A steady-state algebraic model for the time course of covalent enzyme inhibition

Petr Kuzmič

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

Abstract

This report describes a double-exponential algebraic equation for the time course of irreversible enzyme inhibition following the two-step mechanism $E + I \rightleftharpoons E{\cdot}I \rightarrow EI$, under the steady-state approximation. Under the previously invoked rapid-equilibrium approximation [Kitz & Wilson (1962) *J. Biol. Chem.* **237**, 3245] it was assumed that the rate constant for the reversible dissociation of the initial noncovalent complex is very much faster than the rate constant for the irreversible inactivation step. The steady-state algebraic equation reported here removes any restrictions on the relative magnitude of microscopic rate constants. The resulting formula was used in heuristic simulations designed to test the performance of the standard rapid-equilibrium kinetic model. The results show that if the inactivation rate constant is significantly higher than the dissociation rate constant, the conventional "k_{obs}" method is incapable of correctly distinguishing between the two-step inhibition mechanism and a simpler one-step variant, $E + I \rightarrow EI$, even for inhibitors that have very high binding affinity in the reversible noncovalent step.

Key words: enzyme kinetics; inhibition; irreversible inhibition; covalent inhibition; steady-state approximation; mathematical model; algebraic model

1. Introduction

The standard algebraic method of fitting irreversible inhibition data [1, Chap. 9] is based on the simplifying assumption that the reversible formation of the initial noncovalent enzyme–inhibitor complex is essentially instantaneous on the time scale of the experiment. This assumption is based on the classic rapid-equilibrium approximation in enzyme kinetics [2], where it is assumed that the formation of the covalent conjugate is very much slower than the dissociation of the noncovalent complex into its constituent components. However, an examination of existing experimental results reveals that the typical values of rate constants for the covalent inactivation step [3, Fig. 63] are *not* significantly smaller than the typical dissociation rate constants of therapeutically relevant enzyme inhibitors [4]. Thus, the rapid-equilibrium approximation clearly does not hold in many experiments, in which covalent inhibitors are evaluated for potency as possible therapeutic agents. This means that the standard algebraic equations normally used to fit covalent kinetic data might not be appropriate in many cases.

One possible solution to this difficulty is to utilize mathematical models that are based on the numerical solution of systems of simultaneous first-order ordinary differential equations (ODEs). This approach avoids having to make any simplifying assumptions, either about the underlying inhibition mechanism, or about the magnitude of microscopic rate constants. For example, the software package DynaFit [5, 6], which implements a highly advanced numerical ODE solver algorithm [7], was used in the study of covalent inhibition of the EGFR kinase [8]. However, one significant disadvantage of mathematical models that rely on the numerical solution of ODE systems is that that the requisite numerical algorithms are highly complex by comparison with the closed-form algebraic equations. Most importantly, high quality ODE solving algorithms are implemented in only very few off-the-shelf software packages.

In this report we present a closed-form algebraic mathematical model that can be used either to simulate or to fit covalent inhibition data by using any generic software package as long as it allows the input of user-defined algebraic equations. The algebraic model presented here allows that the three microscopic rate constants that fully characterize the two-step covalent inhibition mechanism can have arbitrary values. Only two simplifying assumptions were used in the derivation. First, it is assumed that there is no inhibitor depletion i.e. "tight binding" [9]. This means that the concentration of the inhibitor is assumed to be very much higher than the concentration of the enzyme. The second simplifying assumption is that the uninhibited reaction rate is constant throughout the entire assay, meaning that the positive control progress curve (performed in the absence of the inhibitor) can be mathematically described as a straight line.

The algebraic model presented in this report can be used to gain understanding of the interplay between (a) the microscopic rate constants that define the two-step "slow binding" irreversible inhibition process and (b) experimentally observable measures of covalent inhibition potency, such as the inhibition constant (K_{i}), the inactivation rate constant k_{inact}, and the second-order covalent efficiency constant ($k_{\text{inact}}/K_{\text{i}}$).

2. Methods

This section describes the theoretical and mathematical methods that were used in heuristic simulation described in this re-

port. All computations were performed by using the software package DynaFit [5, 6].

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*

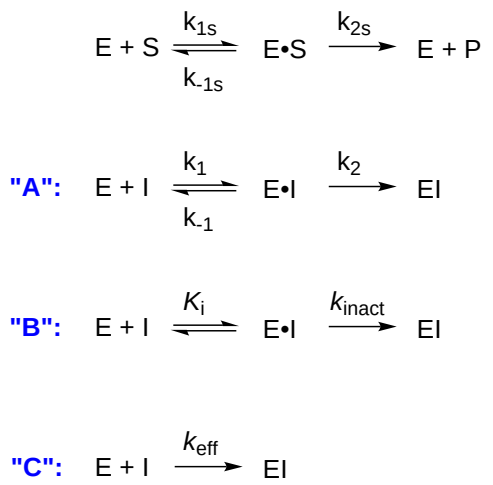

Figure 1: Kinetic mechanisms of substrate catalysis (top) and covalent inhibition (mechanisms **A** – **C**). For details see text.

The top reaction scheme in *Figure 1* represents the basic Michaelis-Menten reaction mechanisms [2]. In the inhibition mechanisms **A** through **C**, it is assumed that the covalent inhibitor I is kinetically *competitive* with the substrate S, because the inhibitor binds only to the free enzyme E and not to the Michaelis complex E·S.

Kinetic mechanisms **A** and **B** both include two consecutive steps, where E·I is a reversibly formed noncovalent initial complex. However, the theoretical assumptions underlying the two kinetic models are different. Mechanism **A** pertains to the steady-state approximation in enzyme kinetics, where the magnitudes of the microscopic rate constants k_1, k_{-1}, and k_2 can have any arbitrary values. In contrast, mechanism **B** is invoked under the rapid equilibrium approximation [10], where it is assumed that the inactivation rate constant k_2 is negligibly small compared to the dissociation rate constant k_{-1} and that the enzyme, inhibitor, and the noncovalent complex are always at equilibrium.

Kinetic mechanism **C** formally describes a direct formation of the irreversibly formed covalent conjugate EI. The second-order bimolecular rate constant k_{eff} (also known as "$k_{\text{inact}}/K_{\text{i}}$") is the covalent efficiency constant. Note that under mechanism **C** there is no distinction between the steady-state and rapid-equilibrium approximations, because the noncovalent initial complex E·I is absent.

2.2. Mathematical models

2.2.1. General mathematical model for the reaction progress

The progress of enzyme reactions is modeled here by using Eqn (1), where F is some experimental signal such as fluorescence intensity; F_0 is the experimental signal observed at time zero (i.e., a baseline signal as a property of the instrument); [P] is the concentration of product P at the reaction time t in some appropriate concentration units, such as micromoles or nanomoles per liter; and r_{P} is the molar response coefficient of the product under the given conditions. The molar response coefficient r_{P} is a proportionality constant that translates the product concentration to an experimentally observable signal, such as UV/Vis absorbance, fluorescence, or peak area, in appropriate instrument units.

$$F = F_0 + r_{\text{P}}\,[\text{P}] \tag{1}$$

2.2.2. Uninhibited substrate kinetics

In the absence of inhibitors, it is assumed that the product concentration changes over time according to the linear Eqn (4), where v_0 is the uninhibited initial rate according to Eqn (2) and t is the reaction time. The linearity of Eqn (4) implies that the uninhibited reaction rate v_0 stays effectively constant under the given experimental conditions. This in turn implies either that the initial substrate concentration is very much higher than the Michaelis constant K_{M} defined by Eqn (3); or that only a negligibly small fraction of the substrate S is converted to the product P at the end of the uninhibited assay; or that both of the above assumptions are satisfied.

$$v_0 = k_{2\text{s}}\,[\text{E}]_0\,\frac{[\text{S}]_0}{[\text{S}]_0 + K_{\text{M}}} \tag{2}$$

$$K_{\text{M}} = \frac{k_{-1\text{s}} + k_{2\text{s}}}{k_{1\text{s}}} \tag{3}$$

$$[\text{P}] = v_0\,t \tag{4}$$

2.2.3. Steady-state model for two-step covalent inhibition

In the presence of a covalent inhibitor following the steady-state mechanism **A**, and under the assumption of zero substrate conversion implied by the linear Eqn (4), the concentration of product P changes over time according to the double-exponential Eqn (5); the two exponential amplitudes a_1, a_2 and the two first-order rate constants r_1, r_1 are defined by Eqns (6)–(9), respectively.

$$[\text{P}] = \frac{v_0}{k^*_{\text{eff}}\,[\text{I}]_0}\left[1 - a_1 \exp(-r_1\,t) - a_2 \exp(-r_2\,t)\right] \tag{5}$$

$$a_1 = \frac{\alpha + \beta}{2\,\alpha} \tag{6}$$

$$a_2 = \frac{\alpha - \beta}{2\,\alpha} \tag{7}$$

$$r_1 = \frac{\gamma - \alpha}{2} \tag{8}$$

$$r_2 = \frac{\gamma + \alpha}{2} \tag{9}$$

$$\alpha = \sqrt{\left([\mathrm{I}]_0\, k_1^* + k_{-1} + k_2\right)^2 - 4\,[\mathrm{I}]_0 k_1^*\, k_2} \tag{10}$$

$$\beta = [\mathrm{I}]_0\, k_1^* \frac{k_{-1} - k_2}{k_{-1} + k_2} + k_{-1} + k_2 \tag{11}$$

$$\gamma = [\mathrm{I}]_0\, k_1^* + k_{-1} + k_2 \tag{12}$$

$$k_1^* = \frac{k_1}{1 + [\mathrm{S}]_0/K_\mathrm{M}} \tag{13}$$

$$k_{\mathrm{eff}}^* = \frac{k_1^*\, k_2}{k_{-1} + k_2} \tag{14}$$

The auxiliary variables α, β, γ, and k_1^* are defined by Eqns (10)–(13); t is the reaction time; $[\mathrm{I}]_0$ is the total or analytic concentration of the inhibitor, assumed to be effectively constant throughout the inhibition assay; and k_{eff}^* is the apparent covalent efficiency constant. The assumption of zero inhibitor depletion implies that the initial concentration of the inhibitor is very much higher than the initial concentration of the enzyme.

The sum of exponential amplitudes $a_1 + a_2$ is by definition equal to unity, because $(\alpha+\beta)/2\alpha+(\alpha-\beta)/2\alpha = 1$. Thus, in this sense a_1 and a_2 are relative amplitudes. The second exponential term, with amplitude a_2, decays faster than the first term, with amplitude a_1, because by definition α and γ are both positive and therefore $(\alpha + \gamma) > (\alpha - \gamma)$, which implies $r_2 > r_1$ for the two first-order rate constants.

The definition of k_1^* in Eqn (13) expresses the assumption that the inhibitor (I) is kinetically *competitive* with the substrate (S), in the sense that S and I bind to the same enzyme form, E. If the inhibitor happened to be kinetically *non-competitive* with the substrate, in the sense that the inhibitor would bind simultaneously and equally strongly to the free enzyme E and to the Michaelis complex E·S, the definition of k_1^* would change such that $k_1^* = k_1$. This situation could arise experimentally for example in covalent inhibition assays of protein kinases following an Ordered Bi-Bi catalytic mechanism [2], in which the inhibitor might be strictly kinetically competitive with ATP (i.e., noncompetitive with peptide substrate) but at the same time the assay might monitor the appearance of the phosphorylated peptide as opposed to ADP.

2.2.4. *ODE model for covalent enzyme inhibition*

In the context of differential-equation modeling, the two-step inhibition mechanism **A** in *Figure 1* is mathematically represented by the ODE system defined by Eqns (15)–(21).

$$\frac{\mathrm{d}[\mathrm{E}]}{\mathrm{d}t} = -k_{1\mathrm{s}}[\mathrm{E}][\mathrm{S}] + (k_{-1\mathrm{s}} + k_{2\mathrm{s}})[\mathrm{E{\cdot}S}] - k_1[\mathrm{E}][\mathrm{I}] + k_{-1}[\mathrm{E{\cdot}I}] \tag{15}$$

$$\frac{\mathrm{d}[\mathrm{S}]}{\mathrm{d}t} = -k_{1\mathrm{s}}[\mathrm{E}][\mathrm{S}] + k_{-1\mathrm{s}}[\mathrm{E{\cdot}S}] \tag{16}$$

$$\frac{\mathrm{d}[\mathrm{E{\cdot}S}]}{\mathrm{d}t} = +k_{1\mathrm{s}}[\mathrm{E}][\mathrm{S}] - (k_{-1\mathrm{s}} + k_{2\mathrm{s}})[\mathrm{E{\cdot}S}] \tag{17}$$

$$\frac{\mathrm{d}[\mathrm{P}]}{\mathrm{d}t} = +k_{2\mathrm{s}}[\mathrm{E{\cdot}S}] \tag{18}$$

$$\frac{\mathrm{d}[\mathrm{I}]}{\mathrm{d}t} = -k_1[\mathrm{E}][\mathrm{I}] + k_{-1}[\mathrm{E{\cdot}I}] \tag{19}$$

$$\frac{\mathrm{d}[\mathrm{E{\cdot}I}]}{\mathrm{d}t} = +k_1[\mathrm{E}][\mathrm{I}] - (k_{-1} + k_2)[\mathrm{E{\cdot}I}] \tag{20}$$

$$\frac{\mathrm{d}[\mathrm{EI}]}{\mathrm{d}t} = +k_2[\mathrm{E{\cdot}I}] \tag{21}$$

The ODE system defined by Eqns (15)–(21) was automatically generated by the software package DynaFit [6] from symbolic input. See the *Supporting Information* document for details.

3. Results

According to Cobelli *et al.* [11], "[t]he notion of *identifiability* addresses the question of whether it is at all possible to obtain unique solutions for unknown parameters of interest in a mathematical model, from data collected in well defined stimulus-response experiments performed on a dynamic system represented by the model." *Structural* identifiability analysis [12] is concerned with idealized data, completely free of the error, whereas *practical* identifiability analysis takes the inevitable random error into account [13].

In this section we first address both the structural and the practical identifiability of the newly derived algebraic model for covalent inhibition. We then use this model to conjure up a hypothetical irreversible inhibitor that has very strong initial binding affinity, as measured by the dissociation equilibrium constant K_i of the initial noncovalent complex, and yet apparently follows the "one-step" kinetic mechanism **C**, as if the complex E·I were absent.

3.1. *Structural identifiability analysis*

The results of structural identifiability analysis are illustrated in *Figure 2*. Idealized, noise-free data were simulated by using the following values of model parameters in Eqn (5): $V_0 = 0.0005$ RFU/sec;[1] $k_1^* = 0.5\ \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$; $k_{-1} = 0.001\ \mathrm{s}^{-1}$; and $k_2 = 0.01\ \mathrm{s}^{-1}$. Note that $k_2/k_{-1} = 10$, meaning that that

[1] RFU stands for relative fluorescence units, but in the more general case it could represent any other appropriate instrument unit such as UV/Vis absorbance units, chromatographic peak areas, radioactive counts, etc.

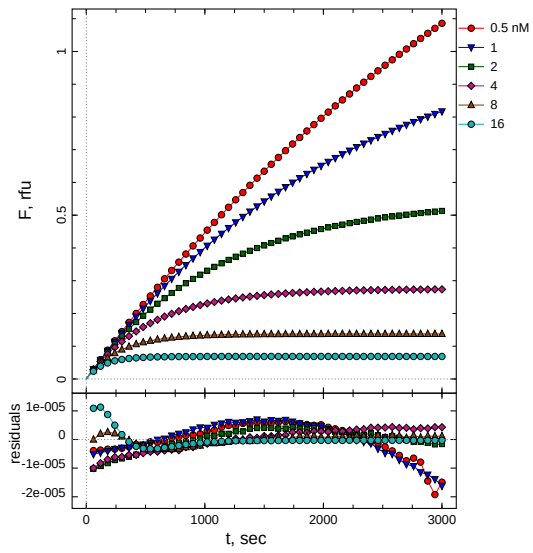

Figure 2: Results of structural identifiability analysis. Idealized, noise-free data were simulated by using the algebraic Eqn (5) and fit by using the ODE system Eqns (15)–(21)

rapid-equilibrium approximation (requiring $k_2 << k_{-1}$) does not hold. The inhibitor concentrations were 0.5, 1, 2, 4, 8, and 16 nM. The simulated time coordinates were 0, 60, 120, 180, ..., 3000 seconds (50 simulated time points, stepping by one minute). The artificial data were subsequently subjected to a global fit [14] to an ODE model defined by Eqns (15)–(21). In the ODE model, the fixed parameters were $[E]_0 = 1$ pM, $[S]_0 = 1\ \mu$M, $k_{1s} = 10\ \mu M^{-1}s^{-1}$, $k_{-1s} = 9.9\ s^{-1}$, $k_{2s} = 0.1\ s^{-1}$, and $r_P = 10^4$ RFU/μM. Accordingly, the Michaelis constant value was $K_M = (0.1 + 9.9)/10 = 1\ \mu$M. The globally optimized model parameters were the molar response coefficient r_P and the rate constants k_1, k_{-1} and k_2; locally optimized[2] model parameters were the eight baseline offsets, F_0, fit separately for each progress curve. DynaFit [6] scripts that were used for the simulation and for the fitting are listed in the *Supporting Information* document.

The best-fit values of the globally optimized model parameters and the associated formal standard error were as follows: $k_1 = (1.0012 \pm 0.0003)\ \mu M^{-1}s^{-1}$; $k_{-1} = (0.00102 \pm 0.00001)\ s^{-1}$; $k_2 = (0.01011 \pm 0.00003)\ s^{-1}$; and $r_P = (10000.10 \pm 0.04)$ RFU/μM. The expected best-fit value of k_1 was 1 $\mu M^{-1}s^{-1}$, because the simulated value of $k_1^* = k_1/(1 + [S]_0/K_M)$ was 0.5 $\mu M^{-1}s^{-1}$ and the adjustment factor $1 + [S]_0/K_M$ in this case is equal to $1 + 1/1 = 2$. Thus, the fitted and theoretically expected values of all adjustable model parameters agree within five significant digits. The best-fit model curves also agree with the simulated data within five significant digits, as is illustrated in the residual plot shown as the bottom panel of *Figure 2*. The very small systematic discrepancies between the algebraic and ODE models, shown in the slightly non-random distribution of the residuals of fit, are due to the inevitable propagation of round-off and truncation errors.

The two main conclusions that can be reached from the results of this heuristic simulation study are as follows. First, the theoretical model represented by Eqn (5) is algebraically correct, because it is congruent with a fully independent mathematical representation provided by the numerical solution of an equivalent ODE system. Second, the algebraic model is structurally identifiable with respect to all three microscopic rate constants that appear in the inhibition mechanism **A**. Thus, in the purely hypothetical case of having access to entirely noise-free experimental data, it would always be possible to determine all three rate constants k_1, k_{-1} and k_2 by performing a global least-squares fit of combined reaction progress curves similar to those shown in *Figure 2*.

3.2. *Practical identifiability analysis*

In the practical identifiability study the roles of the algebraic vs. numerical models were reversed. Artificial data were simulated according to the differential-equation system Eqns (15)–(21), this time with a finite level of experimental noise added to the results. The pseudo-experimental data were subsequently fit by using the algebraic model represented by Eqn (5). The requisite DynaFit [6] input scripts are listed in the *Supporting Information* document.

The simulated concentration plot in the bottom panel of *Figure 3* shows that at $[I]_0 = 80$ nM the noncovalent complex E·I has a significant presence in the evolving reaction mixture and that this noncovalent complex is formed *gradually* on the time scale of the simulated experiment. These observations suggest that there is a good chance of successfully determining all three rate constants that appear in the two-step inhibition mechanism **A**.

Simulated data shown in *Figure 3* were globally fit to the algebraic model represented by Eqn (5). The overlay of the best-fit model curves on the simulated data points was visually indistinguishable from the display of the simulated shown in *Figure 3*. The residual plots were distributed randomly, similar to the residual plot shown in *Figure 3*, upper panel. The best-fit values of globally adjustable model parameters are listed in *Table 1*. The columns labeled "low" and "high" are lower and upper limits, respectively, of non-symmetrical confidence intervals obtained by the profile-*t* method [15, 16] at 5% ΔSSQ level according to the empirical cut-off criterion advocated by Johnson [17, 18].

parameter	best-fit	±	std. err.	low	high
V_0, RFU/s	0.0005011	±	0.000002		
k_1^*, $\mu M^{-1}s^{-1}$	0.046	±	0.002	0.040	0.056
k_{-1}, s^{-1}	0.0007	±	0.0001	0.0004	0.0014
k_2, s^{-1}	0.0008	±	0.0001	0.0003	0.0012

Table 1: Results of practical identifiability analysis. For details see text.

[2] Local optimization means that the optimized parameter is adjustable in the fitting model such that the best-fit value is specific only to a subset of experimental data points, such as in this case each individual progress curve.

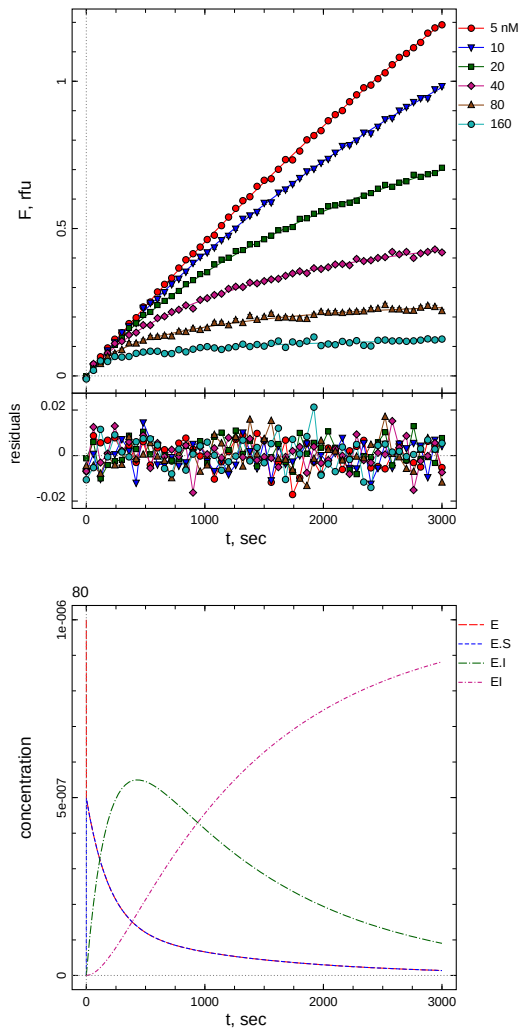

Figure 3: Artificial data simulated by using the ODE system Eqns (15)–(21). **Top:** Pseudo-experimental data (symbols) and the corresponding idealized model curves. **Bottom:** Enzyme species concentrations evolving over time at $[\mathrm{I}]_0 = 80$ nM. The "concentration" axis is in μM units.

The results displayed in *Table 1* show that all three microscopic rate constants appearing in the steady-state two-step covalent inhibition mechanism **A** could be reliably determined from the simulated data. The non-symmetrical confidence interval for the apparent association rate constant k_1^* spanned from 0.046 to 0.056 μM, while the theoretically expected value of $k_1^* = k_1/(1 + [\mathrm{S}]_0/K_\mathrm{M}) = 0.05$ μM. Similarly, the confidence intervals for k_{-1} and k_2 were relatively narrow (high/low ratios were approximately equal to 4) and encompassed the simulated values (0.001 s^{-1} in both cases). The best-fit values of k_{-1} and k_2 were only 20% to 30% lower than the simulated values.

The main conclusion is that, for at least some combinations of microscopic rate constants appearing in the two-step mechanism **A**, all three rate constants (k_1, k_{-1}, and k_2) can be determined in ordinary kinetic measurements, such as those that are typical for plate-reader assays usually performed in preclinical inhibitor screening.

3.3. *"One-step" kinetics of a high-affinity inhibitor*

Pseudo-experimental data were simulated by using the following values of model parameters in Eqn (5): $V_0 = 0.0005$ RFU/sec; $k_1^* = 0.5$ $\mu\mathrm{M}^{-1}\mathrm{s}^{-1}$; $k_{-1} = 0.001$ s^{-1}; and $k_2 = 0.01$ s^{-1}. Note that those are the same parameters that were used in the structural identifiability analysis, as described in section 3.1. However, in this case the simulated signal was perturbed by adding a Gaussian-distributed random noise with the standard deviation equal to 0.5% of the maximum simulated value. Each of the simulated progress curves were fit individually and separately to the standard algebraic model [1, sect. 9.1] for the time course of covalent inhibition, represented by the exponential Eqn (22), where v_i is the initial reaction rate in instrument units and k_obs is the apparent first-order rate constant corresponding to each inhibitor concentration.

$$F = F_0 + \frac{v_\mathrm{i}}{k_\mathrm{obs}}\left[1 - \exp\left(-k_\mathrm{obs}\, t\right)\right] \tag{22}$$

The results of fit are shown graphically in *Figure 4*. Note that the residuals of fit in the bottom panel are distributed randomly, which means that the *single-exponential* Eqn (22) is an adequate fitting model for this data, even though the data were simulated on the basis of a *double-exponential* Eqn (5). The best-fit values of the apparent first-order rate constant k_obs obtained at each inhibitor concentration are collected in *Table 2*.

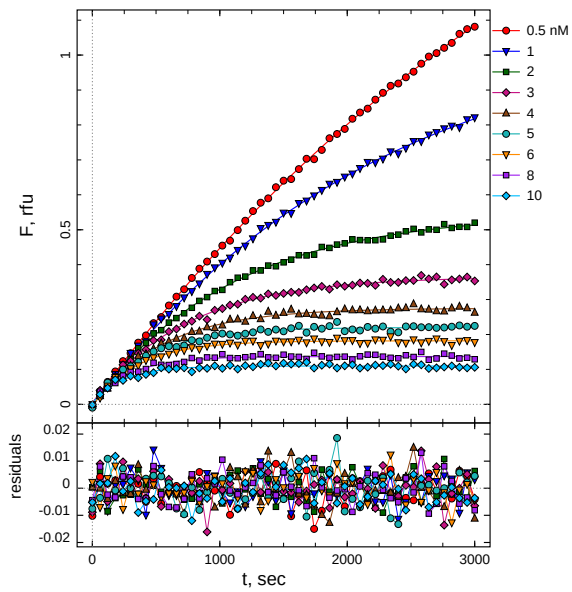

Figure 4: Results of fit of each individual progress curve simulated by using Eqn (5), with parameter values listed in the text, to Eqn (22), in order to determine the k_obs values associated with each inhibitor concentration listed in the right margin.

$[\mathrm{I}]_0$, nM	$1000 \times k_{\mathrm{obs}}$, s^{-1} ± std. err.
0.5	0.225 ± 0.006
1	0.458 ± 0.008
2	0.902 ± 0.016
3	1.327 ± 0.027
4	1.806 ± 0.045
5	2.280 ± 0.070
6	2.715 ± 0.101
8	3.485 ± 0.188
10	4.561 ± 0.320

Table 2: Best-fit values of the apparent first-order rate constant k_{obs} obtained by fitting the progress curves shown in *Figure 4* to the exponential Eqn (22).

The k_{obs} results collected in *Table 2* were subjected to a model discrimination analysis according to the procedure described in ref. [19], considering Eqn (23) and Eqn (24) as the two candidate fitting models, according to the standard treatment described in ref. [1, sect. 9.1] and elsewhere. The hyperbolic Eqn (23), corresponding to the two-step mechanism **B**, could be reliably excluded from consideration, using four independent statistical model selection criteria [19]. The preferred model was the linear Eqn (24), corresponding to the one-step mechanism **C**. The fit to the linear Eqn (24) produced a very well defined value of $k_{\mathrm{eff}} = (0.4496 \pm 0.0030)\ \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$, with the 95% confidence interval computed by the profile-t method [15, 16] spanning from 0.4426 to 0.4566 $\mu\mathrm{M}^{-1}\mathrm{s}^{-1}$. The results of fit to the linear model regression model are shown in *Figure 5*.

$$k_{\mathrm{obs}} = k_{\mathrm{inact}} \frac{[\mathrm{I}]_0}{[\mathrm{I}]_0 + K_{\mathrm{i}}^*} \qquad \text{mechanism } \mathbf{B} \tag{23}$$

$$k_{\mathrm{obs}} = k_{\mathrm{eff}}^* \, [\mathrm{I}]_0 \qquad \text{mechanism } \mathbf{C} \tag{24}$$

The values of microscopic rate constants used in simulating the artificial data shown in *Figure 4* were $k_1^* = 0.5\ \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$, $k_{-1} = 0.001\ \mathrm{s}^{-1}$ and $k_2 = 0.01\ \mathrm{s}^{-1}$. The apparent covalent efficiency constant is defined as $k_{\mathrm{eff}}^* = k_1^* \times k_2/(k_{-1} + k_2)$. Accordingly, the "true" value of the covalent efficiency constant was $k_{\mathrm{eff}}^* = 0.5 \times 0.01/(0.001 + 0.01) = 0.454545\ \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$. Thus the best-fit value $k_{\mathrm{eff}}^* = (0.455 \pm 0.003)\ \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$ and the "true" value $k_{\mathrm{eff}} = 0.454\ \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$ are in good agreement. Importantly, this agreement between the "true" and best-fit values of $k_{\mathrm{eff}}^* \equiv$ "$k_{\mathrm{inact}}/K_{\mathrm{i}}^*$" holds even though the individual values of k_{inact} and K_{i}^* could not be determined from the simulated data.

4. Discussion

Challenges in evaluating covalent inhibitors as potential drugs

Reliable evaluation of irreversible enzyme inhibitors as potential therapeutic agents is exceptionally challenging for the biochemical data analyst engaged in drug discovery. The reason is that the overall potency of covalent inhibitors consists of two separate and yet intertwined contributions. First, the inhibitor's binding affinity is measured by the inhibition constant, K_{i}. Second, the inhibitor's chemical reactivity is measured by the inactivation rate constant k_{inact}. However, we have previously documented that at least some inhibitors currently being prescribed as experimental anti-cancer drugs effectively follow the one-step kinetic mechanism **C** [19]. In those cases, the only available measure of potency is the efficiency constant k_{eff}, also known as "$k_{\mathrm{inact}}/K_{\mathrm{i}}$", which blends together both affinity and reactivity such that those two distinct molecular properties can no longer be evaluated separately.

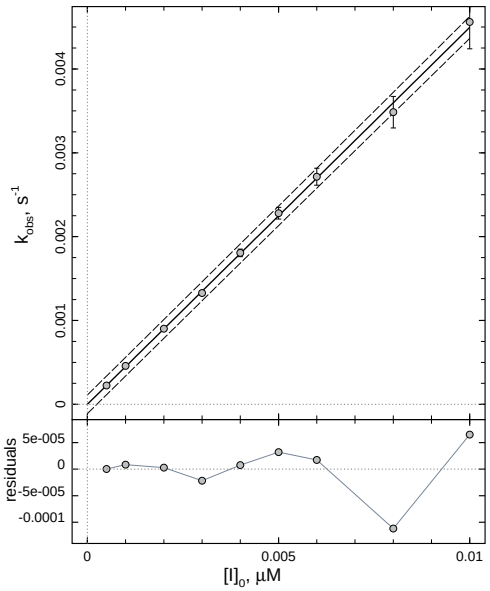

Figure 5: Results of fit of the k_{obs} values listed in *Table 2* to Eqn (24), corresponding to the one-step covalent inhibition mechanism **C**. For further details see text.

Irreversible inhibitors express their potency in a dynamic fashion, in the sense that the residual enzyme activity decreases over time, along with the gradual evolution of the permanent covalent bond between the enzyme and the inhibitor. Thus, in order to evaluate the potency of covalent inhibitors, we need mathematical models that describe the gradual formation of the final reaction product while appropriately taking into account that the rate of product formation (i.e., enzyme activity) inevitably decreases over time. As a result, all mathematical models for the progress of covalent inhibition assays are by definition nonlinear. This complexity presents an additional challenge when compared with measuring the potency of noncovalent inhibitors, where the assay in many cases can be arranged such that the reaction progress is nearly linear.

Existing models for the progress of covalent inhibition

Currently existing nonlinear regression models for the progress of covalent inhibition assays can be divided into two categories,

according to the mathematical formalism involved. In the first category are highly advanced *differential equation* models, which eliminate any simplifying assumptions about the relative magnitude of microscopic rate constants, such as the rapid-equilibrium approximation [2]; about the relative concentration of reactants, such as requiring a very large excess of inhibitor over enzyme; or about the reaction conditions, such as assuming strict linearity of the positive control progress curve. One important disadvantage of these ODE models is that they require highly specialized software algorithms for the numerical (i.e., iterative) solution of ODE systems. Only very few off-the-shelf software packages can be used for this purpose.

In the second category of mathematical models are *algebraic equations*, such as Eqn (22) originally derived by Kitz & Wilson [10], where the definition of k_{obs} is given by Eqns (23)–(24). The Kitz–Wilson equation Eqn (22) applies only to covalent inhibition assays where the uninhibited positive control reaction proceeds at a strictly constant rate, meaning that the plot of product concentration over time is linear. We have previously described a closely related algebraic model that allows for the control assay to be exponential, as opposed to linear [20]. On the one hand, these algebraic models have the major advantage that they can be implemented in any software system that allows the investigator so specify an arbitrary algebraic equation. On the other hand, both of these algebraic equations are based on the *rapid-equilibrium approximation* in enzyme kinetics [2]. Accordingly, it is assumed the chemical inactivation step (rate constant k_2 in *Figure 1*) is very much slower than the dissociation of the noncovalent complex (rate constant k_{-1}).

In the specific case of the Kitz–Wilson single-exponential model represented by Eqn (22), Cornish-Bowden [21, sec. 7.2.2] pointed out that "if k_2 is not small enough to allow formation of E·I to be treated as an equilibrium [...] the loss of activity does not follow simple first-order kinetics: there is no exact analytical solution, but the kinetics may still be analyzed by numerical methods." Translated into non-mathematical language, this brief quote contains three highly relevant statements. First, if the microscopic rate constant k_2 (inactivation) happens to be larger in magnitude than the microscopic rate constant k_{-1} (dissociation), the single-exponential ("first-order") Eqn (22) in principle cannot be used as a fitting model for the reaction progress, and therefore any k_{obs} results derived from that equation are by implication invalid. Second, if we suspect that the chemical reaction is not rate-limiting (k_2 not being "small enough") we should resort to differential equation modeling ("numerical methods") instead of relying on Eqn (22). Third, the author is suggesting that it is in principle impossible to derive an algebraic equation ("analytical solution") that would be applicable under the steady-state approximation ($k_2 > k_{-1}$), as opposed to the rapid-equilibrium approximation ($k_2 << k_{-1}$) implied by Eqn (22).

A need for the newly derived steady-state algebraic model

Cornish-Bowden [21, sec. 7.2.2] stated that a steady-state (i.e., $k_2 > k_{-1}$) algebraic model for the progress of covalent inhibition assays cannot be derived as a matter of principle ("there is no exact analytical solution", as quoted above). However, that statement happens to be incorrect, because the required analytic solution is actually presented in this report, as the double-exponential Eqn (5). Why is it important to have at our disposal a closed-form algebraic model for the time-course of covalent inhibition assays under the steady-state approximation? There are at least two important reasons, which are now addressed in their turn.

The first and most important need for the algebraic model newly presented in this report is that the simplistic rapid-equilibrium approximation ($k_2 << k_{-1}$) is almost certainly violated in many practically relevant cases. This is especially true for the most promising drug candidates. The *Kinetics for Drug Discovery* (K4DD) project [4] revealed that the dissociation rate constants for therapeutically relevant enzyme inhibitors are frequently in the range corresponding to hour-long drug–target residence times, which implies $k_{-1} < 0.0001\ \text{s}^{-1}$ or even $k_{-1} < 0.00001\ \text{s}^{-1}$ in many cases. On the other hand, Abdeldayem *et al.* [3] reported that the large majority of covalent inhibitor drugs and drug candidates are associated with inactivation rate constants in the range from approximately $k_2 = 0.0001\ \text{s}^{-1}$ to $k_2 = 0.01\ \text{s}^{-1}$. Therefore, assuming that the initial (noncovalent, reversible) binding affinity of covalent inhibitors is reasonably similar to the binding affinity of their noncovalent structural analogs, we can conclude that the typical values of dissociation rate constants are *not* very much larger than the covalent inactivation rate constants, as is required by the rapid-equilibrium approximation. Thus, the Kitz–Wilson single-exponential equation Eqn (22) for the reaction progress is very likely to be invalid, especially in the case of highly potent covalent drugs.

The second reason to have available an algebraic equation as the theoretical model for the progress of steady-state covalent inhibition assays is convenience and portability. The newly derived algebraic Eqn (5) can be implemented even in general-purpose software systems such as in Microsoft Excel, as opposed to requiring highly specialized ODE solving algorithms that are only available in very few software packages. In fact, a relevant Microsoft Excel template file is attached as one of the *Supporting Information* documents to this report.

Simplifying assumptions and limitations underlying Eqn (5)

As a reminder, the double-exponential equation Eqn (5) was derived under two simplifying assumptions. First, it is assumed that the inhibitor is not being depleted in the course of the assay. This implies that the initial concentration of the inhibitor is very much larger than the initial concentration of the enzyme. Second, it is assumed that the positive control assay, performed under the same experimental conditions but in the absence of inhibitors, proceeds with a constant rate. This implies that the positive control progress curve can be described mathematically as a straight line. Thus, the only difference in theoretical assumptions underlying the classic Kitz–Wilson single-exponential Eqn (22) and the double-exponential Eqn (5) is that the newly derived algebraic model presented here allows for all three rate constants in mechanism **A** to have entirely arbitrary values.

The utility of the double-exponential steady-state model Eqn (5) lies predominantly in the ability to *simulate* artificial data

set, which can be subsequently used to test the performance of the various established method of evaluating the potency of covalent inhibitors. A typical example is represented by the simulation study described in section 3.3, where we tested the performance of the classic "k_{obs} method". The steady-state model does theoretically allow for all three rate constants k_1, k_{-1}, and k_2, to have any arbitrary values. However, preliminary experience specifically in the *data-fitting* context strongly suggests that it is possible to actually determine all three microscopic rate constants from realistically "noisy" data only for certain special combinations of rate constants. In the majority of cases that were examined, it was possible to determine only either two rate constant, or even just one microscopic rate constant, from the simulated pseudo-experimental data. These investigations are currently ongoing in a more systematic fashion and the final results will be reported elsewhere.

"One-step" kinetics of high-affinity covalent inhibitors

One of the principal motivations for addressing the fundamental difference between the steady-state approximation and the rapid-equilibrium approximation in the analysis of covalent inhibition data has been the puzzling observation that many highly potent covalent inhibitors apparently follow the one-step kinetic mechanism **C**. The inhibition of certain protein kinases by ibrutinib represents a typical example [19, 22]. The occurrence of the one-step mechanism has been described as "non-specific affinity labeling" in textbook literature. Small-molecule inhibitors similar to iodoacetate and N-ethyl maleimide are assumed to simultaneously modify many side-chains on the target protein molecule and have negligibly low initial binding affinity, which presumably explains their one-step kinetic behavior. In contrast, highly specific inhibitors that precisely target the enzyme's active site are assumed to always follow the two-step kinetic mechanism **A** or **B** [21, sec. 7.2.1] [1, sec 9.1]. In this sense, the fact that certainly highly specific and high-affinity inhibitors also follow the one-step mechanism **C** might appear as a paradox.

In order to better understand the unexpected "one-step" kinetics of certain high-affinity inhibitors, in section 3.3 of this report we have conjured up an inhibitor with molecular properties that were perfectly known (i.e., simulated) in advance. The objective was to simulate a compound that might approximate the kinetic properties of ibrutinib and similar "one-step" inhibitors of protein kinases. This hypothetical inhibitor was characterized by high initial binding affinity, with inhibition constant equal to $K_i = 1$ nM. We have simulated assays of this inhibitor at concentrations as high as $[I]_0^{(max)} = 10$ nM, which is ten times higher than the inhibition constant. Under the rapid-equilibrium approximation, it should be easily possible to determine the inhibition constant from the simulated data, because at a ten-fold excess of the inhibitor concentration over the inhibition constant the plot of k_{obs} vs. $[I]_0$ is expected to be highly hyperbolic. The maximum observed k_{obs} value should be closely approaching the asymptotically saturating value, which is by definition equal to k_{inact}. However, the actually observed k_{obs} vs. $[I]_0$ plot was essentially linear, showing no sign of hyperbolic saturation. This means that only the covalent efficiency constant k_{eff} = "k_{inact}/K_i" could be determined from the simulated data, but not the values of k_{inact} and K_i considered separately.

The results of this simulation study confirm what has been observed experimentally for example for ibrutinib inhibition of the TEC and BTK kinases [19, 22]. A highly specific, precisely targeted, and high-affinity irreversible inhibitor characterized by an extremely low *equilibrium* dissociation constant of the reversibly formed noncovalent complex, can indeed behave *kinetically* though the reaction were proceeding via the simple one-step mechanism **C**. There are at least two possible non-mathematical explanations of this potentially puzzling behavior, as described immediately below. In either case, it becomes practically impossible to determine either the inhibition constant K_i or the inactivation rate constant k_{inact}. Thus, the only way to quantitatively characterize the potency of such extremely efficient "one-step" covalent inhibitors is by way of the covalent efficiency constant $k_{eff} \equiv k_{inact}/K_i$.

The first intuitively accessible explanation of "one-step" kinetics has to do with the familiar idea that those reversible (non-covalent) inhibitors that are characterized by extremely low dissociation rate constant k_{-1} can behave as *effectively irreversible* on the time scale of the given kinetic experiment. In fact, the distinction between truly irreversible (covalent) inhibition and effectively irreversible (non-covalent) inhibition can be so blurred that in many cases it can only be established by specialized experiments [1, sec. 5.2]. See also our previous results on extremely potent, non-covalent but nearly irreversible inhibitors of 5-α-ketosteroid reductase [23]. Importantly, if a given covalent inhibitor happens to be "effectively" or "nearly" irreversible already in the first (strictly speaking reversible) binding step, the overall two-step covalent inhibition mechanism will kinetically manifest as a one-step process. This is because the decrease in catalytic activity over time (by definition proportional to the mole fraction of *unmodified* enzyme remaining in the reaction mixture) is exactly identical for the single-step irreversible mechanism $E+I \rightarrow EI$ and also for any multi-step irreversible process of type $E+I \rightarrow (EI)_1 \rightarrow (EI)_2 \rightarrow \cdots \rightarrow (EI)_n$. The rate of product formation is completely insensitive to how many isomers of the EI complex exist, as long as the first one is formed effectively irreversibly.

Another intuitively understandable reason for "one-step" kinetic behavior of certainly highly specific and precisely targeted irreversible inhibitors has to do with the reactivity of the covalent warhead. Indeed if the covalent inactivation rate constant $k_2 \equiv k_{inact}$ happens to be sufficiently high, the initial noncovalent complex may be converted to the final covalent conjugate so rapidly that the mole fraction of the noncovalent complex in the enzyme mass balance is always negligibly small. Under those circumstances, the overall two-step inactivation reaction might manifest kinetically as a one-step bimolecular association.

The steady-state algebraic model for two-step irreversible inhibition, newly presented in this report, should serve as a convenient tool to help increase our understanding of complex data-analytic issues arising in the evaluation of covalent enzyme inhibitors as potential therapeutic agents.

Supporting information

The following supporting files accompany this document:

1. File BioKin-TN-2020-02-SI.pdf: Mathematical details of the derivation of the double-exponential Eqn (5); listing of all DynaFit input scripts that were used to generate this report; instructions for using the Microsoft Excel simulation file BioKin-TN-2020-02-SI2.xls.
2. File BioKin-TN-2020-02-SI2.xls: A Microsoft Excel template file that can be used to simulate the reaction progress curves according to the steady-state kinetic mechanism **A** in *Figure 1*.

References

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

[2] I. H. Segel, Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-state Enzyme Systems, Wiley, New York, 1975.

[3] A. Abdeldayem, Y. S. Raouf, S. N. Constantinescu, R. Moriggl, P. T. Gunning, Advances in irreversible kinase inhibitors, Chem. Soc. Rev. 49 (2020) 2617–2687.

[4] D. Schuetz, W. de Witte, Y. Wong, B. Knasmueller, L. Richter, D. Kokh, S. Sadiq, R. Bosma, I. Nederpelt, L. Heitman, E. Segala, M. Amaral, D. Guo, D. Andres, V. Georgi, L. Stoddart, S. Hill, R. Cooke, C. De Graaf, R. Leurs, M. Frech, R. Wade, E. de Lange, A. IJzerman, A. Mueller-Fahrnow, G. Ecker, Kinetics for Drug Discovery: An industry-driven effort to target drug residence time, Drug Disc. Today 22 (2017) 896–911.

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

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

[7] A. C. Hindmarsh, ODEPACK: a systematized collection of ODE solvers, in: R. S. Stepleman, M. Carver, R. Peskin, W. F. Ames, R. Vichnevetsky (Eds.), Scientific Computing, North Holland, Amsterdam, 1983, pp. 55–64.

[8] 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.

[9] J. F. Morrison, Kinetics of the reversible inhibition of enzyme-catalysed reactions by tight-binding inhibitors, Biochim. Biophys. Acta 185 (1969) 269–286.

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

[11] C. Cobelli, J. J. DiStefano, Parameter and structural identifiability concepts and ambiguities: a critical review and analysis, Am. J. Physiol. Regul. Integr. Comp. Physiol. 239 (1980) 7–24.

[12] R. Bellman, K. Åström, On structural identifiability, Math. Biosci. 7 (1970) 329–339.

[13] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, J. Timmer, Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood, Bioinformatics 25 (2009) 1923–1929.

[14] J. M. Beechem, Global analysis of biochemical and biophysical data, Meth. Enzymol. 210 (1992) 37–54.

[15] D. M. Bates, D. G. Watts, Nonlinear Regression Analysis and its Applications, Wiley, New York, 1988.

[16] D. G. Watts, Parameter estimation from nonlinear models, Meth. Enzymol. 240 (1994) 24–36.

[17] K. A. Johnson, Z. B. Simpson, T. Blom, Global kinetic explorer: A new computer program for dynamic simulation and fitting of kinetic data, Anal. Biochem. 387 (2009) 20–29.

[18] K. A. Johnson, Z. B. Simpson, T. Blom, Fitspace explorer: An algorithm to evaluate multidimensional parameter space in fitting kinetic data, Anal. Biochem. 387 (2009) 30–41.

[19] 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.

[20] 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.

[21] A. Cornish-Bowden, Fundamentals of Enzyme Kinetics, 4th Edition, Wiley-VCH Verlag, Berlin, 2012.

[22] M. Hopper, T. Gururaja, T. Kinoshita, J. P. Dean, R. J. Hill, A. Mongan, Relative selectivity of irreversible inhibitors requires assessment of inactivation kinetics and cellular occupancy: A case study of ibrutinib and acalabrutinib, J. Pharm. Exp. Therap. 372 (2020) 331–338.

[23] M. L. Moss, P. Kuzmič, J. D. Stuart, G. Tian, A. G. Peranteau, S. V. Frye, S. H. Kadwell, T. A. Kost, L. K. Overton, I. R. Patel, Inhibition of human steroid 5-alpha reductases type I and II by 6-aza-steroids: Structural determinants of one-step vs. two-step mechanism, Biochemistry 35 (1996) 3457–3464.