Structure–Activity Relationships for a Collection of Structurally Diverse Inhibitors of Purine Nucleoside Phosphorylase

Adriano D. ANDRICOPULO^{*, a} and Rosendo A. YUNES^b

College of Pharmacy, University of Michigan,^{*a*} *Ann Arbor, MI 48109, U.S.A. and Department of Chemistry, Universidade Federal de Santa Catarina,*^{*b*} *Florianópolis-SC, 88040–900, Brazil.* Received June 12, 2000; accepted November 2, 2000

Values of inhibition constants, K_i , and concentrations required for 50% inhibition, IC₅₀, for a collection of structurally diverse competitive inhibitors of calf spleen purine nucleoside phosphorylase have been determined employing inosine as substrate. These values have been employed to create predictive quantitative structure–activity relationships (QSAR) which link structure to values of K_i and IC₅₀. These QSAR models have substantial power to predict values and the associated uncertainties for K_i and IC₅₀ for unknown, structurally diverse inhibitors of purine nucleoside phosphorylase.

Key words purine nucleoside phosphorylase; enzyme inhibition; QSAR model

Purine nucleoside phosphorylase (PNP, EC 2.4.2.1) is a key enzyme in the purine salvage and degradation pathway.¹) Inhibitors of PNP have potential therapeutic value in the treatment of T-cell proliferative diseases such as T-cell leukemias or lymphomas, for prevention of transplant rejection, and in the treatment of T-cell autoimmune diseases such as rheumatoid arthritis and lupus.^{2,3}) Despite the therapeutic potential of PNP inhibitors, which has been appreciated for more than fifteen years, and the several classes of PNP inhibitors that have been described in the literature,⁴) only one has reached advanced clinical trials.⁵

Recently, we have determined values of K_i and IC₅₀ for a series of inhibitors of purine nucleoside phosphorylase discovered by scientists at BioCryst Pharmaceuticals and developed useful QSAR models based on them.⁶⁾ These QSAR models are, however, less robust than one would like as a consequence of limited structural diversity and a clustering of many of the property values over a rather narrow range. We have now amplified our data set through the inclusion of 24 PNP inhibitors discovered by scientists at Parke–Davis Pharmaceuticals. These inhibitors provide the desired structural diversity and cover a greater range of property values with less clustering. In sum, the amplified data set is an improved data set for the purposes of QSAR model development. The results of modeling this data set are reported herein.

Results

Determination of Inhibition Constants The data set employed in this work consists of two parts: (i) a set of inhibitors from BioCryst Pharmaceuticals, **001**—**034**, and (ii) a set of inhibitors from Parke–Davis Pharmaceuticals, **035**— **058**. We have previously reported values of K_i and IC₅₀ for the BioCryst inhibitors.⁶⁾ The structures of these inhibitors are collected in Figure 1. Table 1 includes structures for the Parke–Davis inhibitors, together with values of K_i , IC₅₀, and the ratio IC₅₀/ K_i . This set of 24 new inhibitors of PNP is structurally distinct from the set of 34 BioCryst inhibitors. Nine of the 24 inhibitors are 9-substituted-8-aminoguanines; one is 8-aminothioguanine (compound **041**); one is 9-substituted-8-thioguanine (compound **049**); one is 8-bromoguanosine (compound **053**); two are 3-substituted-8aminoguanines (compounds **051** and **057**); and a group of 9 inhibitors are structurally diverse from the above. These include compounds **036**, **040**, **042**, **046**, **052**, **054**–056, and **058**.

For 53 of the 58 inhibitors of PNP employed in this work, we have determined values of K_i ; these vary from 0.004 to 190 μ M, a factor of about 48000-fold. Values of IC₅₀ for the complete set of 58 inhibitors vary from 0.0148 to 515 μ M, a factor of 35000-fold. The distribution of measured values of log IC₅₀ and log K_i is provided in Figs. 2 and 3, respectively. These values span approximately five orders of magnitude and property values, although weighted toward the high-potency end of the spectrum, are acceptably distributed across the range of values.

Statistical Modeling This data set of PNP inhibitors is well-suited for QSAR modeling. Inclusion of the Parke–Davis inhibitors has augmented the structural diversity of the data set, increased the dynamic range of property values, and provided a more nearly uniform distribution of property values across the range of values (Figs. 2 and 3).

Two QSAR models were created, one for values of K_i and one for values of IC50. In each case, the data set was divided into five non-overlapping subsets, each subset spanning as much of the range of property values as possible. Sets of molecular descriptors based on earlier experience were selected for the modeling work.^{6,7)} Ten modeling runs were carried out for each data set. In each case, three of the subsets of the data set were employed as training set and the other two were employed as test set. All possible combinations of three subsets (ten) were employed as training sets. Thus, the test sets occupy 67% of the structure space occupied by the training sets. Proceeding in this way, each molecule in the data set appears in a training set six times and in a test set four times. Thus, we generally obtain six estimates of the property value for each compound as a member of a training set and four predictions of the property value for each compound as a member of a test set.

In Table 2, measured, estimated, and predicted values of $\log IC_{50}$ for the Parke–Davis subset of PNP inhibitors are collected. The values reported are averages of all qualified estimates and predictions, together with the associated standard deviations. Differences between measured and estimated or predicted values are also included in Table 2. A correspond-



Fig. 1. Structures for the BioCryst Subset of PNP Inhibitors Employed in This Work

ing set of values for $\log K_i$ for the same subset of inhibitors is provided in Table 3.

CN

Discussion

For all 58 inhibitors employed in this work, a plot of estimated and predicted values of $\log IC_{50}$ against the corresponding measured values is provided in Fig. 4. A corresponding plot for values of $\log K_i$ is provided in Fig. 5. In both cases, there is excellent concordance between estimates and predictions, on the one hand, and measured values, on the other. Work described herein has resulted in two persuasive QSAR models [for IC₅₀ and K_i] for inhibitors of calf spleen PNP. The principal features of these models include the following.

First, the model does an excellent job of characterizing the training sets. Note specifically the excellent agreement between measured and estimated values (Figs. 4 and 5) and the small standard deviations obtained for repeated calculation

Table 1. Values of Inhibition Constants for a Family of Inhibitors of Calf Spleen Purine Nucleoside Phosphorylase^a

Inhibitor	Structure	<i>K</i> _i (µм)	IC ₅₀ (µм)	IC ₅₀ /K _i
035	HN NH2 H2N N NH2	2.71 ± 0.06 n=3	4.17 ± 0.41 n=3	1.54
036	CH ₃ NH ₂ NH NH	3.57 ± 0.20 n=3	7.47 ± 0.18 n=3	2.09
037		0.0214 ± 0.0029 n=3	0.0360 ± 0.0027 n=3	1.68
038		0.187 ± 0.012 n=2	0.362 ± 0.014 n=2	1.94
039	HN H2 H2N H2 OH OCCH3	0.661 ± 0.029 n=2	1.52 ± 0.21 n=2	2.30
040		3.49±0.18 n=2	7.03 ± 0.20 n=2	2.01
041		3.09 ± 0.25 n=3	8.49 ± 0.21 n=3	2.75
042	$ \begin{array}{c} $	21.68±1.99 <i>n</i> =2	39.63 ± 2.28 n=2	1.83

-

Table 1. Continued

Inhibitor	Structure	$K_{ m i}$ (μ м)	IC ₅₀ (µм)	IC_{50}/K_i	
043		0.085 ± 0.015 n=2	0.092 ± 0.0021 n=2	1.10	
044		0.561 ± 0.025 n=2	1.043 ± 0.031 n=2	1.86	
045	HN KN NH2 H2N KN KN	13.26 ± 1.63 n=3	19.32 ± 1.93 n=4	1.46	
046	H2N NH O	0.342 ± 0.038 n=2	0.674 ± 0.032 n=2	1.97	
047		0.445 ± 0.030 n=2	0.852 ± 0.038 n=3	1.91	
048		0.321 ± 0.018 n=2	0.497 ± 0.026 n=2	1.55	
049		1.032 ± 0.18 n=2	2.866 ± 0.29 n=2	2.78	
050	H ₂ N H ₂ N N N N N N N N N N N N N N N N N N N	1.231 ± 0.13 n=2	1.971 ± 0.15 n=2	1.69	

_

Inhibitor	Structure	<i>K</i> _i (<i>µ</i> м)	IC ₅₀ (µм)	IC ₅₀ / <i>K</i> _i
051		1.16±0.19 n=2	2.71 ± 0.22 n=2	2.35
052	H ₂ N NH O	0.806 ± 0.02 n=2	1.839 ± 0.06 n=2	2.28
053	HN HOCH2 HO OH	190.48±7.85 n=2	514.96±17.18 <i>n</i> =3	2.70
054		47.53±4.39 n=2	126.18 ± 11.27 n=2	2.65
055	$H_{3C} \rightarrow H_{2N} \rightarrow H$	6.01±0.66 n=2	13.30±1.02 n=2	2.21
056		7.77 ± 0.43 n=2	22.84 ± 1.82 n=2	2.94
057		0.974 ± 0.08 n=2	1.758 ± 0.09 n=2	1.80
058	NH2 NH2 NH2	104.23 ± 5.85 n=2	389.04 ± 16.97 n=2	3.73

a) All values of K_i and IC₅₀ were determined at pH 7.4, 1 mM phosphate, and 25 °C. Values of IC₅₀ were determined at 10 μ M inosine.



Fig. 2. Distribution of Values of log IC_{50} for the Set of Inhibitors of Calf Spleen Purine Nucleoside Phosphorylase



Fig. 3. Distribution of Values of $\log K_i$ for the Set of Inhibitors of Calf Spleen Purine Nucleoside Phosphorylase

of these estimates (Tables 2 and 3). In every case, the estimated values fall within the limits of the estimated measurement errors. There are no outliers; qualified estimates were obtained for all members of the training sets.

Second, the model does an excellent job of predicting values of molecules in the test sets (Figs. 4 and 5). In nearly all cases, the predictions fall within the estimated measurement error and in no case is the prediction different from the measured value by more than two standard deviations of the measurement error. Here too there are no outliers; qualified predictions were obtained for all members of test sets. Note that the predicted values are independent of the nature of the training set as evidenced by the small standard deviations from the average value obtained from repeated calculations (Tables 2 and 3).

Third, the models cover substantial structure space (Fig. 1 and Table 1). Thus, these models will prove predictive for novel compounds within, and perhaps somewhat outside, this structure space. The models are robust.

Fourth, one has a substantial understanding of the reliability of the predicted values, based on the standard deviations from the average of the calculated values and the agreement between measured and calculated values (Tables 2 and 3).

In sum the QSAR models created in this work meet the

Table 2. Measured, Estimated, and Predicted Values of Log IC_{50} for a Series of Inhibitors of Calf Spleen Purine Nucleoside Phosphorylase^{*a*}

=

T 1 '1 '		Log IC ₅	Δ		
Inhibitor	Measured	Estimated	Predicted ^b	Estimated ^c	Predicted ^d
001	-1.76	-1.58 ± 0.04	-1.52 ± 0.08	-0.18	-0.24
002	-1.78	-1.61 ± 0.05	-1.60 ± 0.01	-0.17	-0.18
003	-1.59	$-1.50 {\pm} 0.03$	-1.50 ± 0.03	-0.09	-0.09
004	-1.65	$-1.55 \!\pm\! 0.04$	-1.59 ± 0.03	-0.10	-0.06
005	-0.99	$-1.08 \!\pm\! 0.04$	-1.05 ± 0.04	0.09	0.06
006	-1.44	$-1.40 \!\pm\! 0.07$	-1.33 ± 0.04	-0.04	-0.11
007	-0.62	$-0.73 \!\pm\! 0.04$	-0.68 ± 0.04	0.11	0.06
008	-1.68	$-1.60 {\pm} 0.04$	-1.62 ± 0.04	-0.08	-0.06
009	-1.69	$-1.63 \!\pm\! 0.01$	-1.63 ± 0.05	-0.06	-0.06
010	-1.83	-1.61 ± 0.05	-1.70 ± 0.10	-0.22	-0.13
011	-1.76	-1.67 ± 0.07	-1.64 ± 0.07	-0.09	-0.12
012	-1.72	-1.65 ± 0.03	-1.60 ± 0.03	-0.07	-0.12
013	-1.45	-1.42 ± 0.08	-1.35 ± 0.06	-0.03	-0.10
014	-1.79	-1.66 ± 0.02	-1.63 ± 0.04	-0.13	-0.16
015	-1.33	-1.28 ± 0.05	-1.29 ± 0.10	-0.05	-0.04
016	-1.57	-1.52 ± 0.05	-1.54 ± 0.07	-0.05	-0.03
017	-0.91	-0.99 ± 0.07	-1.02 ± 0.05	0.08	0.11
018	-1.10	-1.13 ± 0.22	-1.10 ± 0.12	-0.03	-0.06
019	-0.01	-0.39 ± 0.03	-0.71 ± 0.09 -0.75 ± 0.07	-0.02	0.10
020	-0.74	-0.84 ± 0.03	-0.73 ± 0.07 -0.28 ± 0.17	0.10	0.01
021	0.19	0.31 ± 0.09 0.19+0.14	0.23 ± 0.17 0.12+0.19	0.12	0.09
022	0.254	0.15 ± 0.14 0.25 ± 0.11	0.12 ± 0.19 0.20 ± 0.09	-0.08	-0.03
024	-1.61	-1.54 ± 0.04	-1.54 ± 0.04	-0.07	-0.07
025	-1.20	-1.29 ± 0.04	-1.18 ± 0.13	0.09	-0.02
026	-0.87	-0.92 ± 0.07	-0.88 ± 0.08	0.05	0.01
027	1.75	1.66 ± 0.03	1.40 ± 0.18	0.09	0.35
028	1.30	$1.16 {\pm} 0.01$	1.01 ± 0.19	0.14	0.29
029	0.36	$0.35 {\pm} 0.06$	0.33 ± 0.09	0.01	0.03
030	-1.82	$-1.64 \!\pm\! 0.02$	-1.64 ± 0.05	-0.18	-0.18
031	-1.40	-1.44 ± 0.07	-1.50 ± 0.02	0.04	0.10
032	-1.38	-1.34 ± 0.06	-1.30 ± 0.01	-0.04	-0.08
033	-0.82	-0.92 ± 0.03	-0.99 ± 0.08	0.10	0.17
034	-1.10	-1.16 ± 0.06	-1.16 ± 0.07	0.06	0.06
035	0.62	0.55 ± 0.07	0.57 ± 0.12	0.07	0.05
030	0.87	0.91 ± 0.09	$0.8/\pm0.14$	-0.04	0 17
037	-1.45	-1.24 ± 0.04	-1.20 ± 0.03	-0.19	-0.17
030	-0.044	-0.08 ± 0.07 0.10 ± 0.03	$-0.004/3 \pm 0.11$	-0.04	-0.04
040	0.130	0.19 ± 0.03 0.91 ± 0.04	0.25 ± 0.03 0.76 ± 0.11	-0.01	0.07
041	0.929	0.91 ± 0.01 0.92 ± 0.02	0.86 ± 0.09	0.00	0.07
042	1.60	1.60 ± 0.04	1.64 ± 0.12	0	-0.04
043	-1.04	-0.94 ± 0.09	-0.92 ± 0.05	-0.10	-0.12
044	0.018	0.05 ± 0.04	0.07 ± 0.07	-0.03	-0.05
045	1.29	1.21 ± 0.08	1.22 ± 0.08	0.08	0.07
046	-0.172	$-0.25 \!\pm\! 0.06$	-0.10 ± 0.07	0.08	-0.07
047	-0.070	$-0.06 \!\pm\! 0.06$	-0.08 ± 0.07	-0.01	0.01
048	-0.304	-0.39 ± 0.06	-0.25 ± 0.10	0.09	-0.05
049	0.457	0.38 ± 0.06	0.41 ± 0.07	0.08	0.05
050	0.295	0.24 ± 0.04	0.31 ± 0.14	0.05	-0.02
051	0.433	0.44 ± 0.08	0.38 ± 0.09	-0.01	0.05
052	0.264	0.30 ± 0.12	0.29 ± 0.11	-0.04	-0.03
053	2./1	2.00 ± 0.04	2.39 ± 0.05	-0.05	0.31
054	2.10	2.09 ± 0.04 1 05 ± 0.04	2.03 ± 0.19 0.00+0.06	0.01	0.05
056	1 36	1.03 ± 0.00 1.29 ± 0.07	130+0.14	0.07	0.15
057	0.245	0.29 ± 0.04	0.34 ± 0.02	-0.04	-0.09
058	2.59	2.62 ± 0.05	2.28 ± 0.18	-0.03	0.30

a) All values of IC₅₀ (μ M) were determined at pH 7.4, 1 mM phosphate, and 25 °C. *b*) Predicted values are averages of the four values obtained for each inhibitor as a member of test sets; estimated values are averages of the six values obtained for each inhibitor as a member of training sets. *c*) The difference between logarithms of measured and estimated values. *d*) The difference between logarithms of measured and predicted values.

Table 3. Measured, Estimated, and Predicted Values of $Log K_i$ for a Series of Inhibitors of Calf Spleen Purine Nucleoside Phosphorylase^{*a*}

T 1 1 1 1		$\log K_{i}$	Δ		
Innibitor	Measured	Estimated	Predicted ^b	Estimated ^c	Predicted ^d
001	-1.98	-1.80 ± 0.03	-1.80 ± 0.06	-0.18	-0.18
002	-1.96	-1.83 ± 0.03	-1.79 ± 0.02	-0.13	-0.17
003	-1.72	-1.69 ± 0.04	-1.71 ± 0.04	-0.03	-0.01
004	-1.81	-1.74 ± 0.05	-1.69 ± 0.05	-0.07	-0.12
005	-1.19	-1.30 ± 0.03	-1.34 ± 0.03	-0.11	-0.15
006	-1.78	-1.70 ± 0.02	-1.68 ± 0.06	-0.08	-0.10
007	-0.962	-1.02 ± 0.04	-1.01 ± 0.02	0.06	0.05
008	-1.88	-1.84 ± 0.03	-1.93 ± 0.03	-0.04	0.05
009	-2.04	-1.97 ± 0.03	-1.95 ± 0.07	-0.07	-0.09
010	-2.03	-1.87 ± 0.03	-1.79 ± 0.09	-0.16	-0.24
011	-2.07	-1.95 ± 0.04	-1.91 ± 0.08	-0.12	-0.16
012	-1.94	-1.92 ± 0.05	-1.85 ± 0.06	-0.02	-0.09
013	-2.40	-2.30 ± 0.15	-2.29 ± 0.02	-0.10	-0.11
014	-1.96	-1.88 ± 0.05	-1.87 ± 0.04	-0.08	-0.09
015	-2.05	-1.94 ± 0.05	-1.94 ± 0.09	-0.09	-0.09
016	-1.87	-1.78 ± 0.02	-1.75 ± 0.02	-0.09	-0.12
017	-1.42	-1.46 ± 0.03	$-1.46 {\pm} 0.08$	0.04	0.04
018	-1.45	-1.41 ± 0.09	-1.26 ± 0.03	-0.04	-0.19
019	-0.827	-0.981 ± 0.01	-0.853 ± 0.21	0.15	0.02
020	-1.08	-1.14 ± 0.04	-1.10 ± 0.09	0.06	0.02
021	-0.475	-0.513 ± 0.09	$-0.594 {\pm} 0.07$	0.04	0.11
022	-0.065	$-0.132 {\pm} 0.08$	-0.191 ± 0.04	0.06	0.12
023	-0.053	-0.105 ± 0.07	-0.018 ± 0.07	0.05	-0.04
024	-1.73	-1.75 ± 0.02	-1.76 ± 0.02	0.02	0.03
025	-1.60	-1.59 ± 0.05	-1.61 ± 0.04	0	0.02
026	-1.18	-1.24 ± 0.08	-1.22 ± 0.08	0.06	0.04
027	1.39	1.34 ± 0.13	1.05 ± 0.15	0.05	0.34
028	1.01	0.825 ± 0.14	0.692 ± 0.14	0.19	0.32
029	-0.008	-0.021 ± 0.03	-0.10 ± 0.04	0.01	0.09
035	0.433	0.366 ± 0.05	0.392 ± 0.09	0.07	0.04
036	0.552	0.552 ± 0.07	0.512 ± 0.02	0	0.04
037	-1.6/	-1.43 ± 0.07	-1.31 ± 0.15	-0.24	-0.36
038	-0.75	$0.68 / \pm 0.04$	0.081 ± 0.04	-0.04	-0.05
039	-0.180	$-0.1/2 \pm 0.06$	-0.202 ± 0.07	-0.01	0.02
040	0.545	$0.3/3\pm0.04$	0.581 ± 0.04	-0.03	-0.04
041	1.24	0.46 ± 0.00	0.41 ± 0.04	0.01	0.08
042	-1.07	-0.98 ± 0.03	-0.95 ± 0.08	-0.02	-0.00
043	-1.07 -0.25	-0.98 ± 0.04 -0.23 ± 0.06	-0.93 ± 0.08 -0.20 ± 0.12	-0.09	-0.12
044	1.12	0.23 ± 0.00 1 11 ± 0 13	0.20 ± 0.12	0.02	0.05
045	-0.466	-0.555 ± 0.06	-0.562 ± 0.09	0.01	0.10
040	-0.352	-0.324 ± 0.10	-0.273 ± 0.05	-0.03	-0.08
048	-0.493	-0.487 ± 0.04	-0.506 ± 0.14	-0.01	0.00
049	0.0137	0.061 ± 0.04	0.000 ± 0.011 0.093 ± 0.06	-0.05	-0.08
050	0.090	0.10 ± 0.08	0.036 ± 0.09	-0.01	0.05
051	0.063	0.05 ± 0.05	0.07 ± 0.03	0.01	-0.01
052	-0.094	-0.151 ± 0.07	-0.247 ± 0.05	0.06	0.15
053	2.28	2.27 ± 0.05	2.52 ± 0.05	0.01	-0.24
054	1.68	1.66 ± 0.04	1.55 ± 0.08	0.02	0.13
055	0.779	$0.712 {\pm} 0.04$	0.689 ± 0.09	0.07	0.09
056	0.89	$0.89 {\pm} 0.09$	0.79 ± 0.09	0	0.10
057	-0.011 -	0.00114 ± 0.08	0.08 ± 0.02	-0.01	-0.09
058	2.018	$1.97 {\pm} 0.06$	1.81 ± 0.06	0.05	0.21

a) All values of K_i (μ m) were determined at pH 7.4, 1 mm phosphate, and 25 °C. *b*) Predicted values are averages of the four values obtained for each inhibitor as a member of test sets; estimated values are averages of the six values obtained for each inhibitor as a member of training sets. *c*) The difference between logarithms of measured and estimated values. *d*) The difference between logarithms of measured and predicted values.

objectives and standards for such models laid out earlier.

Experimental

All inhibitors employed in this study were synthesized by scientists at BioCryst Pharmaceuticals or Parke–Davis Pharmaceuticals and were the gifts of those organizations. Calf spleen PNP (phosphate-free, purity greater



Fig. 4. Plot of Logarithms of Measured Values of IC_{50} against the Corresponding Estimated (Solid Diamonds) and Predicted (Open Squares) Values for a Set of Inhibitors of Calf Spleen Purine Nucleoside Phosphorylase



Fig. 5. Plot of Logarithms of Measured Values of K_i against the Corresponding Estimated (Solid Diamonds) and Predicted (Open Squares) Values for a Set of Inhibitors of Calf Spleen Purine Nucleoside Phosphorylase

than 98%) and xanthine oxidase were the best grades available from Sigma Chemical and were used without further purification. Other reagents were obtained commercially from either Sigma or Aldrich and were the best grade available.

Kinetic measurements were carried out spectrophotometrically with the aid of an HP 8453 diode array spectrophotometer and employing a standard coupled assay⁸⁾ as previously described.⁶⁾ All measurements were made at 25 °C and pH 7.4. Kinetic parameters were estimated from the collected data employing the Leonora statistics package. Values of pH were determined with a Radiometer PHM240 pH meter.

Statistical modeling was carried out as previously described.^{6,7)} The descriptors employed in the modeling include the traditional 2D and 3D QSAR descriptors, a number of descriptors which identify structural elements within each molecule (*e.g.*, methyl groups, chlorine atoms), as well as a collection of novel descriptors based on transferable atom equivalent technology.^{9,10)} Basically, this work involves the following steps. First, the training set is divided into a large number of overlapping subsets, employing a small number of molecular descriptors and a mixture of regression models algorithm.⁹⁾ Second, several linear subset QSAR models are constructed for each subset but within the training set. Those subsets for which a small number of successful subset QSAR models can be constructed are retained, together with their models, in the final QSAR model. Ordinarily, less than one percent of the subsets initially created meet this criterion and are retained. This process is continued until a large number of subsets have been qualified for

their ability to generate QSAR models which predict the property value for molecules outside the subset but within the training set to within 1-2 standard deviations of the experimental error of the measurements. The final QSAR model employed in this work contained several thousand qualified subsets, each containing 3 to 5 subset QSAR models. Thus, the final QSAR model contained approximately 10000 subset QSAR models. Finally, the descriptors for each of the molecules in the test set are compared with the average of the descriptors for all molecules in each subset. For those cases in which these descriptors for the test set molecule and those in the subset are adequately concordant (suggesting that the structure/descriptor space occupied by the test molecule falls within that spanned by the subset), the subset QSAR models for that subset are employed to make predictions for the property value of the test set molecule. This procedure is continued until all subsets within the final QSAR molecule have been so examined. Thus, at the end of the procedure, each molecule in the test set will have multiple predictions. The number of such predictions varies from several hundreds for some molecules in the test set to perhaps 10 for other molecules. The reported predictions are the simple numerical average of these predictions.

Acknowledgments We are grateful to Dr. E. H. Cordes (University of Michigan) and Dr. J. W. Frazer for continuous support and valuable scientific guidance. Thanks are due to Parke–Davis Pharmaceuticals for the gift of the

inhibitors (035-058) employed in this work, and CAPES (Brazil) for partial financial support.

References

- Mao C., Cook W. J., Zhou M., Federov A. A., Almo S. C., Ealick S. E., Biochemistry, 37, 7135–7146 (1998).
- 2) Sircar J. C., Gilbertsen R. B., Drugs of the Future, 13, 653–688 (1988).
- 3) Montgomery J. A., Exp. Opin. Invest. Drugs, 3, 1303-1313 (1994).
- Beauchamp L. M., Tuttle J. V., Rodriguez M. E., Sznaidman M. L., J. Med. Chem., 39, 949–956 (1996).
- 5) Morris P. E., Montgomery J. A., *Exp. Opin. Ther. Pat.*, **8**, 283–299 (1998).
- Farutin V., Masterson L., Andricopulo A. D., Cheng J., Riley B., Hakimi R., Frazer J. W., Cordes E. H., *J. Med. Chem.*, 42, 2422–2431 (1999).
- Andricopulo A. D., Yunes R. A., Cechinel-Filho V., Nunes R. J., Frazer J.W., Cordes E. H., *Pharmazie*, 54, 698–704 (1999).
- 8) Tuttle J. V., Krenitsky T. A., J. Biol. Chem., 259, 4065-4069 (1984).
- 9) Breneman C. M., Rhem M., J. Comput. Chem., 18, 182-197 (1997).
- Breneman C. R., Thompson T., Rhem M., Dung M., Comput. Chem., 19, 161–179 (1995).