Analytical Biochemistry **281,** 62–67 (2000)
doi:10.1006/abio.2000.4501, available online at http://www.idealibrary.com on IDEAL®

High-Throughput Screening of Enzyme Inhibitors: Automatic Determination of Tight-Binding Inhibition Constants

Petr Kuzmič,*[,1] Steve Sideris,† Lynne M. Cregar,† Kyle C. Elrod,† Kenneth D. Rice,‡ and James W. Janc†

**BioKin, Ltd., P.O. Box 8336, Madison, Wisconsin 53708; and †Department of Enzymology and ‡Department of Medicinal Chemistry, Axys Pharmaceuticals, Inc., 180 Kimball Way, South San Francisco, California 94080*

Received September 6, 1999

Determination of tight-binding inhibition constants by nonlinear least-squares regression requires sufficiently good initial estimates of the best-fit values. Normally an initial estimate of the inhibition constant must be provided by the investigator. This paper describes an automatic procedure for the estimation of tight-binding inhibition constants directly from dose–response data. Because the procedure does not require human intervention, it was incorporated into an algorithm for high-throughput screening of enzyme inhibitors. A suitable computer program is available electronically (http://www.biokin.com). Representative experimental data are shown for the inhibition of human mast-cell tryptase. **© 2000 Academic Press**

Key Words: **enzyme kinetics; mathematics; data analysis; statistics; tight binding; inhibition constant; high-throughput screening; mast-cell tryptase.**

"Tight binding" is defined as a condition where the active enzyme concentration is approximately equal to, or even higher than, the apparent inhibition constant. For tight-binding inhibitors, inhibition constants K_i^{app} (1) are more useful measures of inhibitory potency than IC_{50} values or nominal values of relative enzymatic activity determined at a single concentration of the inhibitor ("percentage inhibition").

Unfortunately, unlike these simpler measures of inhibitory potency, tight-binding inhibition constants can be obtained only by nonlinear least-squares regression (2–4) of dose–response data. In all nonlinear regression analyses, the investigator must provide a sufficiently good initial estimate of the parameter to be determined (here, the tight-binding inhibition constant). For the relatively simple data analysis discussed in this paper, and for most sets of data that can be encountered, the initial estimate of the inhibition constant need only be within about an order of magnitude of the actual values. Nevertheless, the determination of inhibition constants from dose–response data does require human intervention, which presents an obstacle in high-throughput or automated experiments.

In this paper we describe a computational procedure that can be used to make a rough estimate of the apparent tight-binding inhibition constant automatically, without human intervention. "Classical" inhibition constants pertaining to less tightly bound inhibitors (K_i^{app} much greater than enzyme concentration) are also readily determined by the same procedure.

The estimation procedure was incorporated into an algorithm for automated screening of inhibition activity on commercial 96-well plate readers. The initial estimate is subsequently refined by nonlinear least-squares fit. As an illustrative example, we show a set of typical data for the inhibition of human mast-cell tryptase by two different inhibitors, namely, phenylguanidine (a classical inhibitor, $K_i^{app} \approx 50\ \mu M$) and a tight-binding inhibitor APC-1390 (Scheme 1; $K_i^{app} \approx 50$ pM).

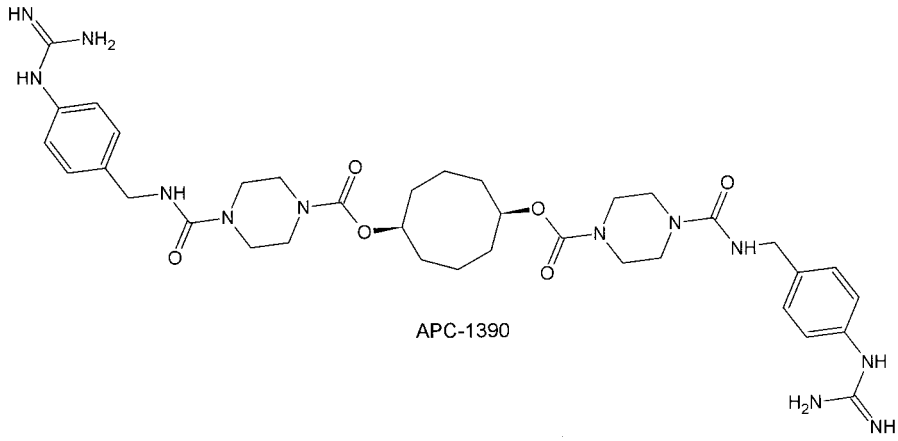
METHODS

Materials

Tos-Gly-Pro-Lys-pNA (Sigma, St. Louis, MO) and phenylguanidine (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 (5). APC-1390 was synthesized at Axys Pharmaceuticals as described (6).

[1] To whom correspondence should be addressed. Fax: (608) 256-1269. E-mail: pkuzmic@biokin.com.

0003-2697/00 $35.00
Copyright © 2000 by Academic Press
All rights of reproduction in any form reserved.



SCHEME 1

Experimental

Kinetic measurements were performed in a total reaction volume of 100 μl in 96-well U-bottom microtiter plates (Falcon). Measurements of substrate hydrolysis were made using a Thermomax kinetic plate reader (Molecular Devices, Sunnyvale, CA). Tryptase (290 pM or 2.5 nM active site) was combined with inhibitor at varying concentrations in a buffer containing 50 mM Tris (pH 8.2), 150 mM NaCl, 0.05% Tween 20, 10% DMSO and 50 μg/ml heparin for 1 h 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 5 min. The apparent inhibition constants were calculated from the enzyme progress curves using the method described in this paper.

Computational

According to Williams and Morrison (1), the apparent tight-binding inhibition constant K_i^{app} (subsequently referred to as K, for brevity) can be determined by nonlinear least-squares fit of reaction velocities to Eq. [1], where v is the initial reaction velocity observed at inhibitor concentration [I], v_0 is the initial reaction velocity observed in the absence of inhibitor, and [E] is the active enzyme concentration. Usually [E] is treated as a constant parameter, while v_0 and K are considered as adjustable parameters for which the investigator must provide initial estimates to be refined in the regression.

$$v = v_0 \frac{[E] - [I] - K + \sqrt{([E] - [I] - K)^2 + 4[E]K}}{2[E]} \qquad [1]$$

Let us assume that the regression analysis is to proceed automatically, for example in an automated high-throughput screening of enzyme inhibitors, some of which might be tight binding. Under those circumstances it is desirable to estimate v_0 and K directly from the experimental data. As an initial guess of the best-fit value of v_0, we use the reaction velocity experimentally observed in the absence of inhibitor. The initial estimate of the apparent inhibition constant K is determined as described below.

Explicit formula for the apparent inhibition constant. Equation [1] can be rearranged in order to express K as shown in Eq. [2]:

$$K = \frac{[I] - [E](1 - v/v_0)}{v_0/v - 1}. \qquad [2]$$

According to Eq. [2], the apparent tight-binding inhibition constant can be computed from any pair of initial velocity measurements v, in the presence of the inhibitor (at concentration [I]), and v_0, in the absence of the inhibitor.

In the hypothetical case of experimental measurements devoid of random errors, it would be sufficient to perform only two measurements of velocity, namely, v and v_0, from which the inhibition constant K would be calculated immediately. In practice, each measurement of v and v_0 is affected by random experimental error. It is therefore advantageous to perform the measurements of reaction velocity at different values of the inhibitor concentration $[I]_j$ ($j = 1, 2, \ldots, n$), calculate n slightly different values of the inhibition constant at each concentration, and finally estimate the true value of K as an average according to Eq. ([3]):

$$K = \frac{1}{n} \sum_{j=1}^{n} K_j, \qquad [3]$$

$$K_j = \frac{[I]_j - [E](1 - v_j/v_0)}{v_0/v_j - 1}. \qquad [4]$$

Iterative reweighting scheme. Inhibition constant K, estimated as a simple average according to Eq. [3], suffers from the drawback that not all values of K_j ($j = 1, 2, \ldots, n$) can be considered equally significant. For example, reaction velocities measured at extremely high inhibitor concentrations (e.g., $[I] > 1000K$) will be practically zero, whereas reaction velocities measured at extremely low inhibitor concentrations (e.g., $[I] < 0.001K$) will be practically equal to v_0. In both extreme cases, the measured reaction velocity contains little or no information about the true value of the inhibition constant K.

To circumvent this difficulty, we estimate the inhibition constant K as a weighted average (Eq. [5]) instead of a simple average (Eq. [3]). The weighting factors w_j (Eq. [6]) are sensitivity coefficients, i.e., partial derivatives of the observable quantity (the reaction velocity v) with respect to the adjustable parameter of interest (the apparent inhibition constant). The modified estimation procedure proceeds iteratively, according to the combined equations [5] and [6]:

$$K^{(m+1)} = \frac{\sum_{j=1}^{n} w_j^{(m)} K_j^{(m)}}{\sum_{j=1}^{n} w_j^{(m)}}, \qquad [5]$$

$$w_j^{(m)} = \frac{v_0}{2[E]}\left(\frac{[E] + [I]_j + K_j^{(m)}}{\sqrt{([E] - [I]_j - K_j^{(m)})^2 + 4[E]K_j^{(m)}}} - 1\right). \qquad [6]$$

In the first iteration ($m = 1$) all weights are set to unity, $w_j^{(1)} = 1$ ($j = 1, 2, \ldots, n$) and the apparent inhibition constant is calculated as a simple average, by using Eq. [3]. In each subsequent iteration ($m = 2, 3, \ldots, m_{max}$), we calculate first the weights according to Eq. [6] and then we update the weighted average according to Eq. [5]. When two subsequent estimates ($K^{(m)}$ and $K^{(m+1)}$) agree to a sufficient number of significant digits, the iteration is terminated.

RESULTS

Monte Carlo Simulations

The accuracy of the newly proposed estimation method was tested in several series of Monte Carlo simulations. Each individual simulation/estimation experiment proceeded in three stages. In the first stage, error-free initial velocities were generated by using Eq. [1]. The values of parameters and variables were selected to resemble a typical dose–response obtained in an automated screening experiment.

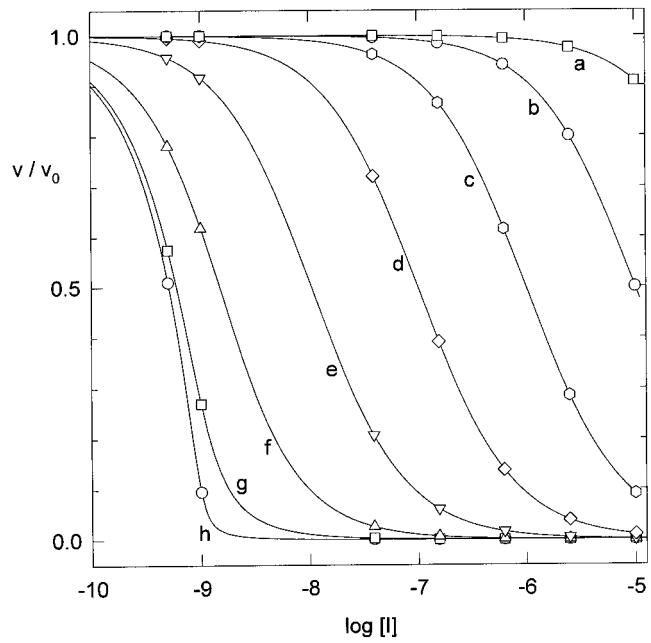


FIG. 1. Layout of inhibitor concentrations in the Monte Carlo simulation study. Dose–response curves labeled a–h correspond to the simulated values of the apparent inhibition constants $K = 100$ μM (curve a) and $K = 10$ pM (curve h), stepping by a factor of 10. Symbols represent the reaction velocities at corresponding inhibitor concentrations (eight concentrations per dose–response curve), simulated by using the Morrison equation [1] (1). To each ideal data point shown here, normally distributed pseudorandom error was added with standard deviation 0.05 (5%). Ten thousand random data sets were generated for each ideal dose–response curve shown here. The resulting "noisy" data were subjected to the weighted-average procedure described in the Methods section to determine the apparent inhibition constant K.

The enzyme concentration was fixed at [E] = 1.0 nM. Each simulated dose–response curve contained eight points, corresponding to the following inhibitor concentrations: 10.0, 2.50, 0.6250, 0.1563, 0.0391, 0.0099, 0.0010, and 0.0005 μM. The reaction velocity in the absence of inhibitor was arbitrarily set to unity, $v_0 = 1.0$. Dose–response curves were generated for the values of apparent inhibition constants K spanning 10 orders of magnitude between 100 μM and 10 pM ($K =$ 100 μM, 10 μM, . . . , 0.1 nM, 0.01 nM). The simulated dose–response curves are shown in Fig. 1.

In the second stage of each simulation/estimation experiment, normally distributed pseudorandom error with standard deviation 0.05 (5% scatter) was superimposed on the simulated data. Finally, the artificial "noisy" data were subjected to the iterative estimation procedure described in the Methods section, which generated an estimate of the inhibition constant K. Each series of simulation/estimation experiments was repeated 10,000 times with different sets of pseudorandom error imposed on the artificial data. The basic statistics (average, standard deviation, maximum

TABLE 1

Simulated and Estimated Values of the Apparent Inhibition Constant *K*

Simulated *K* (nM)	Fitted *K* (nM)			
	Average	Standard deviation	Maximum	Minimum
0.01	0.01002	0.00061	0.012	0.0081
0.1	0.1000	0.0025	0.11	0.092
1	1.001	0.023	1.08	0.92
10	10.2	0.30	11.4	9.1
100	100.4	2.4	115	95
1,000	1,001	20	1,105	933
10,000	10,070	450	14,090	9,070
100,000	102,000	13,000	417,000	73,000

Note. For details of the Monte Carlo simulation experiments, see text in the Methods section.

value, and minimum value) from each series are summarized in Table 1.

The results show that at each tested value of K except $K = 100{,}000$ nM, the estimated values deviated from the simulated values by less than 10%. At the highest value of the inhibition constant, $K = 100{,}000$ nM, the spread of estimated values was much higher but still within one order of magnitude in comparison with the "true" or simulated value.

Comparison with Existing Methods

Henderson (7) proposed a linearized form of Eq. [1] designed for the estimation of apparent inhibition constant K as the slope of a linear plot. When the ratio $[I]/(1 - v/v_0)$ is plotted against v_0/v (a transformed independent variable), the apparent tight-binding inhibition constant K is obtained as the slope:

$$\frac{[I]}{1 - v/v_0} = K\frac{v_0}{v} + [E] \qquad [7]$$

One possible application of the Henderson plot is to subject appropriately transformed dose–response data to a linear least-squares regression, which has the advantage that no initial estimate is needed for the value of K. Accordingly, each of the 100,000 simulated data sets described above (10 series of 10,000 data sets, at different values of the apparent inhibition constant) was transformed according to the Henderson method and then subjected to linear least-squares regression. The value of K was taken as the slope of each best-fit line. The results are summarized in Table 2.

The results in Table 2 show that the Henderson linearization method performs very well for values of inhibition constants that are at or below the nominal enzyme concentration (tight-binding conditions). However, the Henderson method fails on weakly bound inhibitor (classical conditions). The spread of estimated inhibition constants listed in the last four rows of Table 2 is very large and even includes large negative values, which is physically impossible.

Sample Experimental Data

As a representative example, we have analyzed the inhibition of human mast-cell tryptase by APC-1390 (a tight-binding inhibitor) and phenylguanidine (a classical inhibitor). The experimental data and the results of analysis are summarized in Table 3.

It is important to compare the predicted values of apparent, tight-binding inhibition constants, listed on the last line of Table 3, with the corresponding values optimized by the least-squares fit. To that end, we have subjected the same data to a nonlinear least-squares regression analysis by using a procedure developed by Duggleby (8). The results of the least-squares fit (a best-fit curve superimposed on the experimental data) are illustrated in Figs. 2 and 3.

For APC-1390 we obtained the least-squares optimized value of $K = 59.5 \pm 6.7$ pM, which is very close to the initial estimate, $K = 66.3$ pM. The least-squares value for phenylguanidine was $K = 59.0 \pm 2.6$ μM as compared with $K = 62.0$ μM for the initial estimate. In summary, the initially estimated value and the best-fit value of apparent constants were practically identical for both the tight-binding and the classical inhibitor. This is illustrated by the close agreement between the solid (least-squares fit) and dashed (estimated) lines in both Figs. 2 and 3.

DISCUSSION

Greco and Hakala (9) compared the reliability of 11 different methods for the determination of tight-binding enzyme inhibition constants. Most of these early analytical methods relied on graphical (or *linear* re-

TABLE 2

Simulated and Estimated Values of Apparent Inhibition Constants *K* Using the Henderson Equation [7] (7)

Simulated *K* (nM)	Fitted *K* (nM)			
	Average	Standard deviation	Maximum	Minimum
0.01	0.010000	0.000053	0.010194	0.009791
0.1	0.1000	0.00052	0.10190	0.09799
1	0.9999	0.0052	1.0174	0.9808
10	9.999	0.049	10.189	9.831
100	91.8	575.0	6,804.2	−41,829.3
1,000	1,285	13,033	377,224	−1,220,830
10,000	11,643	3,162,220	22,841,200	−311,495,000
100,000	251,964	86,296,500	4,665,960,000	−5,939,910,000

Note. For details of the Monte Carlo simulation experiments, see text in the Methods section.

TABLE 3

Inhibition of Human Mast-Cell Tryptase by APC-1390 and Phenylguanidine

	APC-1390 ([E] = 0.29 nM)			Phenylguanidine ([E] = 2.5 nM)		
j	$[I]_j$ (μM)	$v_j(10^6 \times$ au/s)	K (pM)	$[I]_j$ (μM)	$v_j(10^6 \times$ au/s)	K (μM)
0	0.000	78.55	—	0.00	702.80	—[a]
1	0.234	31.89	42.20	0.46	715.60	—[a]
2	0.351	21.58	53.27	1.37	722.10	—[a]
3	0.526	13.89	61.70	4.12	662.10	66.94
4	0.790	9.55	74.12	12.35	587.20	62.73
5	1.185	8.42	111.26	37.04	454.50	67.80
6	1.770	4.05	81.23	111.10	241.80	58.27
7	4.000	3.17	156.69	333.30	103.30	57.43
8	—	—	—	1000.00	38.19	57.46
		K = 66.3 pM			K = 62.0 μM	

Note. The values of apparent inhibition constants K_j were calculated for each inhibitor concentration by using Eq. [4]. The botton row in the table lists the iteratively reweighted averages computed by using Eq. [5]. The iterations were terminated when two consecutive values of K agreed within six significant digits.

[a] Velocities greater than control are automatically excluded from the analysis.

gression) procedures, such as the Henderson plot (7). Linear regression has the advantage of not requiring an initial estimate of K_i^{app} from the investigator, but unfortunately it is statistically unreliable. A more recent review on tight-binding inhibition (10) confirms this assessment of the linear Henderson plot.

In contrast, the *nonlinear* least-squares fit (2–4) of dose–response data to the Morrison equation [1] (11) was identified statistically as the most reliable method (9). However, the disadvantage is that the investigator must provide for each inhibitor an educated guess of its inhibition constant. This is a serious impediment in developing analytical techniques suitable for automated data processing required in high-throughput biochemical screening laboratories. To the best of our knowledge, no automatic procedure or computer program for the determination of apparent inhibition constants has been described in the available literature. Perhaps the issue is that most companies have their own in-house programs and have not shared these with the research community.

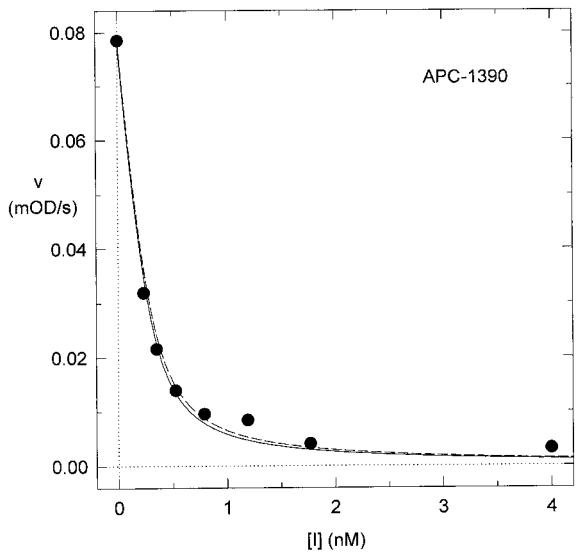


FIG. 2. Inhibition of human mast-cell tryptase by APC-1390. The dashed curve was generated according to the Morrison equation [1] (11), by using the automatically generated initial estimate for the apparent inhibition constant, K = 66.3 pM. The solid curve was generated by using the best least-squares fit value of the inhibition constant, K = 59.5 ± 6.7 pM.

In this paper we have described a novel method for automating the process of making a statistically reliable guess of the apparent enzyme inhibition constant (1) (either tight-binding, where $K \approx$ [E] or even $K <$ [E], or classical, where $K \gg$ [E]) directly from dose–response data. The reliability of the estimation procedure was tested in a series of Monte Carlo simulations. While assuming a realistically large random experimental error (5%), we found that our new estimation procedure yielded estimates of K_i^{app} that differed from the simulated values by, at most, several percent. The only exception was a Monte Carlo experiment in which the maximum concentration of inhibitor (10 μM) was very much lower than the inhibition constant (100 μM). However, even in this case the spread of estimated values falls within one order of magnitude from the "true" or simulated value.

These favorable results are in contrast with those that were obtained by using the linear Henderson plot (7) (see Table 2) and similar methods (12) (data not shown). We found that the linear least-squares fit according to the Henderson method (7) performed quite well under the tight-binding conditions ($K \approx$ [E]) but failed systematically under the classical conditions ($K \gg$ [E]). We considered the possibility of employing for the classical conditions a linear regression procedure

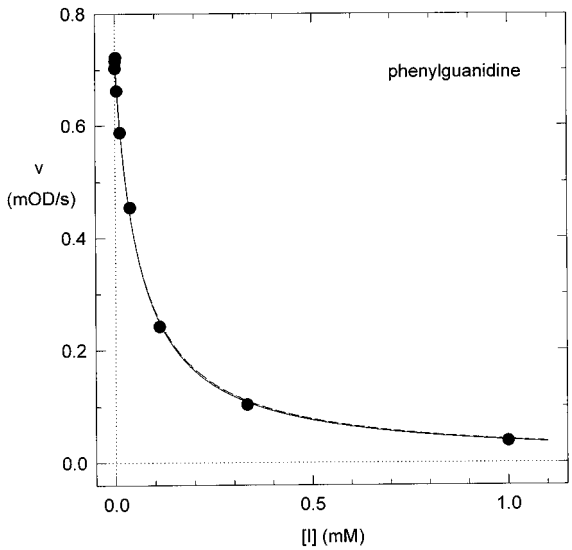


FIG. 3. Inhibition of human mast-cell tryptase by phenylguanidine. See legend to Fig. 2 for details.

based on the Dixon plot (of reciprocal velocity against the inhibitor concentration), from which the classical inhibition constant can be obtained. The fundamental problem is that this approach would require an automatic way to switch methods, from the Henderson model to the Dixon model. In addition, we found that the Dixon method fails on moderately noisy data when the extent of inhibition is small (results not shown).

The analytical procedure proposed here, for the automatic determination of inhibition constants from initial reaction velocities, is similar in principle to the method proposed by Kurganov *et al.* (13), for the computation of receptor-ligand dissociation constants from intrinsic fluorescence data. The authors noted that from any three sufficiently accurate measurements of intrinsic fluorescence (one in the presence of the ligand, one in its absence, and one conducted at a saturating concentration), it is possible to calculate immediately the receptor-ligand dissociation constant. Our procedure differs from the Kurganov method (13) by requiring only two measurements for each estimate and by iterative reweighting, which improves the precision of the estimate approximately by a factor of two to three.

The estimation method described in this paper was incorporated into a comprehensive algorithm for automatic determination of both the tight-binding and the classical inhibition constants during high-throughput screening of drug candidates. The complete algorithm proceeds in two stages. In the first stage, an estimate of K_i^{app} is made by using the weighted-average method. In the second stage, the initial estimate is refined by nonlinear least-squares regression. This technique allowed a successful automatic determination of inhibition constants that span many orders of magnitude. For example, APC-1390 is a picomolar inhibitor while phenylguanidine is a micromolar inhibitor of mast-cell tryptase. In the absence of an automatic method for making initial estimates of K_i^{app}, it would be necessary for a human operator to provide a sufficiently good guess (accurate within an order of magnitude) or else the nonlinear least-squares fit of initial velocities would fail (4).

The combined method for data analysis described in this paper has been successfully deployed to automatically determine over 100,000 inhibition constants from multiple enzyme targets. The inhibitor examples presented, with K_i^{app} ranging from picomolar to micromolar, are representative of the broad utility and robustness of the computational method. A computer program performing these calculations can be obtained via the World Wide Web at the address given in the abstract.

ACKNOWLEDGMENTS

Early stages of this research were partially supported by the American Cancer Society, Grant IRG-35-33-14 ("Artificial Intelligence System for the Evaluation of Antifolate Activity"), to P.K. We thank Sarah McCord for careful editing of the manuscript.

REFERENCES

1. Williams, J. W., and Morrison, J. F. (1979) *Methods Enzymol.* **63,** 437–467.
2. Johnson, M. L. (1992) *Anal. Biochem.* **206,** 215–225.
3. Johnson, M. L., and Faunt, L. M. (1992) *Methods Enzymol.* **210,** 1–37.
4. Johnson, M. L. (1994) *Methods Enzymol.* **240,** 1–23.
5. Butterfield, J. H., Weiler, D. A., Hunt, L. W., Wynn, S. R., and Roche, P. C. (1990) *J. Leukoc. Biol.* **47,** 409–419.
6. Rice, K., Tanaka, R., Katz, B., Numerof, R., and Moore, W. (1998) *Curr. Pharm. Design* **4,** 393–408.
7. Henderson, P. (1972) *Biochem. J.* **127,** 321–333.
8. Duggleby, R. G. (1984) *Comput. Biol. Med.* **14,** 447–455.
9. Greco, W. R., and Hakala, M. T. (1979) *J. Biol. Chem.* **254,** 12104–12109.
10. Szedlacsek, S. E., and Duggleby, R. G. (1995) *Methods Enzymol.* **249,** 144–181.
11. Morrison, J. F. (1969) *Biochim. Biophys. Acta* **185,** 269–286.
12. Williams, M. R., Poe, M., Greenfield, N. J., Hirschfield, J. M., and Hoogsteen, K. (1973) *J. Biol. Chem.* **248,** 6375–6379.
13. Kurganov, B. I., Sugrobova, N. P., and Yakovlev, V. A. (1972) *FEBS Lett.* **19,** 308–310.