4502 J. Med. Chem. 2010, 53, 4502–4510
DOI: 10.1021/jm100301x

Journal of Medicinal Chemistry Article

Analysis of Kinase Inhibitor Selectivity using a Thermodynamics-Based Partition Index

Alan C. Cheng,*,† John Eksterowicz,§ Stephanie Geuns-Meyer,‡ and Yaxiong Sun∥

†Molecular Structure Department, and ‡Medicinal Chemistry Department, Amgen Inc., 360 Binney Street, Cambridge, Massachusetts 02142, §Molecular Structure Department, Amgen Inc., 1120 Veterans Boulevard, South San Francisco, California 94080, and ∥Molecular Structure Department, Amgen Inc., One Amgen Center Drive, Thousand Oaks, California 91320

Received March 6, 2010

In the quest for safe, efficacious kinase inhibitors as drugs, selectivity is often assessed early using kinase profiling panels. Here we present a selectivity index based on thermodynamics principles that can help in analysis of the resulting data. The "partition" selectivity index is easy to calculate and is applicable in certain situations where other widely used indices are not. It is uniquely useful in analysis of small, focused selectivity panel data frequently encountered in medicinal chemistry hit-to-lead and lead optimization. For larger "kinome" panels, the partition index allows assessment of selectivity relative to a kinase or multiple kinases of interest.

Introduction

Reducing dose-limiting off-target pharmacology is a key goal in drug discovery, but achieving selectivity can be challenging for targets from large gene families. Protein kinases are prime examples because they share a large, 518-member gene family[1] and a highly conserved binding site intended for the ubiquitous ATP[a] substrate.[2] Much progress in generating selective kinase inhibitors has been made in recent years, due in part to automation-enabled "selectivity panels" of enzyme inhibition assays as well as advances such as structure-based drug design.[3−8] Selectivity panels in the drug discovery setting generally come in two flavors: smaller panels that include a few choice enzyme assays, and larger panels that represent the kinome. Both types of panels produce large amounts of data that are often analyzed in ad hoc ways. A single number summarizing the selectivity of a particular compound, a selectivity index, can be useful for systematic analysis of the data sets, and allows rank-ordering of compounds and detection of emerging structure−activity trends.

Because of their utility, efforts to create selectivity indices are not new. A ratio of measured IC_{50}'s defines the "fold-selectivity" between a desired target and an undesired target. For more than two targets, composing a selectivity index is not as straightforward. Ambit workers use a counting scheme for reporting the selectivity of compounds in their 420+ kinase profiling panel,[6] which returns either percent of control (POC) or binding affinity (K_d) values. The Ambit score, S_X, represents the fraction of kinases inhibited beyond a cutoff, X, and ranges from zero for perfectly selective compounds to one for completely unselective compounds (also see Figure 1A).[6] For POC values, which range from approximately 100% for no inhibition to 0% for full inhibition, a commonly used score, S_{35}, represents the fraction of kinases inhibited at <35% POC. For K_d or IC_{50} data, indices such as $S_{3\,\mu M}$ are used. The Ambit counting score is simple and easy to interpret but does not capture nuances. A cutoff at 35 POC results in kinases inhibited at 34 POC and 1 POC being placed in the same category, while kinases inhibited at 34 POC and 35 POC are placed in opposite categories. The Gini coefficient was proposed as a more nuanced selectivity index and is calculated by ordering kinase inhibition values from a panel screen by increasing inhibition, creating bar charts of the cumulative inhibition for each kinase and kinases with lesser inhibition, summing up the areas of the bars, and dividing the remaining area of the half box by the total area of the half box,[9] as illustrated in Figure 1B.

In practice, both the S_X scores and Gini coefficient are useful for analyzing larger panels intended to represent the kinome. However, we found that both indices have difficulty distinguishing inhibitor profiles in the smaller focused hit-to-lead screening panels often encountered in medicinal chemistry programs. We developed a "partition" selectivity index to address this issue and found that the new index also has unique advantages for analysis of large panels.

Results and Discussion

We set out to develop a selectivity index based on the thermodynamics of protein−ligand binding that could be used in analyzing both small and large selectivity panels. Conceptually, we define the selectivity problem as a competition assay in which the biological targets in the panel are placed in a test tube and they then compete for binding to a common pool of inhibitor, as depicted in Figure 1C. At thermodynamic equilibrium, the fraction of bound inhibitor that is bound to a particular target is the "partition" index, P. The values of P therefore range from zero, indicating 0% occupancy, to one, indicating 100% occupancy. An inhibitor with

*To whom correspondence should be addressed. Phone: 617-444-5411 Fax: 617-621-3908. E-mail: alan.cheng@amgen.com.

[a]Abbreviations: POC, percent of control; SAR, structure−activity relationships; ATP, adenosine triphosphate; EGFR, epidermal growth factor receptor; FLT3, fms-like tyrosine kinase; PDGFR, platelet-derived growth factor receptor; DYRK, dual-specificity tyrosine phosphorylation-regulated kinase.

 Published on Web 05/11/2010

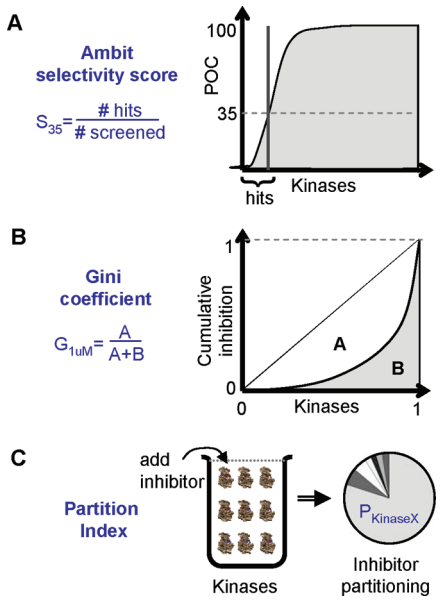
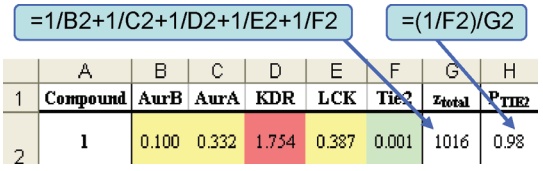


Figure 1. Schematic depiction of selectivity indices. (A) The Ambit selectivity score uses a cutoff value (35 POC in this example) to determine the fraction of kinases that are "hit" by an inhibitor. (B) The Gini coefficient is shown for a similar POC data set and is described in the text and in Graczyk (2007).[9] The *y*-axis is the cumulative fraction of total inhibition, and the *x*-axis is the cumulative fraction of kinases. (C) Conceptual depiction of thermodynamic partitioning index. A pool of inhibitor is added to a flask containing the kinases present in the panel, where the kinases are present at concentrations in vast excess of the inhibitor, and the kinases are present in equal amounts. The pie chart indicates how the inhibitor partitions among the kinases in the panel, with KinaseX being the kinase of interest.



	A	B	C	D	E	F	G	H
1	Compound	AurB	AurA	KDR	LCK	Tie2	z_{total}	P_{TIE2}
2	1	0.100	0.332	1.754	0.387	0.001	1016	0.98

Figure 2. Snapshot of Excel (version 2003, Microsoft Corp) spreadsheet approach to calculating Tie-2 partitioning for compound **1** from Table 1. Excel equations are shown in the blue bubbles.

$P = 0.97$ for the kinase of interest would indicate a quite selective inhibitor, where 97% of the bound inhibitor is bound to the particular kinase and just 3% is bound to other kinases in the test tube. Although we focus on protein kinases here, the selectivity index is general and could be used with any pharmacologic profiling panel.

The partition index is based on the thermodynamics concepts of Boltzmann distributions and partition functions.[10] A simplified derivation of the selectivity index is presented here, with a more detailed version presented in the Methods section. To calculate the index, an inhibitor's occupancy of a particular kinase, "KinaseX", is first described as

$$z_{\mathrm{KinaseX}} = 1/K_{\mathrm{d,KinaseX}}$$

where K_d is the binding affinity of the inhibitor for the kinase in nM units, and IC_{50}'s can be substituted for K_d's. We next compute the sum of z values over all n kinases in the panel, which includes KinaseX:

$$z_{\mathrm{total}} = z_1 + z_2 + \ldots + z_n$$

The thermodynamic partition index can then be defined as partitioning of inhibitor to KinaseX over other kinases in the panel, or

$$P_{\mathrm{KinaseX}} = \frac{z_{\mathrm{KinaseX}}}{z_{\mathrm{total}}}$$

Partition index values are additive based on the thermodynamics theory. For instance, selectivity for two kinases is $P_{\mathrm{both}} = P_{\mathrm{Kinase1}} + P_{\mathrm{Kinase2}}$. Adding up P values for all n kinases in the panel would yield $P = 1.0$. The equations are easily implemented in a spreadsheet program such as Excel, as shown in Figure 2.

Hit-to-Lead Analysis. We first illustrate use of the partition index by analyzing hit-to-lead data from a Tie-2 drug discovery effort.[11] Medicinal chemistry efforts typically receive regular results from a few key selectivity target assays, and the Tie-2 effort is a good example of this. To help identify Tie-2 inhibitors with selectivity over Aurora A, Aurora B, KDR, and LCK, these kinases were screened weekly. Tie-2 data for the dozen compounds we discuss here are shown in Table 1, where compounds are listed in order of increasing P_{TIE2}, the thermodynamic partitioning to Tie-2 enzyme.

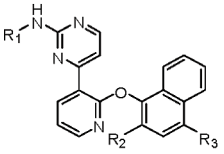
The ranking by P_{TIE2} identifies **1** as the most selective analog with $P_{\mathrm{TIE2}} = 0.98$. The *N*-((1-ethylpiperidin-4-yl)methyl) group contributes substantially to the selectivity of **1**, and this can be seen by comparing P_{TIE2} values for **1** ($P_{\mathrm{TIE2}} = 0.98$) and **5** ($P_{\mathrm{TIE2}} = 0.82$). Incorporation of a methyl on the napthyl ring at the R_2 position (see Table 1) also increases selectivity, and this can be seen by comparing P_{TIE2} values for **3** ($P_{\mathrm{TIE2}} = 0.90$, R_2 = Me) and **5** ($P_{\mathrm{TIE2}} = 0.82$, R_2 = H). In stark contrast, incorporation of a CF_3 group on the benzimidazole ring results in decreased selectivity, and this is reflected by the P_{TIE2} values for **5** ($P_{\mathrm{TIE2}} = 0.82$, des-CF3) and **10** ($P_{\mathrm{TIE2}} = 0.26$). While these structure−activity relationships (SAR) are identifiable through informal perusal of the data, the P_{TIE2} scores help in identifying the SAR quickly, objectively, and systematically.

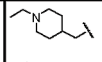
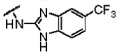
Partition values also highlight SAR that are perhaps less obvious. Although the Tie-2 potencies for **4** and **5** vary by 80-fold, the selectivity profiles actually only vary slightly, and this is reflected by P_{TIE2} values for **4** ($P_{\mathrm{TIE2}} = 0.85$) and **5** ($P_{\mathrm{TIE2}} = 0.82$) as well as by the selectivity ratios in the rightmost columns of Table 1. P_{TIE2} values also suggest a ranking of 5/6 fused ring system substitutions at the R_3 position in order of decreasing selectivity: benzimidazole (**5**) > azabenzimidazole (**6**) > benzoxazole (**8**) > benzothiazole (**9**).

The partition index aids data mining and highlights differences in selectivity profiles. It is particularly useful when analyzing large amounts of data generated from multiple series, especially when multiple off-target activities appear and disappear with small modifications, as often happens in protein kinase inhibitor drug discovery. The index correlates well with the individual selectivity ratios shown in the gray columns of Table 1. If desired, a weighting factor can be applied to each kinase if, for example, selectivity over one kinase is particularly important. Details are provided in the Methods section.

For comparison purposes, the Gini coefficient and $S_{1\ \mu M}$ score for each compound are also provided in Table 1. The Gini coefficient was developed to measure disparities in monetary wealth within defined populations. It appears to perform well when a large number of kinases (the "population") is involved and works particularly well in distinguishing large disparities in potency distributions (the "wealth"). It performs poorly in this example in part because of the small

Table 1. IC_{50}'s for Selected Tie-2 Kinase Inhibitors from Cee et al. (2009)[11] in Units of μM[a]



	R1	R2	R3	Enzyme assay IC_{50} (μM) AurB	AurA	KDR	LCK	Tie2	Selectivity index P_{TIE2}	Gini	$S_{1\mu M}$	S_{100X}	Fold selectivity AurB	AurA	KDR	LCK
1		H		0.100	0.332	1.754	0.387	0.001	**0.98**	0.16	0.80	0.25	100	332	1754	387
2		H		0.047	0.327	2.059	0.269	0.002	**0.94**	0.16	0.80	0.25	21	146	921	120
3	Me	Me		0.077	0.029	10.000	0.405	0.002	**0.90**	0.22	0.80	0.50	33	12	4316	175
4		H		0.759	1.658	10.000	10.000	0.082	**0.85**	0.42	0.40	0.50	9	20	123	123
5	Me	H		0.009	0.019	1.380	0.292	0.001	**0.82**	0.13	0.80	0.50	7	14	998	211
6	Me	H		0.454	1.067	6.358	4.799	0.086	**0.77**	0.35	0.40	1.00	5	12	74	56
7	Me	H		0.023	0.052	0.040	0.048	0.005	**0.63**	0.01	1.00	1.00	4	10	7	9
8	Me	H		0.134	0.141	0.030	0.025	0.016	**0.42**	0.03	1.00	1.00	8	9	2	2
9	Me	H		0.143	0.380	0.088	0.080	0.044	**0.40**	0.05	1.00	1.00	3	9	2	2
10	Me	H		0.056	0.016	0.069	0.132	0.027	**0.26**	0.02	1.00	1.00	2	1	3	5
11	Me	H		0.055	0.225	2.218	4.087	0.215	**0.17**	0.26	0.60	1.00	0	1	10	19
12	Me	H		0.130	0.534	0.048	1.856	0.248	**0.11**	0.16	0.80	1.00	1	2	0	7

[a] On the right half of the table, various selectivity indices are given: P_{TIE2} partition values, Gini coefficients, Ambit S values for counter-screen kinases, and selectivity ratios for each counter-screen kinase enzyme activity relative to Tie-2. IC_{50}'s are colored from green ($IC_{50} < 0.1\ \mu M$) to yellow ($0.1 \leq IC_{50} < 1\ \mu M$) to red ($IC_{50} \geq 1\ \mu M$), and IC_{50}'s reported as greater than 10 μM are defined as 10 μM. All Gini coefficients shown in this paper are calculated using the spreadsheet provided by Graczyk, with details provided in the table in the Supporting Information. On the far right, S100X and selectivity ratios relative to Tie-2 are listed.

"population". The $S_{1\ \mu M}$ score also fails to capture the differences in selectivity profiles here. The partition index performs well because it is a relative measure that takes into account the inhibitor's potency against the desired target. Therefore, a more informative S_X score would be S_{10X} or S_{100X}, which involves counting the number of kinases inhibited within a 10- or 100-fold potency window from a reference kinase potency. As illustrated in Table 1, S_{100X} provides an ordering that is reasonable and consistent with the ranking by P_{TIE2}. However, use of arbitrary cut-offs in defining the selectivity index results in coarse selectivity binning, values of 0.25, 0.50, and 1.0 in this example, that fail to identify nonobvious SAR trends. In contrast, the partition index employs, in effect, a gradual cutoff that is based on thermodynamics theory. The Gini coefficient and S_X scores were introduced, however, to facilitate analysis of data from large kinome panels, which we turn to next.

Kinome Panel Analysis. Both the S score and Gini coefficient are calculated without referencing a specific kinase of interest. The partition index, however, requires a reference kinase which can be the desired kinase of interest, or, in cases where a reference kinase is not known, the most potently inhibited kinase. To illustrate usage of the partition index in a large kinome panel, we analyze the Ambit data set containing K_d's across 290 kinases, with mutant kinases excluded.[6] In Table 2, P_{MAX} represents inhibitor partitioning to the most potently inhibited kinase, and $S_{3\ \mu M}$ is the selectivity index used by Karaman et al.[6] The three most selective inhibitors in the set based on P_{MAX} values, PI-103 (**13**),[12] GW-2580 (**14**),[13] and VX-745 (**15**),[14] fall within the top six compounds when ranked by $S_{3\ \mu M}$ values. The relative selectivity of inhibitors reflected by either P_{MAX} or $S_{3\ \mu M}$ values is similar in the majority of cases. However, there are a few cases where they differ dramatically, and these are indicated by color in Figure 3A.

Three of the dramatic differences involve relatively weak inhibitors. SB-431542 (**21**)[15] is ranked third in selectivity by $S_{3\ \mu M}$ but a low 31st by P_{MAX}. CP-724714 (**18**)[16] is ranked fourth by $S_{3\ \mu M}$ but 22nd by P_{MAX}. Roscovitine (**22**)[17] is ranked sixth by $S_{3\ \mu M}$ but 32nd by P_{MAX}. In each of these three cases, the greatest potencies measured for any kinase in the panel range from 42 to 260 nM. The partition index is a relative measure that is defined by fold-differences from the most potently inhibited kinase, and it essentially shifts the "selectivity window" based on that inhibition value. Put another way, "on-target" activity of a relatively weak inhibitor such as **22** would occur only at elevated concentrations that would cause significant inhibition of many "off-target" kinases as well. $S_{3\ \mu M}$ does not capture this because it uses a coarse filter at 3 μM, while the partition index uses a thermodynamics-based tapering that depends on the affinity observed for the kinase that is inhibited to the greatest extent.

CI-1033 (**16**), an EGFR/Erb inhibitor[18] that is most potent against EGFR,[6] also shows a dramatic difference in rank-ordering between the different indices. Compound **16** is ranked fifth in selectivity by P_{MAX} but 20th by $S_{3\ \mu M}$. The poor showing in the $S_{3\ \mu M}$ ranking is due to **16** binding 43 kinases below the 3 μM threshold.[6] The inhibitor is, however, greater than 500-fold selective for EGFR over 34 of the 43 kinases (79%) and is greater than 50-fold selective over all but two kinases. Interestingly, EKB-569 (**17**),[19] a compound similar in chemical structure to **16**, has a very similar $S_{3\ \mu M}$ but a significantly less selective P_{MAX} value compared to **16**. The different P_{MAX} values reflect the difference in selectivity profiles. Compound **17** is greater than 500-fold selective for only 22 of the 52 kinases (42%) within the $S_{3\ \mu M}$ threshold and is less than 10-fold selective over two of the kinases.

Table 2. Ambit Large Kinome Panel Data Set[6] Sorted by Partition Index[a]

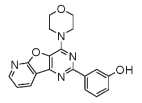
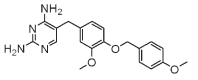
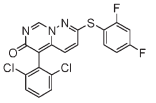
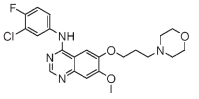
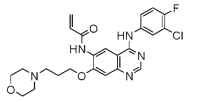
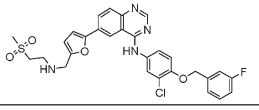
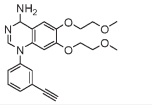
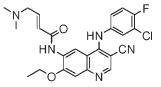
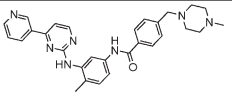
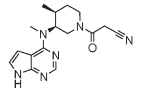
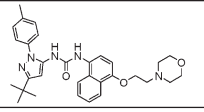
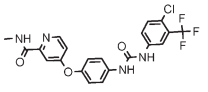
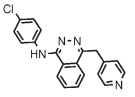
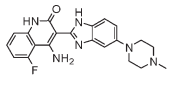
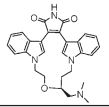
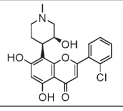
Compound	Structure	best K_d[i] IC50 (nM)	$G_{0.1}$[ii]	$S_{3\mu M}$	S_{10X}[iii]	P_{MAX}
13 (PI-103)		1.5	0.405	0.02	0.0000	0.97
14 (GW-2580)		1.6	0.538	0.01	0.0000	0.92
15 (VX-745)		2.8	0.510	0.03	0.0000	0.90
Gefitinib		1	0.647	0.07	0.0000	0.89
16 (CI-1033)		1	0.722	0.15	0.0000	0.88
Lapatinib		2.4	0.567	0.01	0.0000	0.70
Erlotinib		1	0.748	0.15	0.0035	0.67
17 (EKB-569)		1	0.783	0.18	0.0000	0.67
Imatinib		1	0.786	0.07	0.0209	0.65
CP-690550		2.2	0.655	0.03	0.0069	0.50
BIRB-796		1	0.806	0.16	0.0035	0.47
Sorafenib		1.5	0.815	0.18	0.0972	0.46
PTK-787		5.1	0.714	0.03	0.0242	0.44
CHIR-258/TKI-258		1	0.759	0.33	0.0069	0.44
LY-333531		2.5	0.761	0.16	0.0138	0.42
Flavopiridol		6.4	0.767	0.19	0.1696	0.42

Table 2. Continued

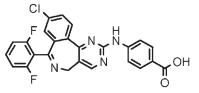
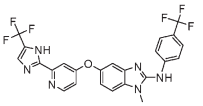
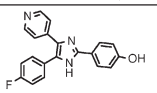
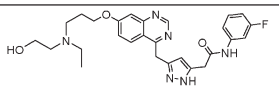
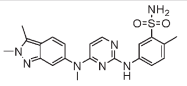
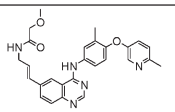
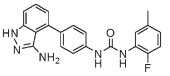
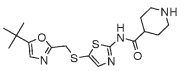
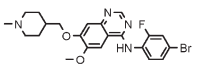
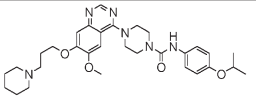
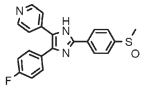
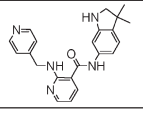
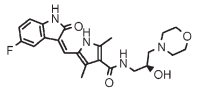
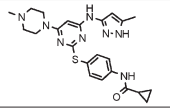
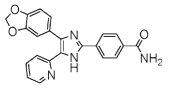
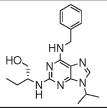
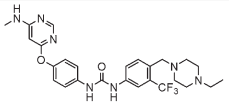
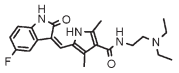
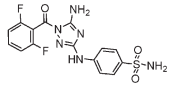
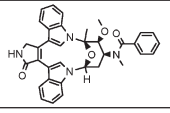
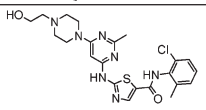
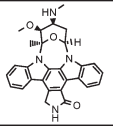
Compound	Structure	best K_d[i] IC50 (nM)	$G_{0.1}$[ii]	$S_{3\mu M}$	S_{10X}[iii]	P_{MAX}
MLN-8054		6.5	0.753	0.13	0.0104	0.36
CHIR-265/RAF-265		13	0.744	0.13	0.1215	0.36
SB-202190		9.8	0.757	0.09	0.0173	0.35
AZD-1152HQPA		4.4	0.777	0.10	0.0173	0.34
GW-786034		2	0.796	0.21	0.0244	0.33
18 (CP-724714)		42	0.488	0.02	0.0069	0.33
ABT-869		1	0.821	0.16	0.0105	0.32
BMS-387032/ SNS-032		7.1	0.796	0.13	0.0588	0.30
ZD-6474		4.6	0.786	0.27	0.2613	0.28
19 (MLN-518)		2.4	0.735	0.06	0.0104	0.26
SB-203580		12	0.773	0.10	0.0242	0.24
20 (AMG-706)		3.7	0.819	0.09	0.0383	0.22
SU-14813		1	0.700	0.51	0.0174	0.18
VX-680/MK-0457		4	0.761	0.38	0.0314	0.15
21 (SB-431542)		170	0.421	0.02	0.0104	0.14
22 (Roscovitine)		260	0.334	0.03	0.0313	0.12

Table 2. Continued

Compound	Structure	best K_d[i] IC50 (nM)	$G_{0.1}$[ii]	$S_{3\mu M}$	S_{10X}[iii]	P_{MAX}
AST-487		1	0.729	0.45	0.0417	0.11
Sunitinib		1	0.666	0.57	0.0174	0.11
JNJ-7706621		21	0.736	0.37	0.0833	0.10
PKC-412		9.3	0.707	0.47	0.0764	0.10
23 (Dasatinib)		1	0.791	0.28	0.1042	0.03
Staurosporine		1	0.353	0.87	0.5017	0.02

[a] [i]Because measured K_d values less than 1 nM tend to have greater discrepancies with reported IC_{50} values,[6] values less than 1 nM are set to 1 nM. Undefined IC_{50} values are set to 15 μM. [ii]To calculate Gini coefficients using the published spreadsheet,[9] the simplified Hill equation was used to convert K_d values to POC values at 0.1 μM inhibitor concentration. [iii]From Supporting Information Table 4 of ref 6.

The partition index is able to distinguish important differences in selectivity profiles that a simpler score does not.

A final compound that is ranked very differently by different indices is MLN-518 (**19**), a FLT3/KIT/PDGFR inhibitor.[20] Compound **19** appears only moderately selective based on $P_{MAX} = 0.26$, even though it binds to very few kinases in the panel, as indicated by $S_{3\ \mu M} = 0.06$. The low ranking based on P_{MAX} is due to **19** being a multikinase inhibitor that binds to four kinases in the Ambit panel roughly equipotently. In the partition index scheme, each of the four kinases binds ~25% of the inhibitor pool. P_{MAX} only accounts for one of these kinases. If all four kinase activities are desirable, then a more relevant partition index would be $P_{FLT3+KIT+PDGFR} = 0.85$, which is simply the sum of the P values for FLT3, KIT, PDGFR-A, and PDGFR-B. Summations of partition index values can provide relevant definitions of selectivity for other multikinase inhibitors in the data set, including **21**,[15] **18**,[16] dasatinib (**23**),[21] and motesanib (**20**, AMG-706).[22] The partition index provides a simple, thermodynamically relevant approach for analysis of desired multikinase inhibition that is difficult to achieve with other indices.

These examples illustrate the advantages of the partition index for kinome panel analysis. We point out, however, that the partition index provides a view of selectivity that is complementary to other indices such as the S score and Gini coefficient. An inhibitor with potent activity of 1 nM against three out of 290 kinases, only one of which is the desired kinase, will score $P_{MAX} = 0.33$, suggesting moderate selectivity for the desired kinase. This is because only a third of the inhibitor pool is binding the desired kinase target. This view is perhaps appropriate for lead optimization. However, one might also call the inhibitor highly selective because it misses 287 of 290 kinases. Selectivity indices such as the S score and Gini coefficient effectively capture this second view and may be more appropriate for lead selection and "high-level"

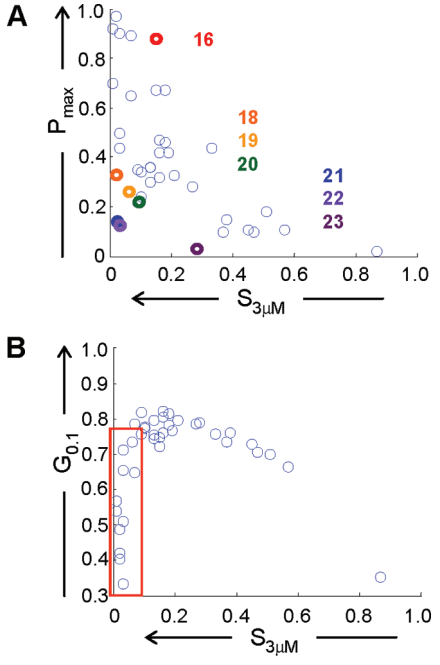


Figure 3. Scatterplot of data in Table 2. (A) $S_{3\ \mu M}$ versus P_{MAX}. (B) $S_{3\ \mu M}$ versus $G_{0.1}$. Arrows indicate direction of increasing selectivity for indices, and the colored circles in (A) indicate compounds discussed in the main text.

kinome data analysis with their low sensitivity to inhibition values and multikinase inhibitors. Relative S scores such as S_{10X} can additionally provide a target-relative view similar to the partition index but with low sensitivity due to use of a hard cutoff. This lessened sensitivity can be seen in Table 2, where many inhibitors have identical S_{10X} values. Interestingly, the Gini coefficient correlates well with the S score except for the very selective compounds highlighted by a red box in Figure 3B. This is because in the Gini procedure, small differences in weak inhibition values, which are essentially noise, are amplified by normalization of cumulative inhibition values.[23]

Kinome Panel Analysis with POC Data. In our final scenario, we turn to a common situation in which a selection of compounds is profiled against a medium-sized, diverse panel of kinases for POC inhibition. Measuring POC inhibition instead of K_d or IC_{50} is popular for pharmacological profiling panels because it requires fewer experiments. To illustrate usage of the partition index in this situation, we use the data shown in Table 3 resulting from an Amgen internal project. Nine undisclosed compounds were screened against 48 kinases[24] at a fixed inhibitor concentration of 1 μM. POC values in Table 3 are colored from red for POC < 15, to yellow for 15 ≤ POC < 35, to green for 35 ≤ POC < 50.

Calculation of the partition index requires converting POC's to estimated IC_{50}'s, and this can be done using the simplified Hill equation:

$$IC_{50} = [C]\left(\frac{100}{100 - POC} - 1\right)$$

where C is the screening concentration. POC values are pruned before conversion by setting all POC's that are greater than 100 to 99.9, and those that are less than zero to 0.1; the offset values ensure mathematically defined IC_{50}'s. Conversion of POC inhibition to IC_{50} values is most accurate for IC_{50}'s within 20-fold of the screening concentration or roughly 5% ≤ POC ≤ 95% (unpublished observations). In particular, IC_{50}'s estimated from POC's are less accurate for potent compounds, which have POC < 5%.

Nevertheless, for our example where the screen was run at a 1 μM compound concentration, the partition index values for compounds **24**−**32** are generally consistent with a visual inspection of the counting scheme shown in the gray-shaded rows of Table 3, with the exception of **28** versus **29**. While the partition index ranks **28** as slightly more selective than **29**, a simple counting approach (gray rows in Table 3) suggests that **29** is slightly more selective than **28**. The reversed relative ranking is largely due to the insignificant difference in DYRK1α POC values (**28**, 2% versus **29**, 1%). While the actual difference in the resulting partition index values is small, the comparison does highlight the sensitivity of the partition index to one or two percent differences in very low POC values. One way to remove experimental uncertainty for targets with POC < 5% is to set those POC values to a fixed value such as 4%. A better way is obtaining more accurate measurements by screening at a lower inhibitor concentration such as 0.1 μM when leads have IC_{50} < 50 nM on the desired target.

The S_{35} score is well-suited for analysis of this type of POC data, and its use of a simple cutoff filter prevents oversensitivity to insignificant differences in POC values < 5%. In comparison, the Gini coefficient performs reasonably for the majority of inhibitors in Table 3, but has difficulty ranking compounds **24** and **28**, largely due to the scaling issues mentioned previously. All three indices perform reasonably for the kinome POC panel shown here. The partition index is particularly helpful when inhibitors do not have single-digit nM inhibition on the desired target, or in the case of designed multikinase inhibitors.

Table 3. POC Kinome Selectivity Panel Example[a]

	24	25	26	27	28	29	30	31	32
Kinase X	0	0	0	0	0	0	0	0	0
ABL	98	77	57	80	62	77	85	27	37
AKT1	96	146	142	152	95	96	103	101	106
AKT2	86	122	134	135	96	91	94	89	95
AMPK	73	87	25	59	29	53	38	18	15
AurA	96	119	62	120	74	94	134	71	72
BTK	95	89	90	90	89	121	100	100	80
CAMK2	98	84	113	86	24	78	53	62	12
CAMK4	98	91	84	83	75	91	80	88	68
CDK2	94	93	44	86	35	37	57	13	49
CHK1	123	121	130	125	113	101	84	94	85
CHK2	105	87	132	133	84	106	104	91	80
CK1δ	97	41	54	15	12	21	14	19	19
c-RAF	95	101	97	97	96	82	110	115	93
c-TAK1	84	115	74	100	71	82	79	23	65
DYRK1α	57	14	56	5	2	1	1	34	3
Erk1	88	122	119	142	89	90	94	114	98
Erk2	92	130	149	136	96	94	105	87	90
FGFR1	97	86	18	81	64	73	82	4	40
FLT3	87	18	4	13	2	10	9	2	0
FYN	99	76	73	86	64	72	86	58	18
GSK3β	85	92	97	96	86	85	139	98	94
HGK	92	113	119	117	86	97	101	99	82
IGF1R	94	74	96	91	91	94	99	109	108
INSR	94	93	88	95	92	94	107	90	91
IRAK4	106	94	88	91	95	113	87	89	82
KDR	103	79	23	74	67	78	96	9	27
LCK	97	75	50	64	40	76	65	36	15
LYN	91	97	79	115	77	87	122	88	40
MAPK2	90	98	101	103	94	97	99	102	98
MARK1	82	95	86	88	75	86	88	54	79
MET	105	108	99	114	98	108	121	93	94
MSK1	95	135	148	128	84	87	88	88	76
MST2	99	98	85	100	80	50	82	77	74
p38α	84	92	96	98	79	98	89	97	91
p70S6K	96	124	118	133	89	112	79	83	89
PAK2	102	106	105	102	100	91	99	96	96
PIM2	96	46	98	26	13	8	95	53	33
PKA	87	140	120	133	89	90	103	77	94
PKCβ2	90	108	109	114	91	95	99	95	94
PKCζ	96	94	102	119	96	96	97	97	96
PKD2	71	70	123	39	17	73	12	28	6
PKGα	98	137	119	149	73	83	86	77	80
PRAK	95	98	100	103	95	96	81	94	98
ROCK2	102	117	138	142	84	84	88	87	82
RSK1	82	126	85	143	60	68	75	67	58
SGK1	93	68	84	41	18	39	29	67	15
SRC	89	97	72	89	74	85	109	88	40
SYK	94	92	102	94	93	97	105	104	87
0≤POC<15	1	2	2	4	5	4	5	5	5
15≤POC<30	0	1	3	1	4	1	1	5	6
30≤POC<50	0	2	1	2	2	3	1	2	6
Gini	0.59	0.71	0.70	0.74	0.52	0.59	0.67	0.57	0.50
S_{35}	0.00	0.04	0.10	0.08	0.16	0.08	0.10	0.20	0.22
P(KinaseX)	0.99	0.96	0.95	0.86	0.84	0.79	0.69	0.66	0.22

[a] Kinases were screened at 1 μM inhibitor concentration, with the exception of KinaseX, with which POC values shown here are converted from IC_{50} values using the simplified Hill equation.

Conclusions

We have described a new thermodynamics-based partition index that is useful for analyzing the selectivity of both small, focused kinase panels as well as larger kinome selectivity panels encountered in medicinal chemistry projects. The partition index is a relative measure calculated for a particular kinase or set of kinases and can be viewed as an extension of the fold-selectivity measure. It is more sensitive than indices such as the Gini coefficient and Ambit S score in its ability to capture differences in inhibitor potencies against individual kinases. These two features, target relativity and sensitivity, provide significant advantages for analysis of small panel counter-screen selectivity data that may be generated on a weekly basis. For larger "kinome" panel data typically generated on a less frequent basis, S_X scores are simple to calculate and perform well, and the loss of sensitivity from use of an arbitrary cutoff value is less of an issue when large numbers of kinases are involved. The partition index provides a graded cutoff based on thermodynamics theory, and this has advantages for providing a target-oriented view of the selectivity data. Because the partition index has a clear thermodynamics interpretation, the index can be easily extended to more complex situations such as analysis of multikinase inhibitor profiles.

Methods

The partition index is based on the thermodynamics partition function, which describes the partitioning of states based on the energy of each state. Using the Boltzmann distribution, we can define the relative partitioning of compound to a kinase, i, relative to reference kinase as:

$$q_i = e^{-\Delta\Delta G_i/\mathrm{RT}} = e^{-(\Delta G_i - \Delta G_{\mathrm{ref}})/\mathrm{RT}} = \frac{e^{-\Delta G_i/\mathrm{RT}}}{e^{-\Delta G_{\mathrm{ref}}/\mathrm{RT}}} = \frac{K_{\mathrm{d}}(\mathrm{ref})}{K_{\mathrm{d}}(i)}$$

where ΔG_i and $K_{\mathrm{d}}(i)$ are the binding free energy and binding affinity, respectively, of compound i. We then define a partition function, Q, which represents the number of occupied kinase states relative to a reference K_{d}:

$$Q = \sum_{i=1}^{n} w_i q_i = \sum_{i=1}^{n} \left(w_i \times \frac{K_{\mathrm{d}}(\mathrm{ref})}{K_{\mathrm{d}}(i)} \right)$$

where the sum is over the n kinases in the kinase panel and IC_{50}'s can be used in place of K_{d}'s. The reference represented by K_{d}(ref) or IC_{50}(ref) can be the most potently inhibited kinase or the kinase of interest.

The "degeneracy" or "weighting" term, w_i, is useful for emphasizing or de-emphasizing certain kinases based on quantitative measures such as measured cell-shifts, or qualitative measures such as the biological importance of not inhibiting a particular kinase. For example, inhibiting some kinases leads to toxicity and these could have a higher weighting. For the current work, a uniform weighting ($w_i = 1$) was used. Finally, the partitioning of inhibitor for a particular kinase is defined by the fraction of time the inhibitor spends bound to a particular kinase ("kinaseA"):

$$P_{\mathrm{kinaseA}} = \frac{w_a \times q_{\mathrm{KkinaseA}}}{Q} = \frac{w_a \cdot \dfrac{K_{\mathrm{d}}(\mathrm{ref})}{K_{\mathrm{d}}(\mathrm{kinaseA})}}{\displaystyle\sum_i \left(w_i \cdot \frac{K_{\mathrm{d}}(\mathrm{ref})}{K_{\mathrm{d}}(i)} \right)} = \frac{\dfrac{w_a}{K_{\mathrm{d}}(\mathrm{kinaseA})}}{\displaystyle\sum_i \frac{w_i}{K_{\mathrm{d}}(i)}}$$

In the main text, we used a simplified derivation without weighting factors, where

$$z_i = \frac{q_i}{K_{\mathrm{d}}(\mathrm{ref})} \text{ and } z_{\mathrm{total}} = \frac{Q}{K_{\mathrm{d}}(\mathrm{ref})}$$

Note Added after ASAP Publication. This paper was published on May 11, 2010 with an error in Figure 2. The revised version was published on May 17, 2010.

Supporting Information Available: Calculated POC values for Tie-2 data set. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The protein kinase complement of the human genome. *Science* **2002**, *298*, 1912–1934.

(2) Noble, M. E.; Endicott, J. A.; Johnson, L. N. Protein kinase inhibitors: insights into drug design from structure. *Science* **2004**, *303*, 1800–1805.

(3) Fabian, M. A.; Biggs, W. H., III; Treiber, D. K.; Atteridge, C. E.; Azimioara, M. D.; Benedetti, M. G.; Carter, T. A.; Ciceri, P.; Edeen, P. T.; Floyd, M.; Ford, J. M.; Galvin, M.; Gerlach, J. L.; Grotzfeld, R. M.; Herrgard, S.; Insko, D. E.; Insko, M. A.; Lai, A. G.; Lélias, J. M.; Mehta, S. A.; Milanov, Z. V.; Velasco, A. M.; Wodicka, L. M.; Patel, H. K.; Zarrinkar, P. P.; Lockhart, D. J. A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat. Biotechnol.* **2005**, *23*, 329–336.

(4) Keri, G.; Orfi, L.; Eros, D.; Hegymegi-Barakonyi, B.; Szántai-Kis, C.; Horváth, Z.; Wáczek, F.; Marosfalvi, J.; Szabadkai, I.; Pató, J.; Greff, Z.; Hafenbradl, D.; Daub, H.; Müller, G; Klebl, B.; Ullrich, A. Signal transduction therapy with rationally designed kinase inhibitors. *Curr. Signal Transduction Ther.* **2006**, *1*, 67–95.

(5) Smyth, L. A.; Collins, I. Measuring and interpreting the selectivity of protein kinase inhibitors. *J. Chem. Biol.* **2009**, *2*, 131–151.

(6) Karaman, M. W.; Herrgard, S.; Treiber, D. K.; Gallant, P.; Atteridge, C. E.; Campbell, B. T.; Chan, K. W.; Ciceri, P.; Davis, M. I.; Edeen, P. T.; Faraoni, R.; Floyd, M.; Hunt, J. P.; Lockhart, D. J.; Milanov, Z. V.; Morrison, M. J.; Pallares, G.; Patel, H. K.; Pritchard, S.; Wodicka, L. M.; Zarrinkar, P. P. A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **2008**, *26*, 127–132.

(7) Fedorov, O.; Marsden, B.; Pogacic, V.; Rellos, P.; Müller, S.; Bullock, A. N.; Schwaller, J.; Sundström, M.; Knapp, S. A systematic interaction map of validated kinase inhibitors with Ser/Thr kinases. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 20523–20528.

(8) Cheng, A. C. Predicting selectivity and druggability in drug discovery. *Annu. Rep. Comp. Chem.* **2008**, *4*, 23–37.

(9) Graczyk, P. P. Gini coefficient: a new way to express selectivity of kinase inhibitors against a family of kinases. *J. Med. Chem.* **2007**, *50*, 5773–5779.

(10) Dill, K. A. Bromberg, S. *Molecular Driving forces: Statistical Thermodynamics in Chemistry and Biology*, 1st ed.; Garland Science: New York, 2002.

(11) Cee, V. J.; Cheng, A. C.; Romero, K.; Bellon, S.; Mohr, C.; Whittington, D. A.; Bak, A.; Bready, J.; Caenepeel, S.; Coxon, A.; Deak, H. L.; Fretland, J.; Gu, Y.; Hodous, B. L.; Huang, X.; Kim, J. L.; Lin, J.; Long, A. M.; Nguyen, H.; Olivieri, P. R.; Patel, V. F.; Wang, L.; Zhou, Y.; Hughes, P.; Geuns-Meyer, S. Pyridylpyrimidine benzimidazole derivatives as potent, selective, and orally bioavailable inhibitors of Tie-2 kinase. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 424–427.

(12) Hayakawa, M.; Kaizawa, H.; Moritomo, H.; Koizumi, T.; Ohishi, T.; Yamano, M.; Okada, M.; Ohta, M.; Tsukamoto, S.; Raynaud, F. I.; Workman, P.; Waterfield, M. D.; Parker, P. Synthesis and biological evaluation of pyrido[3′,2′:4,5]furo[3,2-*d*]pyrimidine derivatives as novel PI3 kinase p110alpha inhibitors. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 2438–2442.

(13) Conway, J. G.; McDonald, B.; Parham, J.; Keith, B.; Rusnak, D. W.; Shaw, E.; Jansen, M.; Lin, P.; Payne, A.; Crosby, R. M.; Johnson, J. H.; Frick, L.; Lin, M. H.; Depee, S.; Tadepalli, S.; Votta, B.; James, I.; Fuller, K.; Chambers, T. J.; Kull, F. C.; Chamberlain, S. D.; Hutchins, J. T. Inhibition of colony-stimulating-factor-1 signaling in vivo with the orally bioavailable cFMS kinase inhibitor GW2580. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 16078–16083.

(14) Fitzgerald, C. E.; Patel, S. B.; Becker, J. W.; Cameron, P. M.; Zaller, D.; Pikounis, V. B.; O'Keefe, S. J.; Scapin, G. Structural basis for p38alpha MAP kinase quinazolinone and pyridol-pyrimidine inhibitor specificity. *Nat. Struct. Biol.* **2003**, *10*, 764–769.

(15) Inman, G. J.; Nicolás, F. J.; Callahan, J. F.; Harling, J. D.; Gaster, L. M.; Reith, A. D.; Laping, N. J.; Hill, C. S. SB-431542 is a potent and specific inhibitor of transforming growth factor-beta superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. *Mol. Pharmacol.* **2002**, *62*, 65–74.

(16) Jani, J.; Barbacci, E.; Bhattacharya, S.; Boos, C.; Campbell, M.; Clark, T.; Coleman, K.; Connell, R.; Cosker, T. Discovery and development of CP-724714, a selective HER2 receptor tyrosine kinase inhibitor. *J. Clin. Oncol.* **2004**, *22* (Suppl.), 3122.

(17) Meijer, L.; Borgne, A.; Mulner, O.; Chong, J. P.; Blow, J. J.; Inagaki, N.; Inagaki, M.; Delcros, J. G.; Moulinoux, J. P. Biochemical and cellular effects of roscovitine, a potent and selective inhibitor of the cyclin-dependent kinases cdc2, cdk2 and cdk5. *Eur. J. Biochem.* **1997**, *243*, 527–536.

(18) Smaill, J. B.; Rewcastle, G. W.; Loo, J. A.; Greis, K. D.; Chan, O. H.; Reyner, E. L.; Lipka, E.; Showalter, H. D.; Vincent, P. W.; Elliott, W. L.; Denny, W. A. Tyrosine kinase inhibitors. 17. Irreversible inhibitors of the epidermal growth factor receptor: 4-(phenylamino)quinazoline- and 4-(phenylamino)pyrido[3,2-*d*]-pyrimidine-6-acrylamides bearing additional solubilizing functions. *J. Med. Chem.* **2000**, *43*, 1380–1397.

(19) Torrance, C. J.; Jackson, P. E.; Montgomery, E.; Kinzler, K. W.; Vogelstein, B.; Wissner, A.; Nunes, M.; Frost, P.; Discafani, C. M. Combinatorial chemoprevention of intestinal neoplasia. *Nature Med.* **2000**, *6*, 1024–1028.

(20) Pandey, A.; Volkots, D. L.; Seroogy, J. M.; Rose, J. W.; Yu, J. C.; Lambing, J. L.; Hutchaleelaha, A.; Hollenbach, S. J.; Abe, K.; Giese, N. A.; Scarborough, R. M. Identification of orally active, potent, and selective 4-piperazinylquinazolines as antagonists of the platelet-derived growth factor receptor tyrosine kinase family. *J. Med. Chem.* **2002**, *45*, 3772–3793.

(21) Lombardo, L. J.; Lee, F. Y.; Chen, P.; Norris, D.; Barrish, J. C.; Behnia, K.; Castaneda, S.; Cornelius, L. A.; Das, J.; Doweyko, A. M.; Fairchild, C.; Hunt, J. T.; Inigo, I.; Johnston, K.; Kamath, A.; Kan, D.; Klei, H.; Marathe, P.; Pang, S.; Peterson, R.; Pitt, S.; Schieven, G. L.; Schmidt, R. J.; Tokarski, J.; Wen, M. L.; Wityak, J.; Borzilleri, R. M. Discovery of *N*-(2-chloro-6-methyl- phenyl)-2-(6-(4-(2-hydroxyethyl)- piperazin-1-yl)-2-methylpyrimidin-4-ylamino)thiazole-5-carboxamide (BMS-354825), a dual Src/Abl kinase inhibitor with potent antitumor activity in preclinical assays. *J. Med. Chem.* **2004**, *47*, 6658–6661.

(22) Polverino, A.; Coxon, A.; Starnes, C.; Diaz, Z.; DeMelfi, T.; Wang, L.; Bready, J.; Estrada, J.; Cattley, R.; Kaufman, S.; Chen, D.; Gan, Y.; Kumar, G.; Meyer, J.; Neervannan, S.; Alva, G.; Talvenheimo, J.; Montestruque, S.; Tasker, A.; Patel, V.; Radinsky, R.; Kendall, R. AMG 706, an oral, multikinase inhibitor that selectively targets vascular endothelial growth factor, platelet-derived growth factor, and kit receptors, potently inhibits angiogenesis and induces regression in tumor xenografts. *Cancer Res.* **2006**, *66*, 8715–8721.

(23) This is due to normalization of the cumulative inhibition values in the Gini procedure—see ref 9 for details. A very selective compound will have a low cumulative inhibition value relative to a less selective compound. The cumulative inhibition values are normalized to the same value of one. The resulting effect is that with very selective compounds, low POC values end up being scaled up to achieve normalization, and small differences (e.g., 99 POC versus 97 POC) are magnified.

(24) Caliper ProfilerPro Assay Kits 1 and 2 from Caliper Life Sciences, Hopkinton, MA. Assays performed in 2008.