

# Kinetic determination of tight binding impurities in enzyme inhibitors

Petr Kuzmic <sup>\*,1</sup>, Craig Hill <sup>2</sup>, Matthew P. Kirtley <sup>2</sup>, and James W. Janc <sup>2</sup>

<sup>1</sup> *BioKin, Ltd.*

1652 South Grand Avenue, Suite 337, Pullman, WA 99163

<sup>2</sup> *Celera Genomics*

Department of Enzymology and High-Throughput Screening

180 Kimball Way, South San Francisco, CA 94080

**Journal**            *Analytical Biochemistry*  
**Ms. No.**            *ABIO537*  
**Revision Date**    *March 21, 2003*  
**Section**            *Enzyme assays and analyses*  
**Running title**    *Mixtures of enzyme inhibitors*

# Abstract

A novel rate equation is derived to characterize the dose-response behavior of a moderately potent ("classical") enzyme inhibitor contaminated with a very potent ("tight binding") impurity. Mathematical properties of the new rate equation show that, for such contaminated materials, experimentally observed  $I_{50}$  values are ambiguous. The four-parameter logistic equation, conventionally used to determine  $I_{50}$  values, cannot be used to detect the presence of tight binding impurities in inhibitor samples. In contrast, fitting the newly derived rate equation to inhibitor dose-response curves can, in favorable cases, reveal whether the unknown material is chemically homogeneous or whether it is contaminated with a tight binding impurity. The limitations of our method with respect to the detectable range of inhibition constants (both classical and tight binding) were examined by using Monte-Carlo simulations. To test the new analytical procedure experimentally, we added a small amount (0.02 mole %) of a tight binding impurity ( $K_i = 0.065$  nM) to an otherwise weak inhibitor of human mast-cell tryptase ( $K_i = 50.4$   $\mu$ M). The resulting material was treated as "unknown". Our kinetic equation predicts that such adulterated material should show  $I_{50} = 0.40$   $\mu$ M, which was identical to the experimentally observed value. The best-fit value of the apparent inhibition constants for the tight binding inhibitor was  $K_i = (0.107 \pm 0.035)$  nM, close to the true value of 0.065 nM.

## Keywords

*enzyme kinetics; mathematics; statistics; regression analysis; tight binding; enzymes; inhibition constant; impurities; inhibitor mixtures; IC-50; four-parameter logistic equation; mast cell tryptase*

# Introduction

Classical enzyme inhibitors are characterized by apparent inhibition constants ( $K_i$ ) that are very much larger than the enzyme concentration ( $[E]_0$ ) in the given assay. In contrast, tight binding inhibitors are characterized by  $K_i$  values comparable in magnitude with  $[E]_0$ , or even much smaller ( $K_i \ll [E]_0$ ). In this paper, we propose a new method of analyzing initial velocity data obtained with *mixtures* of classical and tight binding inhibitors.

Our research has been motivated by anecdotal evidence that occasionally enzyme inhibitor samples encountered in screening for drug discovery are not chemically pure. For example, using analytical methods recently described in this Journal [1, 2], we have determined  $K_i$  values for hundreds of thousands of inhibitors. However, for certain inhibitors we found that the dose-response behavior deviated from the expected theoretical model. The question arose whether this anomalous kinetic behavior could be explained by the presence of tight binding impurities in inhibitors that are weakly potent, or even completely inactive.

To address this problem, we derived a rate equation that describes the dose-response kinetics of inhibitor mixtures consisting of one tight binding inhibitor and one classical inhibitor. The mathematical properties of our equation show that the experimentally observed  $I_{50}$  values are ambiguous. Each observed value of  $I_{50}$  can arise in principle by infinitely many combinations of the classical inhibition constant, the tight binding inhibition constant, and the molar fraction of tight binding impurity. This is in contrast with chemically pure samples, for which Cha's equation [3] or the Cheng-Prusoff equation [4] give an unambiguous relationship between  $K_i$  and  $I_{50}$ .

The predictive power of the new rate equation was established in a series of Monte-Carlo simulations, using more than a million synthetic data sets that were similar in experimental design to those typically generated in the laboratory. We found that, under certain conditions, it is indeed possible to diagnose the presence of a tight binding impurity in an otherwise moderately active or inactive sample. In most cases, the inhibition constant of the tight binding contaminant could be determined reliably, but the "classical" inhibition constant of the weakly potent major component could not.

To verify our new kinetic method experimentally, we created an artificially impure inhibitor sample. A weakly potent inhibitor of human mast-cell tryptase (benzamidine,  $K_i = 50 \mu\text{M}$ ) was doped with a very small amount (0.02%) of the tight binding inhibitor CRA-001390 ( $K_i = 0.065 \text{ nM}$ ) [1]. The resulting "unknown" inhibitor had  $I_{50} = 0.4 \mu\text{M}$ , exactly as predicted by our theoretical model.

The nonlinear least-squares fit to the commonly used four-parameter logistic equation gave no indication that the sample was not chemically homogeneous. In contrast, fitting the same experimental data to the mechanism-based Morrison equation [5, 6] clearly revealed systematic discrepancies, suggesting the presence of a tight binding impurity. Using our newly derived kinetic equation as the fitting model, we obtained a very satisfactory fit. The best-fit values were  $K_i = (0.107 \pm 0.035) \text{ nM}$  for the tight binding inhibition constant and  $(0.09 \pm 0.02)\%$  for the molar fraction of the tight binding impurity. The apparent inhibition constant for benzamidine could not be determined.

Our new method of kinetic analysis should be of interest in all biochemical laboratories where large numbers of enzyme inhibitors are evaluated with the goal of extracting apparent inhibition constants or  $I_{50}$  values from dose-response curves. Many such laboratories are involved in secondary screening of enzyme inhibitors as potential therapeutics.

## Methods

### Materials

Tos-Gly-Pro-Lys-pNA (Sigma, St. Louis, MO) and benzamidine (Sigma-Aldrich, Milwaukee, WI) were purchased from the indicated commercial sources. Mast cell tryptase was purified from the immortalized human mast cell line, HMC-1, as previously reported [7]. CRA-001390 was synthesized at Celera Genomics (formerly Axys Pharmaceuticals compound APC-1390) as described [8].

### Experimental

Kinetic measurements were performed in a total reaction volume of  $100 \mu\text{L}$  in 96-well U-bottom microtiter plates (Falcon). Measurements of substrate hydrolysis were made using a SpectraMax 250 kinetic plate reader (Molecular Devices, Sunnyvale, CA). Tryptase ( $50 \text{ pM}$  active site) was combined with inhibitor at varying concentrations in a buffer containing  $50 \text{ mM}$  Tris (pH 8.2),  $150 \text{ mM}$  NaCl,  $0.05\%$  Tween 20,  $10\%$  DMSO and

50 µg/ml heparin for one hour at room temperature. Control reactions in the absence of inhibitor were performed in replicates of eight. Reactions were initiated by the addition of substrate (0.5 mM Tos-Gly-Pro-Lys-pNA) and the rate of substrate hydrolysis was measured by monitoring the change in absorbance at 405 nm over 10 minutes. Initial velocities were determined by linear regression of progress curves.

## Computational

$I_{50}$  values were determined by nonlinear regression of percent inhibition data by using the four-parameter logistic equation (1), where  $p$  ("percent inhibition") is the relative decrease in enzyme activity due to the inhibitor concentration  $[I]_0$ . The regression analysis was performed by the commercial software packages Prism (GraphPad) and SigmaPlot (SPSS).

$$p = p_{\min} + \frac{P_{\max} - P_{\min}}{1 + ([I]_0 / I_{50})^n} \quad (1)$$

Initial velocity data were analyzed by nonlinear regression using a modification of the Marquardt algorithm [9]. The following equations have been employed as theoretical models. Equation (2) applies to a pure tight binding or classical inhibitor. Equation (3) applies to a single tight binding or classical inhibitor present as an impurity (molar fraction  $\alpha$ ) in a completely inactive bulk material.

$$v = v_0 \frac{[E]_0 - [I]_0 - K_T + \sqrt{([E]_0 - [I]_0 - K_T)^2 + 4[E]_0 K_T}}{2[E]_0} \quad (2)$$

$$v = v_0 \frac{[E]_0 - \alpha[I]_0 - K_T + \sqrt{([E]_0 - \alpha[I]_0 - K_T)^2 + 4[E]_0 K_T}}{2[E]_0} \quad (3)$$

The kinetics of a tight binding inhibitor (apparent inhibition constant  $K_T$ , molar fraction  $\alpha$ ) mixed with a classical inhibitor (apparent inhibition constant  $K_C$ ) is described by

equation (4), where the auxiliary variable  $\beta$  is defined by equation (5). We have derived equation (4) by using the algebraic method of Segel (ref. [10], p. 22) based on the rapid-equilibrium approximation.

$$v = v_0 \frac{[E]_0 - \alpha[I]_0 - \beta K_T + \sqrt{([E]_0 - \alpha[I]_0 - \beta K_T)^2 + 4[E]_0 \beta K_T}}{2\beta[E]_0} \quad (4)$$

$$\beta = 1 + (1 - \alpha)[I]_0 / K_C \quad (5)$$

The dependence of  $I_{50}$  on inhibition constants for a mixture of a classical and a tight binding inhibitor is described by equation (6). When the molar fraction of the tight binding inhibitor is set to unity ( $\alpha = 1$ , representing a pure tight binding inhibitor), equation (6) turns into equation (7) according to Cha [3]. On the other hand, when  $\alpha$  is set to zero (representing a pure classical inhibitor), we obtain the Cheng-Prusoff equation,  $I_{50} = K_C$ , as a special case for classical inhibitors [4].

$$I_{50} = \frac{[E]_0 / 2 + K_T}{([E]_0 / 2 + K_T)(1 - \alpha) / K_C + \alpha} \quad (6)$$

$$I_{50} = [E]_0 / 2 + K_T \quad (7)$$

Under special circumstances, both apparent inhibition constants  $K_C$  and  $K_T$  and the enzyme concentration  $[E]_0$  are known, and the task is to compute the molar fraction of tight binding impurity from the observed value of  $I_{50}$ , using equation (8).

$$\alpha = \frac{K_C / I_{50} + 1}{K_C / ([E]_0 / 2 + K_T) - 1} \quad (8)$$

# Results

## **$I_{50}$ for Inhibitor Mixtures**

For chemically homogeneous classical inhibitors,  $I_{50}$  is identical to  $K_C$  by definition (Cheng-Prusoff equation [4]). For homogeneous tight binding inhibitors, the plot of  $I_{50}$  vs.  $K_T$  also is a straight line with unit slope, but this time with nonzero intercept equal to  $[E]_0/2$  (Cha's equation (7), ref. [3]). In contrast, the  $I_{50}$  for inhibitor mixtures depends on the apparent inhibition constants as a rectangular hyperbola, defined by equation (6).

It is obvious from equation (6) that any given value of  $I_{50}$  can be produced by infinitely many combinations of  $K_T$ ,  $K_C$ , and  $\alpha$ . This result suggests that any statistical method of analyzing dose-response curves that is based merely on determining  $I_{50}$  values (for example, fitting dose-response curves to the four-parameter logistic equation (1)) will produce ambiguous results. In contrast, nonlinear regression of dose-response data using the mechanism-based equation (4) offers at least the possibility of distinguishing between chemically homogeneous and heterogeneous samples.

A major concern is that equation (4) contains four adjustable parameters ( $K_C$ ,  $K_T$ ,  $\alpha$ , and  $v_0$ ). This large number of adjustable parameters can cause a large degree of mutual dependency (coupling, or correlation) among adjustable parameters. An extremely high degree of correlation among parameters could prohibit the accurate determination of all parameter values from real-world experimental data. A standard approach to investigating the possibility of such parameter correlation is to apply Monte-Carlo simulation techniques.

## **Monte-Carlo simulations**

A tight binding impurity in an inhibitor preparation can be detected kinetically if and when the parameters  $K_C$ ,  $K_T$ , and  $\alpha$  can be determined simultaneously from dose-response data. To ascertain which (if any) combinations of  $K_C$ ,  $K_T$ , and  $\alpha$  can be so determined, we chose a typical experimental design and a typical measurement error (e.g., two or three percent) for a series of Monte-Carlo simulations. The experimental design included nine inhibitor concentrations, starting from the maximum  $[I]_0 = 1.0 \mu\text{M}$ , and including seven two-fold serial dilutions (0.5, 0.25, 0.125 ...  $\mu\text{M}$ ) as well as the negative control ( $[I]_0 = 0$ ).

Using this layout, we employed equation (4) to simulate ten thousand dose-response curves for each particular combination of  $K_C$ ,  $K_T$ , and  $\alpha$ . "Experimental" data sets were generated by taking the theoretical data set and superimposing on it normally distributed pseudo-random noise with standard deviation equal to 1, 2, or 3% of the simulated velocity. The equation (4) was then fit to the synthetic data sets to obtain least-squares values of adjustable parameters  $K_C$ ,  $K_T$ , and  $\alpha$ . The purpose of this analysis is to compare these best-fit parameters with their simulated "true" values.

A representative example is shown in *Figure 1*. The dashed curve represents a theoretical dose-response curve simulated with  $[E]_0 = 1.0$  nM,  $K_C = 0.1$   $\mu$ M,  $K_T = 0.1$  nM, and  $\alpha = 1.0$  %. The data points (open circles) were generated by superimposing pseudo-random error with standard deviation equal to 3% of the simulated velocity. The least-squares fit of the simulated data (solid curve) yielded  $K_C = 0.065$   $\mu$ M,  $K_T = 0.041$  nM, and  $\alpha = 0.6$  %. Clearly, the 3% pseudo-random noise significantly distorted the values of the fitting parameters. In some simulations the best-fit values of model parameters  $K_C$ ,  $K_T$ , and  $\alpha$  were much closer to their true values, but in other simulations the best-fit values were even more distorted.

Figure 1 near here

A summary of thirty thousand representative regression analyses for three different combinations of  $K_C$ ,  $K_T$ , and  $\alpha$  (ten thousand data sets for each combination) is shown in *Figure 2*. Panels A and B use the same values of model parameters ( $K_C = 0.1$   $\mu$ M,  $K_T = 0.1$  nM, and  $\alpha = 1.0$  %), only the simulated experimental error was different. A slight increase in experimental error, from one percent in panel A to two percent in panel B, causes a significant increase in the uncertainty of the inhibition constants  $K_C$  and  $K_T$ . This is illustrated by the widening of the histograms. Additionally, in many of the Monte-Carlo simulations the classical inhibition constant  $K_C$  could not be determined at all. The solid black rectangles represent all regression analyses where the best-fit value of  $K_C$  was several orders of magnitude larger than the simulated value, corresponding to effectively no inhibition. In these cases the 99% bulk material behaved as completely inactive, and only the tight binding constant  $K_T$  could be determined at all. These results suggest that the measurements of initial velocities must be very accurate (experimental error approximately one percent) in order to determine the tight binding inhibition constants for impurities.

Figure 2 near here

The relative magnitude of classical versus tight binding inhibition constants had even greater effect on the possibility of determining simultaneously the values of  $K_C$ ,  $K_T$ , and  $\alpha$  in Monte-Carlo simulations. This is illustrated in panel C (bottom) in *Figure 2*, where both  $K_C$  and  $K_T$  have ten times greater values compared to panel A (top). The uncertainties of both  $K_C$  and  $K_T$  in the bottom panel ( $K_C = 1000 \times [E]_0$ ,  $K_T = [E]_0$ ) are much larger compared to the top panel ( $K_C = 100 \times [E]_0$ ,  $K_T = [E]_0/10$ ). These results suggest that simultaneous determination of both apparent inhibition constants ( $K_C$  and  $K_T$ ) will be possible only in a restricted range of their values.

To investigate this range, we have performed 1302 Monte-Carlo simulations, each with ten thousand dose-response data sets. The apparent inhibition constant  $K_C$  had 21 different values and the apparent inhibition constant  $K_T$  had 62 different values ( $21 \times 62 = 1302$ ).  $K_C$  was varied from 0.1  $\mu\text{M}$  to 8.6736  $\mu\text{M}$  stepping logarithmically by a factor of 1.25, whereas  $K_T$  was varied from 0.001 nM to 815.66 nM, again stepping logarithmically by the same factor. For each of the 1302 combinations of inhibition constant values (assuming  $[E]_0 = 1.0$  nM and standard error of measurement 1%) we assessed the spread of best-fit values, as illustrated for three such Monte-Carlo runs in *Figure 1*. The conclusions can be summarized as follows.

The classical inhibition constant  $K_C$  could be determined from dose-response curves only if two conditions were met simultaneously: (i) the true value of  $K_C$  was lower than approximately  $500 \times [E]_0$ ; (ii) the true value of  $K_T$  was lower than approximately  $0.5 \times [E]_0$ . Outside of this range, the best-fit values of  $K_C$  were many orders of magnitude larger than the true values, suggesting that the classical inhibitor had virtually zero activity. Thus, for most combinations of  $K_C$  and  $K_T$ , the inhibitor sample behaved as if it were composed of a completely inactive (ballast) material contaminated with a tight binding impurity. In other words, for most combinations of  $K_C$  and  $K_T$ , equation (3) was the most suitable fitting model.

On the other hand, the tight binding inhibition constant  $K_T$  (and also the molar fraction of impurity  $\alpha$ ) could be reliably determined regardless of the true value of  $K_C$ , but only if  $K_T$  was in the range between  $0.005 \times [E]_0$  and  $0.5 \times [E]_0$ . These results suggest that the detection of a tight binding impurity in an inhibitor sample will be most reliable when the enzyme concentration is at least comparable in magnitude with  $K_T$ , or higher ( $[E]_0 \geq K_T$ ). However, extremely high enzyme concentrations relative to  $K_T$  ( $[E]_0 > 200 \times K_T$ ) again are not useful.

## Experimental example

To test our method based on equation (4), we have contaminated a weak mast-cell tryptase inhibitor with a tight binding impurity. Benzamidine ( $K_C = 50.4 \mu\text{M}$ ) was doped with a tryptase inhibitor CRA-001390 ( $K_T = 0.065 \text{ nM}$ , [1]) at the molar ratio 5000:1 (0.02% impurity). For both pure and contaminated benzamidine, initial reaction velocities were measured at eight different inhibitor concentrations, plus negative control ( $[I]_0 = 0$ ). The enzyme concentration was 0.05 nM. Using equation (6), we can predict the  $I_{50}$  for this mixture of inhibitors to be approximately 0.4  $\mu\text{M}$ .

Figure 3 near here

The initial velocities were analyzed by two different methods. The first method consisted of nonlinear regression using the conventional four-parameter logistic equation (1). The results are summarized graphically in *Figure 3*. The best least-squares fit values of  $I_{50}$  were  $(50.4 \pm 3.3) \mu\text{M}$  for pure benzamidine and  $(0.40 \pm 0.05) \mu\text{M}$  for benzamidine contaminated with CRA-001390. Dose response curves for both pure benzamidine and contaminated benzamidine fit to equation (1) very well, as is seen from the fact that the experimental data points (open circles in *Figure 3*) deviate very little from the best-fit curves. Thus, the results of fit to equation (1) do not indicate in any way that the dose-response curve B in *Figure 3* is generated by a chemically heterogeneous material.

The second method of statistical analysis consisted of nonlinear regression of the same initial velocity data using (in turn) equations (2), (3), and (4), instead of the four-parameter logistic equation (1). The results are graphically summarized in *Figure 4*.

Figure 4 near here

Least-squares fit of dose-response data for contaminated benzamidine to equation (2), describing a chemically homogeneous tight binding or classical inhibitor, is represented by the dashed curve in *Figure 4*. It is important to note a pronounced lack of fit. Indeed, within the experimental errors represented by the error bars shown in *Figure 4*, only two out of nine experimental data points overlap with the best-fit curve.

Least-squares fit of the same data to equation (4), describing a chemically heterogeneous mixture of a classical inhibitor contaminated by a tight binding impurity, is represented by solid curve in *Figure 4*. The best-fit values and formal standard errors of adjustable parameters were  $v_0 = (7.90 \pm 0.31)$  a.u./min,  $\alpha = (0.087 \pm 0.015)\%$ ,  $K_C = (1.9 \pm 49.8)$  M, and  $K_T = (0.107 \pm 0.035)$  nM. The best-fit value of the apparent inhibition constant for benzamidine is not only extremely large (38000 times larger than the true value  $K_C = 50$   $\mu$ M), but also it is accompanied by a large formal standard error. These results mean that the inhibition constant for benzamidine cannot be determined at all from the available data, and that the sample behaves effectively as a completely inactive material containing a small amount of a tight binding impurity. Accordingly, we repeated the regression analysis using equation (3) instead of equation (4). The results were identical for the relevant model parameters:  $v_0 = (7.90 \pm 0.31)$  a.u./min,  $\alpha = (0.087 \pm 0.015)\%$ , and  $K_T = (0.107 \pm 0.035)$  nM. These results suggest that the inhibitor is not chemically pure. The regression analysis accurately diagnosed that the adulterated sample contains less than 100 ppm of a tight binding impurity with  $K_i$  approximately 100 pM (true values  $\alpha = 20$  ppm,  $K_i = 65$  pM).

## Discussion

We have recently described methods for kinetic analysis that are suitable for fully automatic determination of inhibition constants and  $I_{50s}$  [1, 2]. The underlying theoretical model for dose-response curves is the equation (2) for "tight binding" inhibition [5, 6], which is also applicable to weakly active or "loose-binding" inhibitors. These methods have been used successfully in our laboratories to analyze hundreds of thousands of inhibitor samples. However, occasionally we observe that a certain dose-response curve does not fit equation (2) very well.

There are many possible explanations for such deviations between experimental data and the given theoretical model. In this paper, we focused on one explanation that seems reasonable based on the complex realities of drug research. Indeed, *de novo* enzyme inhibitors are often products of multi-step chemical syntheses and purification methods whose efficiency is necessarily finite. Such syntheses almost always yield a preparation containing a single inhibitory species. However, is it possible that an occasional anomaly in dose-response behavior could be caused by the presence of tight binding impurities in otherwise weakly potent or even entirely inactive material?

To answer this question, we applied two different methods of analyzing inhibitor dose-response data. The first method was traditional, based on the four-parameter logistic equation. A systematic survey of biochemical literature on the evaluation of enzyme inhibitors reveals that, year after year (e.g., [11-13]),  $I_{50}$  remains a favorite measure of inhibitory potency. Similarly, the four-parameter logistic equation remains a favorite mathematical model for inhibitor dose-response curves.

In this paper, we have demonstrated one potential pitfall of this traditional approach. In particular, the four-parameter logistic equation cannot discriminate between a chemically homogeneous inhibitor (classical or tight binding) and a heterogeneous sample consisting of (a) bulk material that has moderate or no inhibitory potency and (b) a highly potent impurity. Indeed both experimental dose-response curves in *Figure 3* (curve A for a chemically pure inhibitor, and curve B for an inhibitor mixture) fit the available data extremely well.

Because the fit to the four-parameter logistic equation (1) is very good in both cases, there is no "red flag" that would alert the investigator clearly to the presence of a tight binding impurity. Indeed, curve B in *Figure 3* appears as just another half-micromolar inhibitor, instead of an essentially inactive material containing a small amount of compound with  $K_i = 65$  pM.

The only possible giveaway is the Hill slope parameter  $n$  in equation (1). For the contaminated benzamidine (curve B in *Figure 3*) we obtained  $n = (1.53 \pm 0.14)$ , whereas for pure benzamidine (curve A) the slope parameter was  $n = (0.98 \pm 0.07)$ . However, in practice the Hill slope parameter was found not sufficiently reliable to reveal conclusively the presence of tight binding impurities. Even for experimental data that are affected by small measurement error (less than one percent), for nearly optimally designed experiments, and for kinetically pure inhibitors, the Hill slope parameter varied widely between 0.7 and 1.5 simply because of random errors. This was especially true when both  $p_{\max}$  and  $p_{\min}$  in equation (1) were treated as adjustable parameters. Although these observations are merely anecdotal in the scope of this report, a more detailed examination of the Hill slope parameter  $n$  as an indicator of anomalous inhibitor behavior will be presented elsewhere.

A more successful approach to correctly diagnosing the presence of tight binding impurities in enzyme inhibitors is based on our newly derived rate equation (4), used in conjunction with equation (2). Our approach is easily amenable to automation in a high-throughput setting. Such automated analysis of inhibitor dose-response curves consists

of three separate steps. In the first step, the data are analyzed by using equation (2), which assumes that the inhibitor preparation is chemically pure. In the second step, the same data are fit to equation (4), which assumes the presence of a tight binding impurity. In the third step, we apply a standard model-discrimination analysis based on the  $F$ -statistic [14].

According to our method of analysis, an inhibitor sample is flagged as *possibly* containing a tight binding impurity if two conditions are satisfied simultaneously. The first condition is that the traditional equation (2) gives a poor fit based on the residual sum of squared deviations. The second condition is that the calculated  $F$ -statistic suggests a statistically significant improvement in the goodness of fit, going from the simple equation (2) to the more complex equation (4). When both conditions are satisfied, the experimentally determined values of  $I_{50}$  and  $K_i$  should not be trusted, and the sample should be carefully scrutinized for trace amounts of tight binding impurities.

The approach described in this paper, as in any other analytical methodology, has distinct limitations and requirements. Our mathematical model does not take into account the detailed mechanism of inhibition (competitive vs. non-competitive) and only operates on *apparent* inhibition constants. Another restriction arises from the assumed composition of the inhibitor mixture. In the hypothetical case of two tight binding inhibitors present as a mixture, it would be necessary to solve a cubic equation [15]. However, we have purposely restricted ourselves to the less general case, where only one inhibitor is tight binding. Therefore, our theoretical model arises more simply, as solution to a quadratic equation.

A more important limitation of our analytical method lies in that the inhibition constants of the impurity and the bulk material must be significantly different from each other. Only then, both inhibition constants can be determined reliably at the same time, as is illustrated by Monte-Carlo simulations. Our experimental example illustrates the opposite scenario, where simultaneous determination of  $K_C$  and  $K_T$  was not possible. Furthermore, the impurity's inhibition constant must fall into a certain specific range (several orders of magnitude wide) relative to the enzyme concentration used in the assay. The model-discrimination method described here requires that the initial velocities are measured very accurately and reproducibly (less than 2% experimental error), and that the experiment is nearly optimally designed in terms of the inhibitor concentrations chosen. Finally, in fitting dose-response data to equation (4), we assumed that the enzyme concentration  $[E]_0$  can be reasonably treated as a fixed parameter. Our previous paper

(ref. [2]) describes in detail the appropriate statistical treatment of those experimental situations where the assumption of constant  $[E]_0$  does not hold.

We hope that other researchers active in the kinetic analysis of enzyme inhibitors, especially those evaluating the inhibitory potency in the drug discovery process, will find good use for our new analytical method and the underlying mathematical model. In our experience, abandoning the traditional exclusive reliance on  $I_{50}$  - a step taken in these laboratories several years ago - certainly did enhance the information content of a massive data stream flowing from various inhibitor screening projects.

## Acknowledgments

We thank Sarah K. McCord (Washington State University, Health Sciences Library) for careful reading of the manuscript and helpful suggestions. We also acknowledge with gratitude all comments offered by anonymous Celera Genomics internal reviewers.

## References

1. Kuzmic, P., Sideris, S., Cregar, L. M., Elrod, K. C., Rice, K. D., and Janc, J. W. (2000) "High-throughput screening of enzyme inhibitors: Automatic determination of tight-binding inhibition constants" *Anal. Biochem.* **281**, 62-67.
2. Kuzmic, P., Elrod, K. C., Cregar, L. M., Sideris, S., Rai, R., and Janc, J. W. (2000) "High-throughput enzyme kinetics: Simultaneous determination of tight-binding inhibition constants and enzyme concentration" *Anal. Biochem.* **286**, 45-50.
3. Cha, S. (1975) "Tight binding inhibitors - I. Kinetic behavior" *Biochem. Pharmacol.* **24**, 2177-2185.
4. Cheng, Y.-C., and Prusoff, W. H. (1972) "Relationship between the inhibition constant ( $K_i$ ) and the concentration of inhibitor which causes 50 percent inhibition ( $IC_{50}$ ) of an enzymatic reaction" *Biochem. Pharmacol.* **22**, 3099-3108.
5. Morrison, J. F. (1969) "Kinetics of the reversible inhibition of enzyme-catalysed reactions by tight-binding inhibitors" *Biochim. Biophys. Acta* **185**, 269-286.
6. Williams, J. W., and Morrison, J. F. (1979) "The kinetics of reversible tight-binding inhibition" *Meth. Enzymol.* **63**, 437-467.

7. Butterfield, J. H., Weiler, D. A., Hunt, L. W., Wynn, S. R., and Roche, P. C. (1990) "Purification of tryptase from a human mast cell line" *J. Leukoc. Biol.* **47**, 409-419.
8. Rice, K., Tanaka, R., Katz, B., Numerof, R., and Moore, W. (1998) "Inhibitors of tryptase for the treatment of mast cell-mediated diseases" *Curr. Pharm. Design* **4**, 393-408.
9. Marquardt, D. W. (1963) "An algorithm for least-squares estimation of nonlinear parameters" *J. Soc. Ind. Appl. Math.* **11**, 431-441.
10. Segel, I. H. (1975) *Enzyme Kinetics*, Wiley, New York.
11. Lisnock, J., Griffin, P., Calaycay, J., Frantz, B., Parsons, J., O'Keefe, S. J., and LoGrasso, P. (2000) "Activation of JNK31 requires both MKK4 and MKK7: kinetic characterization of in vitro phosphorylated JNK31" *Biochemistry* **39**, 3141-3148.
12. Adlington, R. M., Baldwin, J. E., Becker, G. W., Chen, B., Cheng, L., Cooper, S. L., Hermann, R. B., Howe, T. J., McCoull, W., McNulty, A. M., Neubauer, B. L., and Pritchard, G. J. (2001) "Design, synthesis, and proposed active site binding analysis of monocyclic 2-azetidinone inhibitors of prostate specific antigen" *J. Med. Chem.* **44**, 1491-1508.
13. Farr, C. D., Burd, C., Tabet, M. R., Wang, X., Welsh, W. J., and Ball, W. J. (2002) "Three-dimensional quantitative structure-activity relationship study of the inhibition of Na<sup>+</sup>,K<sup>+</sup>-ATPase by cardiotonic steroids using comparative molecular field analysis" *Biochemistry* **41**, 1137-1148.
14. Schmider, J., Greenblatt, D. J., Harmatz, J. S., and Shader, R. I. (1996) "Enzyme kinetic modelling as a tool to analyse the behaviour of cytochrome P450 catalysed reactions: application to amitriptyline N-demethylation" *Brit. J. Clin. Pharmacol.* **41**, 593-604.
15. Kuzmic, P., Ng, K. Y., and Heath, T. D. (1992) "Mixtures of tight-binding enzyme inhibitors. Kinetic analysis by a recursive rate equation" *Anal. Biochem.* **200**, 68-73.

## Figure Legends

### *Figure 1*

A typical simulation / regression sequence from a Monte-Carlo study of Equation (4), using  $[E]_0 = 1.0$  nM. Synthetic data points were generated at  $[I]_0 = 1, 0.5, 0.25, \dots, 0.0078125$ , and 0 nM. Pseudo-random error with standard deviation of 3% relative to

the generated value was superimposed on the ideal data, and the resulting data set was fit to Equation (4). Ten thousand such simulation / regression runs were performed for each combination of inhibition constants  $K_T$ ,  $K_C$ , and the molar fraction of tight binding impurity,  $\alpha$ .

### Figure 2

Summary of three typical Monte-Carlo simulations, all conducted at  $[E]_0 = 1.0$  nM and the molar fraction of impurity  $\alpha = 1\%$ . *Upper panel (A)*: Simulated values were  $K_C = 0.1$   $\mu$ M,  $K_T = 0.1$  nM, with the pseudo-random error 1%. *Middle panel (B)*: Same as **A**, except for pseudo-random error 2%. *Lower panel (C)*: Simulated values  $K_C = 1.0$   $\mu$ M,  $K_T = 1.0$  nM, with the pseudo-random error 1%. For explanation see text. In total, 1302 such Monte-Carlo simulations were performed at 1302 different combinations of  $K_T$ ,  $K_C$  and  $\alpha$  (approximately 1.3 million synthetic data sets).

### Figure 3

Dose-response curves for the inhibition of human mast-cell tryptase by **A** pure benzamidine ( $K_C = I_{50} = 50$   $\mu$ M) or **B** benzamidine doped with 0.02% of the tight binding inhibitor CRA-001390 ( $K_T = 0.065$  nM). Data points are averages from triplicate measurements of initial velocities. Percent inhibition values ( $p$ ) were fitted to the four-parameter logistic equation (1) to obtain  $I_{50}$  values. Equation (6) predicts for "adulterated" benzamidine  $I_{50} = 0.40$   $\mu$ M.

### Figure 4

Nonlinear least-squares fit of inhibition data for human mast-cell tryptase ( $[E]_0 = 0.05$  nM). The inhibitor was benzamidine doped with the tight binding inhibitor CRA-001390. Curve **A** is the best-least squares fit to Equation (2) for a *pure* tight binding or classical inhibitor. Note the pronounced lack of fit (dashed curve misses most of the data points). Curve **B** is the best least squares fit to Equation (4) for a *mixture* of a classical inhibitor contaminated with a tight binding inhibitor. The best-fit values were molar fraction  $\alpha = (0.09 \pm 0.02)\%$ , apparent inhibition constant  $K_T = (107 \pm 35)$  pM. The actual values for inhibitor CRA-001390 were  $\alpha = 0.02\%$ ,  $K_T = 65$  pM (ref. [1]).







