

High-Throughput Screening of Enzyme Inhibitors: Simultaneous Determination of Tight-Binding Inhibition Constants and Enzyme Concentration

Petr Kuzmič,*¹ Kyle C. Elrod,† Lynne M. Cregar,† Steve Sideris,† Roopa Rai,‡ and James W. Janc†

*BioKin, Ltd., 1652 South Grand Avenue, Suite 337, Pullman, Washington 99163; and †Department of Enzymology and

‡Department of Medicinal Chemistry, Axys Pharmaceuticals, Inc., 180 Kimball Way, South San Francisco, California 94080

Received February 14, 2000

Active site titration by a reversible tight-binding inhibitor normally depends on prior knowledge of the inhibition constant. Conversely, the determination of tight-binding inhibition constants normally requires prior knowledge of the active enzyme concentration. Often, neither of these quantities is known with sufficient accuracy. This paper describes experimental conditions under which both the enzyme active site concentration and the tight-binding inhibition constant can be determined simultaneously from a single dose-response curve. Representative experimental data are shown for the inhibition of human kallikrein.

© 2000 Academic Press

Key Words: enzyme kinetics; mathematics; statistics; regression analysis; tight binding; inhibition constant; active site titration; enzyme concentration; high-throughput screening; kallikrein.

The concentration of enzyme active sites $[E]_0$ is usually determined in a titration experiment, using an irreversible or a reversible tight-binding inhibitor with a *known* apparent inhibition constant (1). Conversely, tight-binding inhibition constants K_i^{app} are usually determined from initial velocity measurements using a *known* concentration of the enzyme (2). Thus, the precise measurement of either quantity, $[E]_0$ or K_i^{app} , depends on the prior knowledge of the other.

In this paper we show under what conditions it is possible to measure $[E]_0$ and K_i^{app} simultaneously, from

a single dose-response curve. We also describe an optimal method for the determination of tight-binding inhibition constants in high-throughput, automated measurements. The method consists of fitting certain dose-response curves twice. In an initial least-squares fit the presumed enzyme concentration is *held constant*. If the apparent inhibition constant so determined falls into a certain range, the analysis is repeated while the enzyme concentration is treated as an *adjustable parameter*.

The merits of the method are exemplified on a series of Monte-Carlo simulations. We found that under some conditions even a relatively small bias in the presumed enzyme concentration (e.g., 50%) causes a large error (e.g., more than an order of magnitude) in the apparent inhibition constant. This result shows the need for our two-stage regression method during automated high-throughput analysis of pharmaceutically important tight-binding inhibitors.

As an example, we show a set of laboratory data for the inhibition of human kallikrein. The raw data, and a sample computer code to perform the calculations proposed here, can be obtained electronically via the Internet at the address given in footnote 1.

METHODS

Materials

Human plasma kallikrein (Athens Research and Technology) and kallikrein substrate, H-Pro-Phe-Arg-AMC (Bachem Bioscience) were purchased from the indicated commercial sources. The compound RR-101, a competitive reversible kallikrein inhibitor, was synthesized at Axys Pharmaceuticals.

¹ To whom correspondence should be addressed at BioKin, Ltd., 1652 South Grand Avenue, Suite 337, Pullman, WA 99163. E-mail: pkuzmic@biokin.com. www: <http://www.biokin.com>.

Experimental

Inhibitor potency and enzyme activity measurements were performed at room temperature using a fluorescence microplate reader (Molecular Devices). Plasma kallikrein (nominal concentration 4.0 nM) was incubated with inhibitor RR-101 (final concentration 1, 2, 3, 4, 6, 8, 16, and 32 nM) in 50 mM Tris (pH 7.4), 150 mM NaCl, 0.05% Tween 20, 500 μ M EDTA, and 10% DMSO for 30 min. Control reactions in the absence of inhibitor were performed in replicates of 8. The reactions were initiated by the addition of substrate (350 μ M H-Pro-Phe-Arg-AMC) and the rate of substrate hydrolysis was measured by monitoring the increase in fluorescence (excitation 355 nm, emission 460 nm) over a 5-min period. The commercial vendor provided an initial estimate of enzyme active site concentration; the final enzyme titer and inhibition constants were determined by the methods described in this paper.

Computational: Monte-Carlo Simulations

Dose-response curves for tight-binding inhibitors were simulated by using the Morrison equation [1], where $[E]_0$ is the active enzyme concentration, K_i^{app} is the apparent inhibition constant, v_0 is the control velocity observed in the absence of inhibitors ($[I]_0 = 0$), and v is the initial reaction velocity observed at inhibitor concentration $[I]_0$.

$$v = v_0 \frac{[E]_0 - [I]_0 - K_i^{\text{app}} + \sqrt{([E]_0 - [I]_0 - K_i^{\text{app}})^2 + 4[E]_0 K_i^{\text{app}}}}{2[E]_0}. \quad [1]$$

Each simulated dose-response curve contained nine points corresponding to the following inhibitor concentrations: 10.0, 2.50, 0.6250, 0.1563, 0.0391, 0.0099, 0.0010, and 0.0005 μ M, and a data point simulated in the absence of inhibitors ($[I]_0 = 0$). The reaction velocity in the absence of inhibitor was arbitrarily set to unity, $v_0 = 1.0$.

At each chosen value of K_i^{app} or $[E]_0$, 10,000 data sets were generated by superimposing normally distributed pseudo-random noise, with constant variance and standard deviation set to 5% of the maximum simulated reaction velocity. Each set of simulated data was subjected to nonlinear least-squares regression by using the Levenberg-Marquardt algorithm (3).

RESULTS

Propagation of Systematic Errors

In the first series of Monte-Carlo simulations, we investigated how systematic errors in enzyme concen-

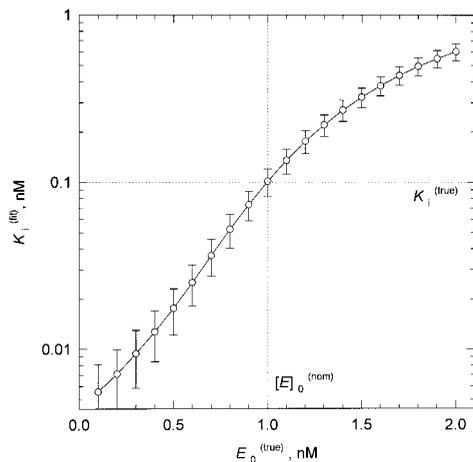


FIG. 1. Propagation of systematic errors from the active enzyme concentration into the apparent inhibition constant. Circles and error bars represent averages and standard deviations, respectively, for each set of 10,000 fitted values of $K_0^{(\text{fit})}$. The true value of the inhibition constant was $K_i^{(\text{true})} = 0.01$ nM. The nominal enzyme concentration ($[E]_0^{(\text{nom})} = 1.0$ nM in all cases) was held as a constant parameter in the least-squares fit of dose-response curves to Eq. [1]. The true enzyme concentration $[E]_0^{(\text{true})}$ was varied from 0.1 to 2.0 nM. Each dose-response curve contained nine data points, at inhibitor concentrations described in the experimental section.

tration propagate into systematic errors in the apparent inhibition constant. For each simulated value of the inhibition constant, $K_i^{(\text{true})}$, 200,000 artificial dose-response curves were generated and subsequently subjected to nonlinear least-squares fit.

All simulated dose-response curves contained nine data points, representing initial velocities measured at different inhibitor concentrations (see experimental section for the corresponding values of $[I]_0$).

Each dose-response curve was generated by using a fixed $K_i^{(\text{true})}$ for the inhibition constant (for example, 0.1 nM) and the correct enzyme concentration $[E]_0^{(\text{true})}$. In a subsequent least-squares fit of the same curve to Eq. [1], the enzyme concentration was held constant at a value that deviated from the correct concentration, $[E]_0^{(\text{nom})}$ (for example, 1.0 nM instead of 0.9 nM). Because of this systematic error in enzyme concentration, the fitted inhibition constant $K_i^{(\text{fit})}$ was different from the simulated inhibition constant $K_i^{(\text{true})}$.

At each combination of $[E]_0$ and K_i^{app} , 10,000 slightly different dose-response curves were simulated and subsequently fit to Eq. [1], yielding a spread of values for the fitted inhibition constant. The distribution of the $K_i^{(\text{fit})}$ values was approximately Gaussian. Therefore, it was possible to summarize the results of 10,000 simulations in the form of an average and a standard deviation. One series of numerical experiments, obtained for $K_i^{(\text{true})} = 0.1$ nM, is summarized in Fig. 1.

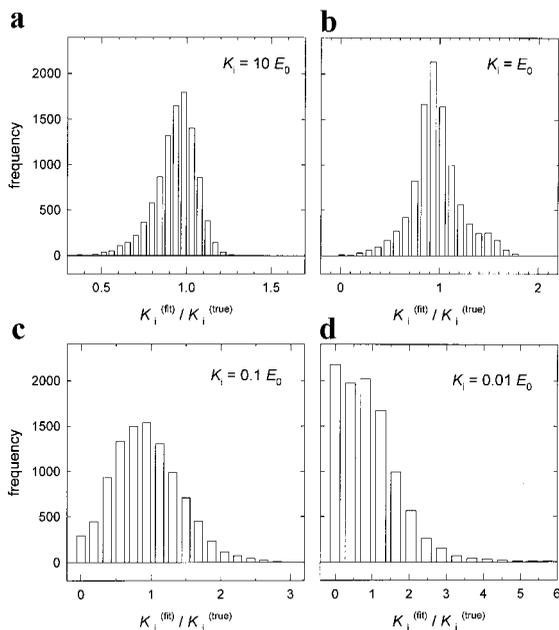


FIG. 2. Uncertainty in the determination of K_i^{app} from dose-response data. Dose-response curves were simulated at $[E]_0^{(\text{true})} = 1.0$ nM in all cases. For each panel a through d, 10,000 slightly different dose-response curves were simulated and subsequently subjected to least-squares fit to Eq. [1]. The nominal enzyme concentration was treated as an adjustable parameter. Note the non-Gaussian character of the histogram in panel d; approximately 20% of all data sets in that series lead to the value of $K_i^{(\text{fit})}$ which was practically zero.

Figure 1 represents 20 sets of simulations assuming the true value $K_i^{(\text{true})} = 0.1$ nM and the nominal enzyme concentration $[E]_0^{(\text{nom})} = 1.0$ nM, which was held constant in the least-squares fit. The true enzyme concentration $[E]_0^{(\text{true})}$ was varied between 0.1 and 2.0 nM. The results show that comparatively small systematic errors in enzyme concentrations propagate into large systematic errors in inhibition constants. For example, if the true enzyme concentration were 0.5 nM instead of the assumed 1.0 nM, the inhibition constant would be underestimated on the average by a factor of 9 ($K_i^{(\text{fit})} = 0.012$). Even a systematic error as small as 10% in enzyme concentration ($[E]_0^{(\text{true})} = 0.9$ nM) leads to fitted values of inhibition constants that are underestimated on the average by a factor of 2 to 3.

Even more prominent distortion in the fitted value of K_i^{app} was seen in a similar series of experiments at $K_i^{(\text{true})} = 0.01$ nM, under otherwise identical conditions ($[E]_0^{(\text{nom})} = 1.0$ nM). In that series of Monte-Carlo simulations (data not shown), at some values of the true enzyme concentration (e.g., at $[E]_0^{(\text{true})} = 0.1$ nM) a significant number of the dose-response curves (approximately 2000 of the 10,000 total data sets) could

not be fit to Eq. [1] at all. In those cases the best-fit value $K_i^{(\text{true})}$ was practically zero. This is illustrated also in Fig. 2d below.

Simultaneous Determination of K_i^{app} and $[E]_0$

In the second series of Monte-Carlo simulations, we examined the conditions under which it is possible to determine both the inhibition constant and the enzyme active site concentration from a single dose-response curve. This simultaneous determination is based on fitting dose-response data to Eq. [1], while treating not only K_i^{app} and v_0 but also $[E]_0$ as adjustable parameters.

The simulated dose-response curves again consisted of nine data points (zero inhibitor concentration plus eight data points at inhibitor concentrations described in the experimental section). The simulated enzyme concentration was always 1.0 nM, while the inhibition constant was varied from 10 nM to 0.01 nM, stepping by a factor of 10. At each value of the apparent inhibition constant, 10,000 slightly different dose-response

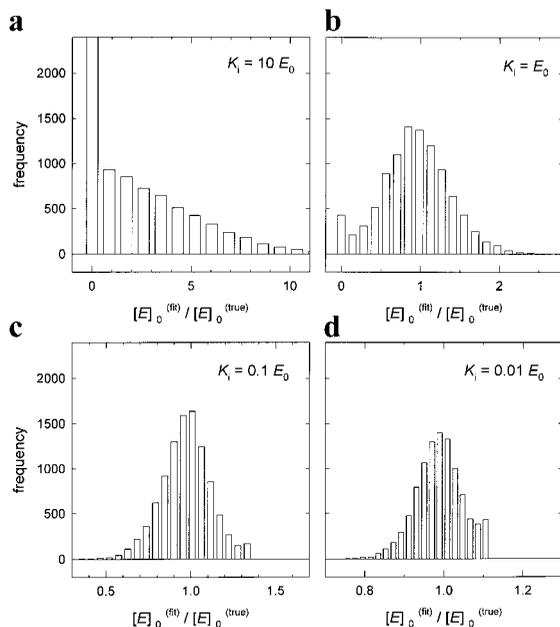


FIG. 3. Uncertainty in the determination of $[E]_0$ from dose-response data. For each panel a through d, 10,000 slightly different dose-response curves were simulated and subsequently subjected to least-squares fit to Eq. [1]. The nominal enzyme concentration was treated as an adjustable parameter. Note the non-Gaussian character of the histogram in panel a; most data sets in that series lead to the value of $[E]_0^{(\text{fit})}$ which was practically zero, which means that the enzyme concentration cannot be estimated under these conditions. However, the inhibition constant is estimated very reliably (see Fig. 2a).

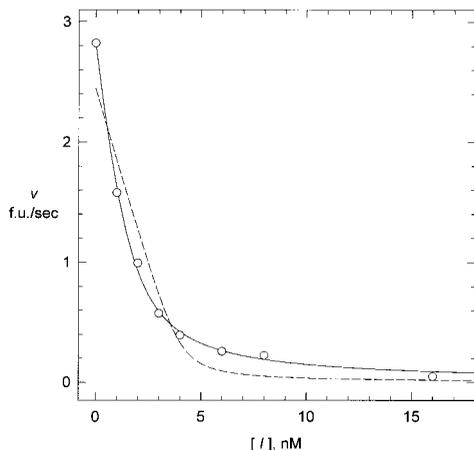


FIG. 4. Inhibition of human kallikrein. The initial velocities (relative fluorescence units per second) were fit to Eq. [1]. Dashed curve: enzyme concentration was kept constant at the nominal value $[E]_0 = 4.00$ nM; the fitted value of the inhibition constant was $K_i^{\text{app}} = 85$ pM. Solid curve: enzyme concentration was treated as an adjustable parameter; the fitted values were $[E]_0 = 1.53$ nM and $K_i^{\text{app}} = 478$ pM.

curves were simulated. Each simulated curve was subsequently fit to Eq. [1] and the best-fit values of K_i^{app} and $[E]_0$ were recorded. The spread of best-fit values of K_i^{app} is shown in Fig. 2, while the spread of best-fit values of $[E]_0$ is shown in Fig. 3.

Figure 2 shows that the histogram of distribution for the fitted values of K_i^{app} is reasonably narrow only for those values that are larger than 10% of the true enzyme concentration. On the other hand, Fig. 3 shows that the enzyme concentration can be determined with certainty only if the inhibition constant is, at most, equal to the true enzyme concentration itself. Thus, between $K_i^{\text{app}} > 0.1 \times [E]_0$ and $K_i^{\text{app}} < 1.0 \times [E]_0$, both quantities can be determined simultaneously from the same dose-response curve.

Sample Experimental Data

Human kallikrein was assayed at the nominal concentration $[E]_0 = 4.00$ nM, based on the nominal concentration of the enzyme stock. Protease inhibitor RR-101 was used to inhibit the enzyme in 96-well plate reader assays. The initial reaction velocities at different inhibitor concentrations are shown as open circles in Fig. 4.

The Morrison equation [1] was used to fit the experimental data in two different ways. In the first round of analysis, represented by the dashed curve in Fig. 4, the concentration of the enzyme was held constant at its nominal value $[E]_0 = 4.00$ nM. The best-fit value of the apparent inhibition constant so obtained was 85 ± 97 pM. The relatively large standard error is caused by

the systematic discrepancy between the experimental data and the theoretical model (the dashed curve in Fig. 4 deviates significantly from the data points). The sum of squared deviations was 0.3798.

In the second round of analysis, represented by the solid curve in Fig. 4, the enzyme concentration was treated as an adjustable parameter yielding the best-fit value of $[E]_0 = 1.53 \pm 0.19$ nM. This value represents approximately 40% of the nominal enzyme concentration. The best-fit value of the apparent inhibition constant was 478 ± 57 pM, which is approximately six times the value observed in the first round of nonlinear regression. The sum of squared deviations between the experimental data and the theoretical model was 0.0141.

DISCUSSION

Laboratories engaged in high-throughput kinetic characterization of enzyme inhibitors have an increasing need to develop primary and secondary (follow-up) screening assays in which the concentrations of the enzyme and the inhibitor are comparable in magnitude ("tight binding"). The automated two-stage regression analysis described in this paper represents a new data-analysis technique that can be used to improve the accuracy of such high-throughput enzyme kinetic analysis without any additional expenditures in time or materials.

We have recently described a new method for automated determination of tight-binding enzyme inhibition constants (4). The method has been applied successfully to more than 100,000 inhibitors of mast-cell tryptase and other therapeutic targets. In the course of this work, we noticed that the inhibition constants seemed remarkably sensitive to the presumed enzyme concentration. Particularly for extremely tight binding inhibitors, it appeared that small variations in the enzyme concentration produced large variations in the observed inhibition constant. To our knowledge, no systematic statistical treatment of such error propagation exists in the analytical literature.

In this study we present the results of Monte-Carlo simulations predicting that for very tight-binding inhibitors ($K_i^{\text{app}} \approx 0.1 \times [E]_0$ or $0.01 \times [E]_0$, see Fig. 1), a systematic error in enzyme concentration as small as 10–20% might produce as much as an order of magnitude systematic distortion of the apparent inhibition constant. This prediction was confirmed on experimental data for the kallikrein inhibitor RR-101. At the presumed enzyme concentration $[E]_0 = 4.00$ nM the inhibition constant was 85 pM, while at the best-fit enzyme concentration $[E]_0 = 1.53$ nM, the apparent inhibition constant was 478 pM.

Systematic errors in the concentration of enzyme active sites are probably quite common in biochemical laboratories, including high-throughput screening establishments. Enzyme preparations easily change specific activity due to a host of denaturation processes or small changes in the assay conditions (pH, temperature, ionic strength). The question is how to avoid the consequences, in the form of large systematic errors in the apparent inhibition constants, of such inadvertent changes in total enzyme concentration.

The strategy proposed in this report emerged from the results of Monte-Carlo simulations shown in Figs. 2 and 3. We found that when the inhibition constant is lower than the total enzyme concentration, it is possible to extract both quantities from a single dose-response curve. On the other hand, when the inhibition constant is relatively high ($K_i^{\text{app}} > [E]_0$), the enzyme concentration *cannot* be extracted from the experiment and therefore must be treated as a constant. Fortunately, under such “loose binding” conditions even a very large systematic error in enzyme concentration has no effect on the observed inhibition constant, because the Morrison equation [1] reduces to Eq. [2].

$$v = v_0 \frac{K_i^{\text{app}}}{[I]_0 + K_i^{\text{app}}}. \quad [2]$$

Based on these results, the proper computational algorithm for high-throughput analysis of enzyme inhibitors is the *two-stage* regression analysis proposed in this paper. In the first round of analysis, the presumed enzyme concentration is treated as a constant parameter. The apparent inhibition constant so determined is accepted as the final result if it is higher than the nominal enzyme concentration. Otherwise the least-squares fit of the same data to Eq. [1] is repeated, while the enzyme concentration is treated as an adjustable parameter.

Based on Monte-Carlo simulation studies similar to those reported here, Greco and Hakala (2) proposed that the accuracy of measuring tight-binding inhibition constants can be improved by lowering the enzyme concentration, if possible even to the point where tight-binding is avoided (that is, where $[E]_0 \ll K_i^{\text{app}}$). However, for many enzyme systems this strategy is not possible or practical, because commonly used synthetic substrates are too inefficient. A relatively large amount of enzyme would then be required to complete the assay in a reasonably short time. For example, assays of viral proteases such as cytomegalovirus assembling (5, 6) may require very high concentrations, in the micromolar range, in order to detect substantial proteolytic activity. Some mammalian proteases that are

potential therapeutic targets such as prostate specific antigen (7) or urokinase (8) are routinely assayed at concentrations approaching 100 nM. Under such conditions even inhibitors with $K_i^{\text{app}} \approx 10$ nM behave as “tight binding.”

The two-stage regression method described in this paper has two important limitations. First, the nominal enzyme concentration must not be too inaccurate. In exploratory Monte-Carlo simulations (data not shown), we found that the nominal enzyme concentration must be accurate approximately within an order of magnitude. Outside this range, the nonlinear least-squares regression might produce an erroneous estimate of the apparent inhibition constant, ending in a local minimum on the least-squares surface.

The second limitation of the two-stage regression method is illustrated by Fig. 2d. A significant number of the simulated dose-response curves (approximately 2000 or 20% from the total) produced artificially low estimates for the inhibition constants, practically approaching zero. This problem can be minimized by a proper design of the kinetic experiment, that is, by optimally choosing the inhibitor concentrations. Our preliminary results suggest that an optimally designed series of inhibitor concentrations includes several points near the nominal enzyme concentration. The inhibitor dilution series chosen in this study (see experimental section) is not optimal in that regard. The results of a detailed study on the optimal design of high-throughput kinetic experiments will be reported elsewhere.

It has been said (9) that “in the 100 years to 1995, the pharmaceutical industry worked on about 500 targets with a limited number of compounds, whereas now, using new technologies like genomics, high throughput screening, and combinatorial chemistry, drug companies will see an explosion in the number of targets and leads [they] can explore. To improve the transition from research to development it is necessary to automate the research and development process [. . .] using information technologies to make better use of existing data.” Our two-stage regression method represents a step toward this goal. A suitable computer program for performing the kinetic analysis is available free of charge via the internet (<http://www.biokin.com>).

ACKNOWLEDGMENT

We thank Sarah McCord for helpful discussions and a careful reading of the manuscript.

REFERENCES

1. Williams, J. W., and Morrison, J. F. (1979) *Meth. Enzymol.* **63**, 437–467.

2. Greco, W. R., and Hakala, M. T. (1979) *J. Biol. Chem.* **254**, 12104–12109.
3. Marquardt, D. W. (1963) *J. Soc. Ind. Appl. Math.* **11**, 431–441.
4. Kuzmič, P., Sideris, S., Cregar, L. M., Elrod, K. C., Rice, K. D., and Janc, J. W. (2000) *Anal. Biochem.*, in press.
5. Baum, E. Z., Ding, W. D., Siegel, M. M., Hulmes, J., Bebernitz, G. A., Sridharan, L., Tabei, K., Krishnamurthy, G., Carofiglio, T., Groves, J. T., Bloom, J. D., DiGrandi, M., Bradley, M., Ellestad, G., Seddon, A. P., and Gluzman, Y. (1996) *Biochemistry* **35**, 5847–5855.
6. Sardana, V. V., Wolfgang, J. A., Veloski, C. A., Long, W. J., LeGrow, K., Wolanski, B., Emini, E. A., and LaFemina, R. L. (1994) *J. Biol. Chem.* **269**, 14337–14340.
7. Yang, C. F., Porter, E. S., Boths, J., Kanyi, D., Hsieh, M., and Cooperman, B. S. (1999) *J. Pept. Res.* **54**, 444–448.
8. Jankum, J., Keck, R. W., Skrzypczak-Jankun, E., and Swiercz, R. (1997) *Cancer Res.* **57**, 559–563.
9. Kuhlmann, J. (1999) *Int. J. Clin. Pharmacol. Ther.* **37**, 575–583.