular value of the partition ratio k_2/k_{-1}. (iii) Compounds **33** – **48** associated with an intermediate value of $k_2 = 0.001$ s^{-1} followed either of the two kinetic mechanisms, depending on the partition ratio k_2/k_{-1}. In particular, compounds **33** – **44** with the exception of **37** displayed the two-step kinetic mechanism **B**. In all those cases, the chemical step is slower than the dissociation step ($k_2/k_{-1} < 1$). In contrast, compounds **45** – **48**, for which $k_2/k_{-1} = 1$, conformed to the one-step kinetic mechanism **C**.

3.4.2. Calculated values of macroscopic kinetic constants

The calculated values of the second-order covalent efficiency constant k_{eff} for all 64 simulated inhibitors, as determined by the two-point method, are summarized in *Figure 7*. See *Supporting Information* for details. Note that the semantics of k_{eff} differs depending on the kinetic mechanism of inhibition (either **B** or **C**) assigned to each individual compound, as shown in *Table 2*. Thus, $k_{\text{eff}} = k_1$ for compounds that follow the one-step kinetic mechanism **C**, whereas $k_{\text{eff}} = k_{\text{inact}}/K_{\text{i}}$ for compounds that follow the two-step kinetic mechanism **B**.

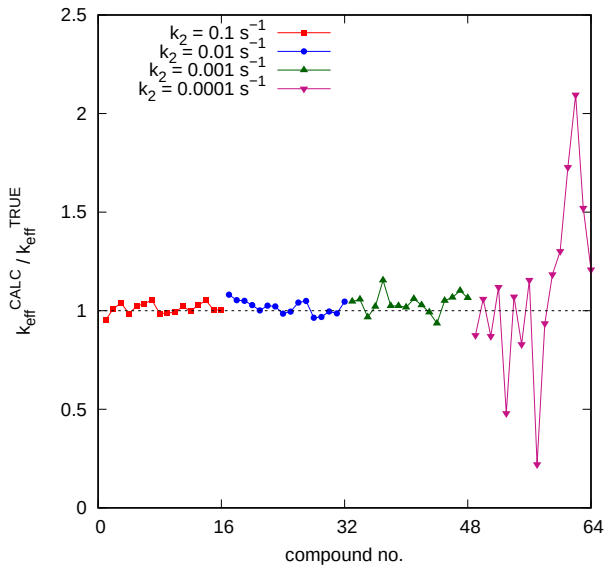

Figure 7: The ratios of calculated vs. "true" (i.e., simulated) values of k_{eff} a.k.a. $k_{\text{inact}}/K_{\text{i}}$. For details see text.

The results displayed in *Figure 7* indicate that k_{eff} is determined with better than approximately 25% accuracy for all inhibitors except compounds **49**– **64**, which are associated with very slow chemical inactivation step ($k_{\text{inact}} = k_2 = 0.0001$ s^{-1}). In the latter group of compounds, the ratio of the calculated over "true" i.e. simulated k_{eff} varies approximately from 0.2 to 2.0. Note that for all inhibitors following the two-step kinetic mechanism, the efficiency constant k_{eff} is computed after the fact, as a *ratio* of independently determined k_{inact} and K_{i} values. Thus, the question remains which of these two contributing factor, if any, is principally responsible for the lack of accuracy.

An explanation for the relatively large uncertainty of k_{eff} seen in most "slow" inactivators is presented in *Figure 8*, which shows that in most cases (in particular, compounds **49** – **60**) the inhibition constant K_{i} is determined quite accurately, whereas the inactivation constant k_{inact} shows a large degree of uncertainty. Only compounds **61** – **64** show a relatively large discrepancy between the "true" and calculated values for both k_{inact} and K_{i}. Note that compounds **61** – **64** are genuinely exceptional in two different respects. Not only their chemical reactivity is exceptionally low, as measured by $k_{\text{inact}} = k_2 = 0.0001$ s^{-1}, but also their dissociation rate constant $k_{-1} = 0.0001$ s^{-1} is the lowest in the entire compound collection. An examination of instantaneous rate plots for these four compounds (see *Supporting Information*) shows that the there is a kinetic transient (a "slow binding" phenomenon [12, 13]) that dramatically distorts the I_{50} values determined at $t = 30$ min.

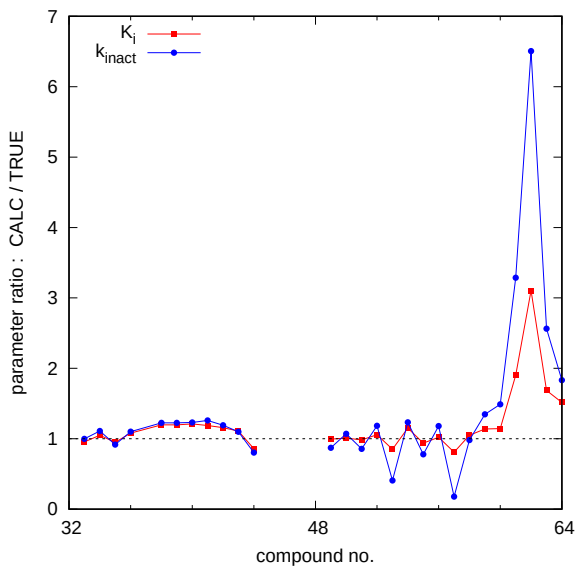

Figure 8: The ratios of calculated vs. "true" (i.e., simulated) values of k_{inact} and K_{i}. For details see text.

4. Discussion

Assumptions and limitations of the present method

The theoretical model represented by Eqns (1)–(7) is based on two simplifying assumptions. Note that the two assumptions are those that also underlie the Kitz–Wilson "k_{obs}" method [5, 14] and the Krippendorff method [10] of analyzing the time-dependence of I_{50} from multiple measurements. First, it is assumed that there is no inhibitor depletion, in the sense that during the assay only a negligibly small mole fraction of the inhibitor is bound to the enzyme, either covalently or non-covalently. This in turn implies that the total or analytic concentration of the inhibitor is always significantly higher than the initial concentration of the enzyme. In practical terms, we found that an approximately three fold excess of the lowest inhibitor concentration in a dilution series over the enzyme concentration is satisfactory. A situation that should be strenuously avoided is allowing any of the inhibitor concentrations become lower than the enzyme concentration, if and when those inhibitor concentrations are associated with any observable inhibitory effect.

Of course, depending on the nature of the assay, it may not be practically possible to lower the enzyme concentration sufficiently and still maintain assay sensitivity. For example, it may not be practically possible to use enzyme concentrations as low as $[\text{E}]_0 = 1$ pM, as was done in the simulation study presented here. In fact, in many assays it becomes necessary to use enzyme concentrations as high as $[\text{E}]_0 \approx 10$ nM or even higher, because of sensitivity concerns. However, note that the binding affinity of many therapeutically important enzyme inhibitors also lies in the nanomolar region, which means that these molecules express their inhibitory potency already at $[\text{I}]_0 \approx 1$ nM or lower. Neither the Kitz–Wilson "k_{obs}" method [5, 14], the Krippendorff method [10], nor the method presented here, can be used under such experimental circumstances, where the zero-inhibitor depletion assumption is violated. The only remedy is to deploy a data-analytic procedure that does not rely on any simplifying assumptions, meaning a mathematical model based on the numerical solution of differential equation. For an illustrative example involving the inhibition of drug-resistant EGFR mutants, see ref. [15].

The second simplifying assumption underlying the data-analytic procedure presented here is that the reaction rate in the positive control experiment ($[\text{I}]_0 = 0$) remains strictly constant over the entire duration of the assay. In other words, it is assumed that the positive control progress curve (time vs. experimental signal) is strictly linear. This can only be achieved if the mole fraction of the substrate ultimately consumed in the control assay remains negligibly low; if the initial concentration of the substrate is very much higher than the corresponding Michaelis constant ($[\text{S}]_0 >> K_{\text{M}}$); or if both of the above conditions are satisfied simultaneously. It should be noted that simple visual inspection can often be extremely misleading when it comes to judging linearity vs. nonlinearity of positive control assays. Instead of relying on a subjective assessment, it is preferable to deploy an objective *cross-validation procedure* described in ref. [16].

Experimental design

Krippendorff's [10] implicit algebraic Eqn (10) for time-dependence of I_{50}, as well as the data-analytic formulas derived in this report, are both based on the important assumption that there is *no preincubation of the enzyme with inhibitor prior to adding the substrate* to trigger the enzymatic assay. Instead, the enzyme's interactions with the substrate and with the inhibitor need to be initiated at the same time, by adding the enzyme catalyst as the last component into the assay.

At least some practitioners in enzyme kinetics apparently misunderstand this very important aspect of covalent I_{50} assays analyzed specifically by Krippendorff's method [10] (and also by the two-point method presented here). For example, Fassunke *et al.* [17] reported that "for kinetic characterization ($k_{\text{inact}}/K_{\text{i}}$), the inhibitors were incubated with EGFR-mutants over different periods of time (2–90 min), whereas durations of enzymatic reactions [25 min after adding the substrate at the end of enzyme–inhibitor preincubation, note added by P.K.] were kept constant. [...] Calculated IC_{50}-values were [...] fit as described in the literature to determine k_{inact} and K_{i}". The "literature" method mentioned immediately above is a reference to the Krippendorff method [10]. However, to repeat for emphasis, Krippendorff's equation Eqn (10) was derived under the assumption that the onset of product formation occurs simultaneously with the onset of enzyme inhibition, otherwise Eqn (10) cannot be used. On that basis, the k_{inact} and K_{i} values reported for EGFR inhibitors in ref. [17] are almost certainly invalid.

For the purposes of utilizing the newly proposed two-point I_{50} method, it is important to arrange the experiment such that the two stopping times are spaced sufficiently widely. Based on preliminary investigations, it appears sufficient to maintain at least a four-fold difference between $t_{50}^{(1)}$ and $t_{50}^{(2)}$, for example $t_{50}^{(1)} = 15$ min and $t_{50}^{(2)} = 1$ hr, or alternately 30 min and 2 hr. The objective is to assure that the two I_{50} values are sufficiently

different from each other, such that it becomes possible to discern whether or not the two resulting I_{50} values are inversely proportional to the stopping time according to Eqn (1). In this respect, it is advantageous to choose $t_{50}^{(2)}$ as high as is practically possible, while also keeping in mind that substrate depletion, enzyme deactivation, and other "nuisance" factors might cause the positive control progress curve to become nonlinear, which should be avoided as much as possible.

The optimal duration of the covalent inhibition assay is also closely related to the expected inactivation rate constant k_{inact}. In the hypothetical scenario where the enzyme is instantaneously saturated with the inhibitor because the inhibitor concentration is very much higher than the covalent inhibition constant K_{i}, the covalent conjugate EI is formed with the first-order rate constant k_{inact}. Under these hypothetical circumstances, the half-time for inactivation is equal to $\ln(2)/k_{\text{inact}}$. For example, in the specific case of $k_{\text{inact}} = 0.0001\ \text{s}^{-1}$, the expected half-time for inactivation is $t_{1/2} = \ln(2)/0.0001 = 0.693/0.0001 = 6930$ s = 115 min, or approximately two hours. Assuming that nearly full inactivation is achieved at about $t_{\max} \approx 3 \times t_{1/2}$, the assay would have to last almost six hours in order to see the enzyme fully inactivated. An enzyme assay that long of course might not be possible for numerous practical reasons, which also means that covalent inhibitors with $k_{\text{inact}} \leq 0.0001\ \text{s}^{-1}$ are exceedingly difficult to characterize accurately; see also the results reported here for simulated compounds **49** – **64**.

Choice of the model selection criterion σ_g

A successful application of the two-point data analytic procedure described in this report depends on a suitable choice of the model selection criterion σ_g. Recall that σ_g is the geometric standard deviation between two numerical values of I_{50}, defined by Eqn (4), and is used to decide between the one-step kinetic mechanism **C** and the two-step kinetic mechanism **B**. The optimal choice σ_g depends on the nature of assay and also on the choice of the two stopping times, $t_{50}^{(1)}$ and $t_{50}^{(2)}$. We found that for more closely spaced stopping time values, $\sigma_g \approx 1.25$ performs satisfactorily, whereas for stopping times separated by a factor of five or higher, $\sigma_g \approx 1.5$ works better. The optimal value of σ_g may need to be adjusted in the course of an inhibitor screening campaign, as practical experiences are being accumulated.

Similarities and differences with the Krippendorff method

The present method is similar to the method of Krippendorff *at al.* [10] in that both methods use the same theoretical foundation represented by Eqn (10). There are also three major differences. First, our method requires only two measurements of I_{50} whereas the Krippendorff method requires several times as many data points. Second, our method can be computationally implemented very simply by using a common spreadsheet program, whereas the Krippendorff method requires a highly specialized software package that allow nonlinear least-squares fit to an *implicitly* defined algebraic model, of the general form $f(X, Y) = 0$, as opposed to the much more common *explicit* algebraic equation, of the general form $Y = f(X)$. Third, and most important, the Krippendorff method is based on an *assumption* that all inhibitors follow the two-step kinetic mechanism **B**, whereas our method allows the actual kinetic mechanism to be detected from the experimental data, without making prior assumptions.

In fact, we have previously documented [6, 9] that covalent inhibitors characterized by high initial binding affinity, high chemical reactivity, or both, will outwardly display one-step kinetics. In the specific case of highly "tight binding" inhibitors [13, 18], which are characterized by relatively low dissociation rate constant k_{-1} in the initial noncovalent step, the noncovalent complex dissociates only very slowly on the time-scale of the experiment, which renders even the first (strictly speaking, noncovalent) binding step effectively irreversible. In the case of highly reactive covalent inhibitors, which are characterized by relatively high inactivation rate constant k_2, the initial noncovalent complex may be pulled forward through the reaction pathway so rapidly that the mole fraction of E·I remains kinetically undetectable. That is why covalent inhibitors characterized by relatively high inactivation constant ($k_2 \geq 0.01\ \text{s}^{-1}$) are not expected to yield a meaningful value of the dissociation constant K_{i}, even though the initial noncovalent complex must be physically present – however fleetingly.

Significance and utility of the two-point I_{50} method

The cost of successfully developing new medications and bringing them to market past unavoidable regulatory hurdles is enormous, amounting to approximately 2.6 billion US dollars per compound in 2016 [19]. Even assuming that the largest fraction of the overall expenditure is taken up by clinical trials, the cost of pre-clinical discovery processes such as the evaluation of enzyme inhibitors for biochemical potency is very significant, both in terms of human energy and in terms of material supplies.

In this context, irreversible enzyme inhibitors present an exceptional challenge, because even the "simple" task of meaningfully *ranking* a series of drug candidates by biochemical potency is complicated by the fact that the overall potency of covalent inhibitors consists of two entirely separate components, namely, their initial binding affinity (K_{i}) and their chemical reactivity (k_{inact}). This conceptual difficulty often leads to low-information experiments that are potentially wasteful.

For example, a number of drug discovery projects begin by testing each irreversible inhibitor in a "one-hour I_{50}" assay, or in a similar single time-point I_{50} assay. However, any value of covalent I_{50} observed at a single time-point is by definition non-unique, because it could be produced either by an inhibitor that has high affinity (low K_{i}) and low reactivity (low k_{inact}) or alternately by another inhibitor that has low affinity (high K_{i}) and high reactivity (high k_{inact}). Because of this inherent redundancy and ambiguity, a covalent I_{50} value determined at a single time-point cannot be used to rank irreversible inhibitors by potency in a meaningful way. In contrast, the two-point I_{50} method presented here is guaranteed to produce a *unique* value of the covalent efficiency constant k_{eff} for all inhibitors, irrespective of the kinetic mechanism, and additionally also a *unique* value of k_{inact} and K_{i} for those inhibitors that formally follow the two-step kinetic mechanism **B**.

Acknowledgements

I thank Dr. Claire McWhirter (Artios Pharma, Cambridge, UK) for helpful discussions.

Supporting information

1. File BioKin-TN-2020-03-SI.pdf: Details of algebraic derivations; listing of DynaFit script files; detailed kinetic analysis of all 64 simulated inhibitors.
2. File BioKin-TN-2020-03-SI.zip: Simulated pseudo-experimental data files.

References

[1] J. M. Strelow, A perspective on the kinetics of covalent and irreversible inhibition, SLAS Discovery 22 (2016) 3–22, doi:10.1177/1087057116671509.
URL http://doi.org/10.1177/1087057116671509

[2] Z. Zhao, P. E. Bourne, Progress with covalent small-molecule kinase inhibitors, Drug Disc. Today 23 (2018) 727–735.

[3] A. K. Ghosh, I. Samanta, A. Mondal, W. R. Liu, Covalent inhibition in drug discovery, ChemMedChem 14 (2019) 889–906.

[4] A. Abdeldayem, Y. S. Raouf, S. N. Constantinescu, R. Moriggl, P. T. Gunning, Advances in covalent kinase inhibitors, Chem. Soc. Rev. 49 (2020) 2617–2687, doi:10.1039/C9CS00720B.
URL http://doi.org/10.1039/C9CS00720B

[5] R. A. Copeland, Evaluation of Enzyme Inhibitors in Drug Discovery, 2nd Edition, John Wiley, New York, 2013.

[6] P. Kuzmič, Deciding between one-step and two-step irreversible inhibition mechanisms on the basis of "k_{obs}" data: A statistical approach, ChemRxiv. Preprint. (2020) doi:10.26434/chemrxiv.12407105.v1.
URL http://doi.org/10.26434/chemrxiv.12407105.v1

[7] P. Kuzmič, Program DYNAFIT for the analysis of enzyme kinetic data: Application to HIV proteinase, Anal. Biochem. 237 (1996) 260–273.

[8] P. Kuzmič, DynaFit - A software package for enzymology, Meth. Enzymol. 467 (2009) 247–280.

[9] P. Kuzmič, A steady-state algebraic model for the time course of covalent enzyme inhibition, ChemRxiv. Preprint. (2020) doi:10.26434/chemrxiv.12445394.v1.
URL http://doi.org/10.26434/chemrxiv.12445394.v1

[10] B.-F. Krippendorff, R. Neuhaus, P. Lienau, A. Reichel, W. Huisinga, Mechanism-based inhibition: Deriving K_I and k_{inact} directly from time-dependent IC_{50} values, J. Biomol. Screen. 14 (2009) 913–923.

[11] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical Recipes in C, Cambridge University Press, Cambridge, 1992.

[12] J. F. Morrison, C. T. Walsh, The behavior and significance of slow-binding enzyme inhibitors, Adv. Enzymol. Related Areas Mol. Biol. 61 (1988) 201–301.

[13] S. Szedlacsek, R. G. Duggleby, Kinetics of slow and tight-binding inhibitors, Meth. Enzymol. 249 (1995) 144–180.

[14] R. Kitz, I. B. Wilson, Esters of methanesulfonic acid as irreversible inhibitors of acetylcholinesterase, J. Biol. Chem. 237 (1962) 3245–3249.

[15] P. A. Schwartz, P. Kuzmič, J. Solowiej, S. Bergqvist, B. Bolanos, C. Almaden, A. Nagata, K. Ryan, J. Feng, D. Dalvie, J. Kath, M. Xu, R. Wani, B. W. Murray, Covalent egfr inhibitor analysis reveals importance of reversible interactions to potency and mechanisms of drug resistance, Proc. Natl. Acad. Sci. U.S.A. 111 (2014) 173–178.

[16] P. Kuzmič, J. Solowiej, B. W. Murray, An algebraic model for the kinetics of covalent enzyme inhibition at low substrate concentrations, Anal. Biochem. 484 (2015) 82–90.

[17] J. Fassunke, F. Mueller, M. Keul, S. Michels, M. A. Dammert, A. Schmitt, D. Plenker, J. Lategahn, C. Heydt, J. Braegelmann, H. L. Tumbrink, Y. Alber, S. Klein, A. Heimsoeth, I. Dahmen, R. N. Fischer, M. Scheffler, M. A. Ihle, V. Priesner, A. H. Scheel, S. Wagener, A. Kron, K. Frank, K. Garbert, T. Persigehl, M. Puesken, S. Haneder, B. Schaaf, E. Rodermann, W. Engel-Riedel, E. Felip, E. F. Smit, S. Merkelbach-Bruse, H. C. Reinhardt, S. M. Kast, J. Wolf, D. Rauh, R. Buettner, M. L. Sos, Overcoming EFGR G724S-mediated osimertinib resistance through unique binding characteristics of second-generation EFGR inhibitors, Nat. Commun. 9 (2018) 4655, doi:10.1038/s41467-018-07078-0.
URL http://doi.org/10.1038/s41467-018-07078-0

[18] J. W. Williams, J. F. Morrison, The kinetics of reversible tight-binding inhibition, Meth. Enzymol. 63 (1979) 437–467.

[19] J. A. DiMasi, H. G. Grabowski, R. W. Hansen, Innovation in the pharmaceutical industry: New estimates of R & D costs, J. Health Econ. 47 (2016) 20–33, doi:10.1016/j.jhealeco.2016.01.012.
URL http://doi.org/10.1016/j.jhealeco.2016.01.012