Gini Coefficient: A New Way To Express Selectivity of Kinase Inhibitors against a Family of Kinases†

Piotr P. Graczyk*

Eisai London Research Laboratories Ltd., Medicinal Chemistry Department, University College London, Bernard Katz Building, Gower Street, London WC1E 6BT, United Kingdom

Received May 14, 2007

A novel application of the Gini coefficient for expressing selectivity of kinase inhibitors against a panel of kinases is proposed. This has been illustrated using single-point inhibition data for 40 commercially available kinase inhibitors screened against 85 kinases. Nonselective inhibitors are characterized by Gini values close to zero (Staurosporine, Gini 0.150). Highly selective compounds exhibit Gini values close to 1 (PD184352 Gini 0.905). The relative selectivity of inhibitors does not depend on the ATP concentration.

Introduction

Recent advances in genomics have allowed the identification of about 518 kinases[1] with many of them interesting from a medicinal chemistry viewpoint[2] as targets for cancer,[3] inflammatory,[4,5] neurodegenerative,[6] and autoimmune[7,8] disorders. Since kinases are involved in complex molecular signaling mechanisms, insufficient selectivity of an inhibitor may result in undesirable side effects.[9] The desire to limit such effects has stimulated various approaches to measure, understand, and therefore control compound selectivity.[10]

Discussion

The term "selectivity" of an inhibitor is usually understood as the ratio of K_i values for different kinases. Thus, a set of K_i values for an inhibitor is commonly used to evaluate its selectivity profile.[11] More often than not IC_{50} values are used for this purpose.[12] These approaches may be quite costly for compound profiling since several data points are needed for each K_i or IC_{50} value to be established. For a large number of kinases, the information hidden within hundreds of K_i values is difficult to interpret, particularly if selectivity of several compounds needs to be compared. The interaction maps based on cluster analysis of specificity profiles[13] allow visualization of multiple kinase−inhibitor interactions. This approach is focused on the identification of individual unanticipated interactions with kinases present within the screening panel. However, it requires a nonstandard assay to determine the affinity of the compounds and is limited to ATP-competitive inhibitors. Furthermore, it does not allow the ranking of compounds in terms of selectivity and does not offer any insight into a compound's likely selectivity against the rest of the kinome.

If one considers the terms "selectivity" and "inequality" to have analogous meaning, the selectivity of a compound against a larger kinase population can be evaluated in statistical terms by the Gini coefficient. This is frequently used by economists to measure income inequality,[14] although Gini has also been applied for other purposes.[15] For the selectivity evaluation, instead of income, one could use the magnitude of inhibition measured for each kinase at a single point at a given ATP concentration.

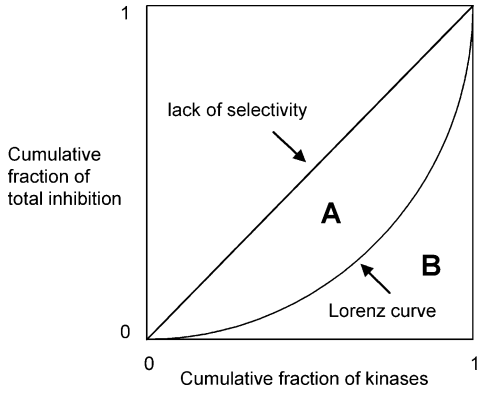

Figure 1. Gini coefficient as the ratio of areas $A/(A + B)$.

The total inhibition ("total wealth") can be calculated as the sum of magnitudes of inhibition for all kinases. Then, the kinases need to be sorted in order of increasing inhibition. The cumulative fraction of total inhibition is plotted against the cumulative fraction of kinases as shown in Figure 1. If all kinases are inhibited to an equal extent, the cumulative fraction of total inhibition increases linearly with the cumulative fraction of kinases. This diagonal line corresponds to complete lack of selectivity. If the compound strongly inhibits a small fraction of kinases, the cumulative fraction of total inhibition will initially increase slowly following the Lorenz curve[16] and then steeply increase to 1 for the last few potently inhibited kinases.

If the area between the diagonal line and Lorenz curve is A, and the area under the Lorenz curve is B, the Gini coefficient is $G = A/(A + B)$. Taking into account that $A + B = 0.5$, the Gini coefficient equals $G = 1 - 2 \times B$.

An example of Gini calculations for AG1024 at 10 μM concentration in the presence of 10 μM ATP is presented below using the inhibition data provided by Upstate/Millipore[17] for 20 kinases shown in Table 1.

First, the residual activities are sorted in decreasing order and normalized by changing to 100 all values greater than 100 and all negative values to zero (Table 2).

The activity values are converted into a percentage of inhibition and assigned a sequential position number. Using the sum of inhibition percentages of 513, the fractions of total inhibition are then calculated and converted into a cumulative inhibition fraction. The area under the Lorenz curve (B, Figure

† This work is dedicated to Professor Yoshito Kishi on the occasion of his 70th birthday.

* E-mail: piotr_graczyk@eisai.net. Fax: +442074131121. Phone: +442074131147.

10.1021/jm070562u CCC: $37.00
Published on Web 10/19/2007

Table 1. Inhibition Data for AG1024 at 10 μM Concentration in the Presence of 10 μM ATP

kinase	residual activity (%)
Abl(m)	1
AMPK(r)	84
Arg(m)	6
Aurora-A(h)	104
Axl(h)	58
Blk(m)	4
Bmx(h)	84
CaMKII(r)	89
CaMKIV(h)	81
CDK1/cyclinB(h)	67
CDK2/cyclinA(h)	101
CDK2/cyclinE(h)	95
CDK3/cyclinE(h)	87
CDK5/p35(h)	78
CDK6/cyclinD3(h)	92
CDK7/cyclinH/MAT1(h)	93
CHK1(h)	97
CHK2(h)	88
CK1(y)	87
CK2(h)	96

1) is established using the trapezium method. For instance, the area of an individual trapezium ending at position 17 is

$$\left(\frac{0.4366 - 0.3548}{2} + 0.3548\right) \times (0.85 - 0.80) = 0.0198$$

With the total area of 0.211, the Gini coefficient is equal in this case $G = 1 - 2 \times 0.211 = 0.578$.

Such calculations can easily be automated by applying a properly programmed Microsoft Excel spreadsheet (available in the Supporting Information). In a similar manner, the Gini coefficients were computed for 40 commercially available kinase inhibitors, which were profiled by Upstate/Millipore against a panel of 85 kinases at ATP concentrations of 10 and 100 μM. The relevant Gini values are shown in Table 3. For reference purposes the relevant hit rates, equal to the number of kinases inhibited by >50%, are also shown. Example Lorenz curves for PD184352, AG1024, and Staurosporine are presented in Figure 2.

It is interesting to see that the Gini coefficients cover a broad range from 0.093 (1 μM staurosporine at 10 μM ATP; denoted by $G_{1/10}$) up to 0.905 (10 μM PD184352 at 100 μM ATP; $G_{10/100}$). Staurosporine is considered to be a highly nonselective kinase inhibitor known for its toxicity.[18] PD184352 is a very selective inhibitor of MAPKK1,[19] selectivity of which is due to its noncompetitive mode of binding.[20] AG1024 represents an inhibitor with moderate selectivity ($G_{10/10}$ 0.568).

The Lorenz curves corresponding to the clusters of Gini coefficient between 0.417 and 0.460 and 0.553−0.595 at 10 μM ATP form two distinct groups, which are also characterized by similar shapes as shown in Figure 3.

The hit rate depends on the ATP concentration and varies in the range from 0 (e.g., Genistein, PP3) to 78 (Staurosporine). It is moderately correlated with the Gini coefficient (Figure 4; $r^2 = 0.657$). Thus, the magnitude of Gini coefficient seems to reflect the overall selectivity of these compounds.

Gini coefficients at 10 μM and 100 μM ATP (G_{10} and G_{100}, respectively) for the same compound differ slightly. However, there is a good linear correlation between them (Figure 5; $r^2 = 0.884$). A practical consequence of this finding is that in order to rank the compounds in terms of their overall selectivity it might be enough to profile the compounds at a single ATP concentration. Regardless of this concentration the most selective compounds include PD184352, PP1 analogue, Roscovitine, PP2, Wortmannin, Hispidin, Lavendustin A, PP3, and LY294002 with slight variation in their order depending on ATP concentration. Thus, the Gini coefficient removes the subjective element of judgment and allows ranking of inhibitors in terms of selectivity, which has not been possible to date. This approach has been used at our laboratories to identify structural motifs with higher potential for creating selective inhibitors.[21]

One may notice that the Gini value of 0.578 calculated as an example for AG1024 based on the data for 20 kinases (Table 1) does not change markedly when the set of kinases is expanded to 85 (Gini 0.568, Table 3). The 20 kinases in Table 1 represent the first 20 out of the whole set of 85. To check whether the Gini coefficient depends on the choice of a kinase subset, the subsets of 10, 20, 40, 60, and 80 kinases were chosen randomly, each 13 times, to calculate the relevant Gini coefficients (see Supporting Information). The results are presented in Table 4.

It is clear that the mean values of Gini coefficients for AG1024 do not depend markedly on the subset size as they

Table 2. Calculation of Area under the Lorenz Curve for AG1024 at 10 μM Concentration in the Presence of 10 μM ATP Using Residual Activities from Table 1

activities sorted	activities normalized	position	cumulative sample fraction	inhibition (%)	fraction of total inhibition	cumulative inhibition fraction	area
104	100	1	0.05	0	0.0000	0.0000	0.0000
101	100	2	0.10	0	0.0000	0.0000	0.0000
97	97	3	0.15	3	0.0058	0.0058	0.0001
96	96	4	0.20	4	0.0078	0.0136	0.0005
95	95	5	0.25	5	0.0097	0.0234	0.0009
93	93	6	0.30	7	0.0136	0.0370	0.0015
92	92	7	0.35	8	0.0156	0.0526	0.0022
89	89	8	0.40	11	0.0214	0.0741	0.0032
88	88	9	0.45	12	0.0234	0.0975	0.0043
87	87	10	0.50	13	0.0253	0.1228	0.0055
87	87	11	0.55	13	0.0253	0.1481	0.0068
84	84	12	0.60	16	0.0312	0.1793	0.0082
84	84	13	0.65	16	0.0312	0.2105	0.0097
81	81	14	0.70	19	0.0370	0.2476	0.0115
78	78	15	0.75	22	0.0429	0.2904	0.0135
67	67	16	0.80	33	0.0643	0.3548	0.0161
58	58	17	0.85	42	0.0819	0.4366	0.0198
6	6	18	0.90	94	0.1832	0.6199	0.0264
4	4	19	0.95	96	0.1871	0.8070	0.0357
1	1	20	1.00	99	0.1930	1.0000	0.0452
			total	513		total area	0.211

Table 3. Gini Coefficients and Hit Rate (>50% Inhibition) for 40 Inhibitors against 85 Kinases (see Supporting Information) at 10 μM (G_{10}) and 100 μM (G_{100}) ATP

inhibitor	concn (μM)	G_{10}	hit rate	G_{100}	hit rate
AG1024	10	0.568	11	0.637	7
AG1296	10	0.498	4	0.541	2
AG1478	10	0.500	25	0.649	16
AG18	10	0.680	1	0.652	1
AG183	10	0.460	14	0.518	7
AG538	10	0.417	17	0.500	7
Alsterpaullone	1	0.633	22	0.704	14
Calphostin C	10	0.606	2	0.651	1
Cdk2/5 inh.	10	0.555	7	0.678	3
Curcumin	50	0.417	38	0.538	24
EGCG	10	0.495	21	0.621	23
Genistein	10	0.582	0	0.623	0
H89	10	0.442	34	0.518	27
HA1077	10	0.650	19	0.677	5
Herbimycin A	10	0.534	0	0.590	0
Hispidin	10	0.790	2	0.722	0
Indirubin	10	0.291	52	0.407	45
JNK inh II	10	0.445	35	0.533	17
K252c	10	0.236	58	0.384	41
KT5720	1	0.423	33	0.537	21
Lavendustin A	1	0.726	0	0.779	0
Lavendustin B	1	0.515	0	0.548	0
LY294002	50	0.619	6	0.733	2
Olomoucin	50	0.556	13	0.660	7
PD153035	0.001	0.616	0	0.623	0
PD184352	10	0.802	1	0.905	1
PP1 analogue	1	0.758	15	0.837	11
PP2	1	0.722	14	0.799	11
PP3	1	0.634	0	0.750	0
Ro31-8220	1	0.432	39	0.527	32
Roscovitine	10	0.744	10	0.805	6
Rottlerin	10	0.595	3	0.627	2
SB202190	10	0.553	29	0.700	14
SB203580	10	0.621	17	0.731	10
ST638	10	0.568	14	0.624	5
Staurosporine	1	0.093	78	0.150	72
SU6656	1	0.607	19	0.679	12
Wortmannin	1	0.775	0	0.739	0
Y27632	10	0.628	13	0.703	8
ZM336372	10	0.635	11	0.609	9

vary in the range from 0.527 to 0.571. However, the individual Gini coefficients can vary in a much wider range (e.g., 0.385−0.658), which is reflected in the standard deviation values increasing from 0.007 to 0.091 when the subset size is reduced from 80 to 10 kinases (94% to 12% of the panel). Analogous observations were also made for the nonselective 1 μM Staurosporine at 100 μM ATP and the very selective 10 μM PD184352 at 10 μM ATP (Supporting Information). Such variability of individual Gini coefficients would reduce their value as measures of selectivity.

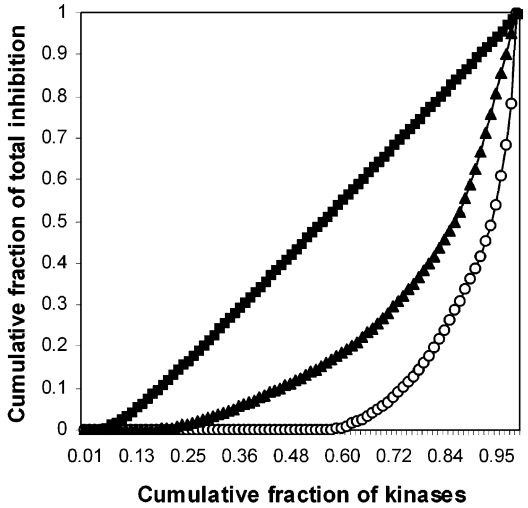

Figure 2. Lorenz curves for 10 μM PD184352 (○), 10 μM AG1024 (▲), and 1 μM Staurosporine (■) at 10 μM ATP (based on data in Supporting Information).

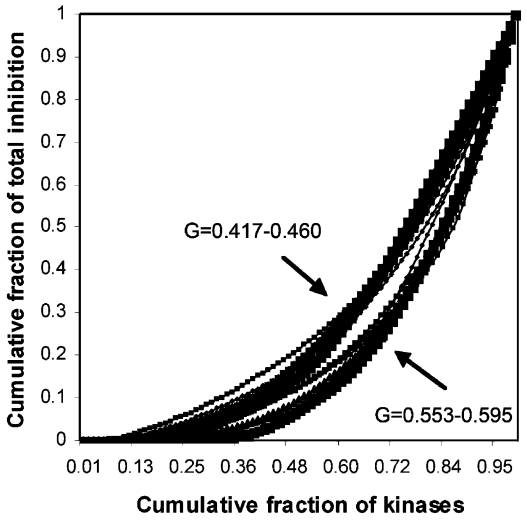

Figure 3. Lorenz curves for clusters of Gini G_{10} = 0.417−0.460 (AG538, Curcumin, KT5720, Ro31−8220, H89, JNK inh II, AG183) and G_{10} = 0.553−0.595 (SB202190, Cdk2/5 inhibitor, Olomoucin, AG1024, ST638, Genistein, Rottlerin) at 10 μM ATP (based on data in Supporting Information).

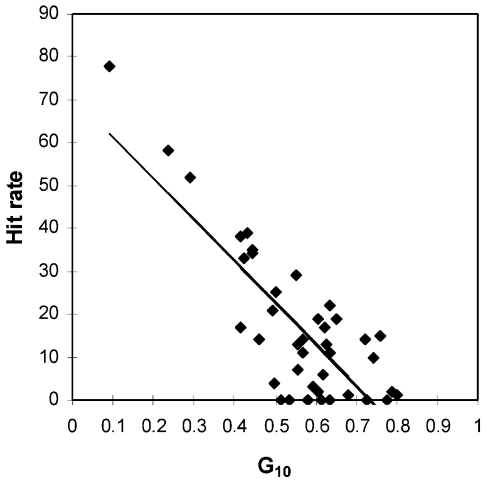

Figure 4. Correlation between Gini at 10 μM (G_{10}) and the relevant hit rate (based on data in Table 3).

However, the variability of the Gini coefficient might decrease if subsets of a larger set of kinases, preferably a whole kinome, are evaluated. To verify this hypothesis using experimental datasets is, unfortunately, not possible for any compound, yet. Nevertheless, a set of 510 kinases ("virtual kinome") can be modeled using the current results for 85 kinases by artificially increasing the population of data points 6-fold. To achieve this, one can introduce an additional five data points identical to each experimentally derived one. The use of such an approach would preserve the selectivity profile of the inhibitor and allow for comparisons to be made. The results for AG1024, Staurosporine, and PD184352 for the subsets of 50, 100, 200, 300, and 400 kinases are shown in Table 5 (Supporting Information) and graphically presented in Figure 6. As expected, the mean values of Gini coefficients for AG1024 fall now in a very narrow range from 0.556 to 0.568 and are much closer to the Gini coefficients calculated for the "virtual kinome" of 510 kinases (0.568). The individual Gini coefficients (0.504−0.642) are not varied as

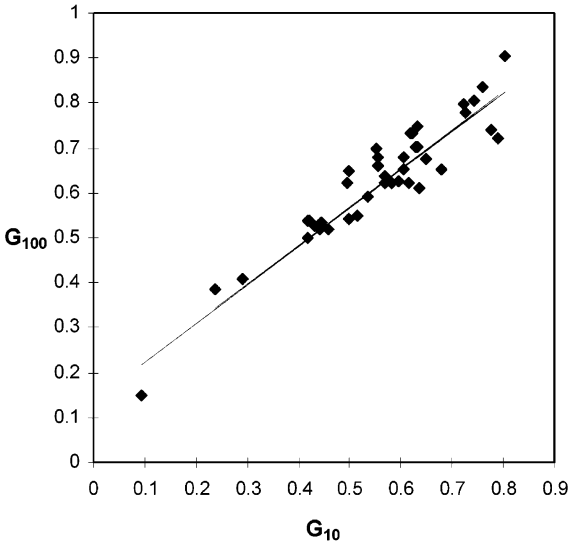

Figure 5. Relationship between G_{10} and G_{100} based on data in Table 3.

Table 4. Gini Coefficients Calculated 13 Times for Each Subset of 10, 20, 40, 60, and 80 Kinases for 10 μM AG1024 at 10 μM ATP Concentration (see Supporting Information)

subset size	10	20	40	60	80
fraction of panel (%)	12	24	47	71	94
			Gini Coefficient		
	0.429	0.590	0.599	0.605	0.572
	0.619	0.576	0.532	0.578	0.572
	0.510	0.488	0.534	0.586	0.572
	0.476	0.540	0.489	0.555	0.558
	0.398	0.514	0.525	0.580	0.567
	0.639	0.579	0.587	0.542	0.556
	0.511	0.505	0.547	0.578	0.570
	0.546	0.544	0.543	0.567	0.578
	0.658	0.575	0.551	0.565	0.563
	0.565	0.561	0.560	0.554	0.572
	0.385	0.477	0.591	0.575	0.573
	0.487	0.566	0.593	0.555	0.558
	0.623	0.505	0.572	0.586	0.573
mean	0.527	0.540	0.556	0.571	0.568
std dev	0.091	0.038	0.032	0.017	0.007

much as before, which results in about 3-fold decrease in standard deviation (0.033 for a subset of 50 kinases which constitute about 10% of the panel). In practice, assuming normal distribution,[22] the Gini coefficient can, therefore, be estimated with 95% confidence and accuracy $\pm$ 0.07 using a random panel of only 50 kinases. Increasing the number of kinases in the screening panel to 200 would improve the accuracy to $\pm$0.02 (standard deviation 0.011; Table 5, Supporting Information). Analogous comments are also valid for Staurosporine and PD184352 (Table 5, Supporting Information). The above demonstrates that the Gini coefficient is population independent for a random selection of kinases. Thus, the overall selectivity of an inhibitor against the whole kinome may be evaluated by profiling an inhibitor against a randomly chosen subset of the kinome. However, since the choice of kinases is limited by the content of the current screening panels, the Gini coefficient determined for 50 kinases is as much representative of the selectivity against the whole kinome as the current screening panels reflect the kinome itself.

It is important to note that the Gini coefficient depends on compound concentration, which potentially may affect the ranking of compounds. This effect of concentration can be assessed by using a mathematical model of a full kinome

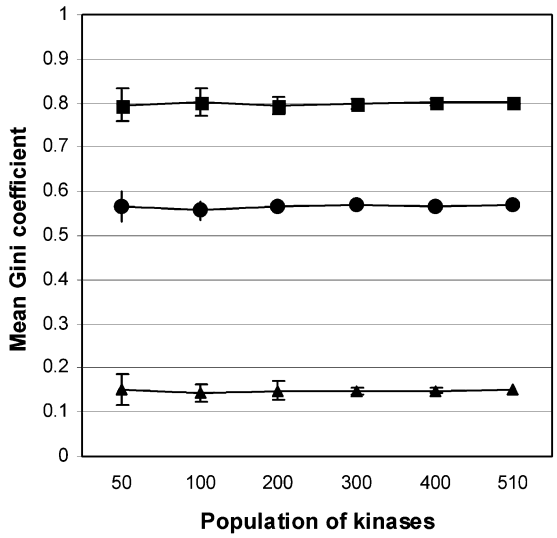

Figure 6. Gini coefficients calculated 13 times for each "virtual kinome" subset of 50, 100, 200, 300, and 400 kinases for 10 μM PD184352 at 10 μM ATP (■), 10 μM AG1024 at 10 μM ATP (●), and 1 μM Staurosporine 100 μM ATP (▲) (based on Table 5 in Supporting Information). Error bars represent standard deviation.

profiling (Supporting Information) wherein the IC_{50} values of an inhibitor against each of the kinases is arbitrarily expressed as:

$$IC_{50} = a \times \text{kinase\_number}^{b}$$

wherein:

a = highest inhibitory potency (in nM) of the inhibitor against the kinome (assumed to be against kinase number 1, usually the target kinase).

kinase_number = sequential number of a kinase within the kinome (518 kinases). Kinases are arranged in the order of increasing IC_{50} value against the profiled compound.

b = selectivity modulator. Values of $0 < b < 1$ allow modeling of relatively nonselective compounds. Selective compounds can be modeled using $b > 1$.

Furthermore, the percentage inhibition of each kinase at a given compound concentration c (in nM) is approximated using a simplified[23] Hill equation

$$\text{inhibition} = \frac{100}{1 + \left(\frac{IC_{50}}{c}\right)^{S}}$$

wherein S is the slope of the dose−response curve.

The percentages of inhibition can then be converted as discussed above (cf. the fifth column in Table 2) into the relevant Gini coefficient against a full kinome at a particular compound concentration c.

The inhibition profile has been modeled for 5 inhibitors having the same maximum potency of IC_{50} = 10 nM (a = 10 nM) and showing different selectivity profile as determined by varying b between 0.5 (low selectivity) and 1.5 (high selectivity) and assuming S = 1.[24] Figure 7 shows how the relevant Gini coefficients depend on compound concentration c.

The Gini coefficient varies little with concentration up to about 1 nM ($^1/_{10}$ of the IC_{50} value). At higher concentrations all the Gini coefficients decrease with increasing concentration, which means that the compounds become less selective. However, the ranking of compounds in terms of selectivity remains unaltered. The Gini coefficient for a more selective

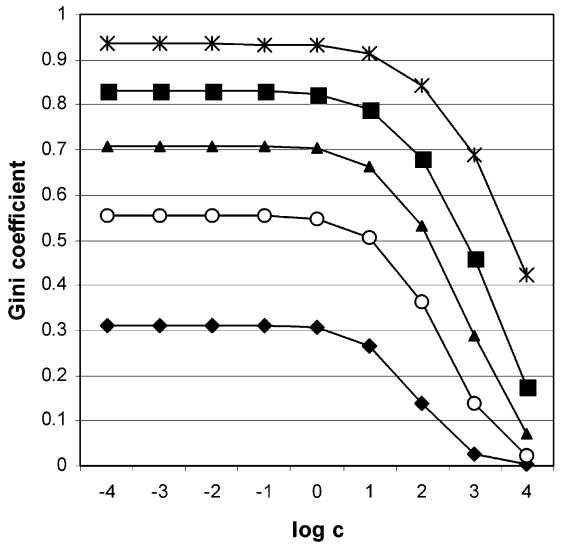

Figure 7. Gini coefficients for compounds with $a = 10$ nM with $b = 0.5$ (◆), 0.8 (○), 1.0 (▲), 1.2 (■), and 1.5 (*) in the concentration range from $c = 10^{-4}$ to 10^4 nM (based on data in Table 6, columns with $r = 0\%$).

compound is always larger than that for the less selective one. Analogous conclusions are also valid for five inhibitors having the same maximum potency of $IC_{50} = 100$ nM ($a = 100$ nM; see Supporting Information).

A closer examination of the model reveals that at very low concentrations the Gini coefficient based on the magnitude of inhibition and the Gini coefficient calculated using the reciprocal of IC_{50} instead of the magnitude of inhibition (last row in Table 6) become equal.[25] Since the IC_{50} values are considered to be a more reliable measure of selectivity, one might prefer to calculate Gini coefficients using the magnitude of inhibition at low concentrations. Unfortunately, such magnitudes are of the order of 0.1% or less (Supporting Information) and are below the level of experimental accuracy.

An estimate of the impact of experimental accuracy r on the Gini coefficient can be obtained by calculating the Gini coefficient values using the percentage inhibition values rounded to the nearest 1% and 2% (Table 6; for examples of the calculation, see Supporting Information). One can notice that rounding the values of percentage inhibition leads to an increase in the Gini coefficients to about 1 at low concentrations, regardless of the actual selectivity of the compound. Such overestimation of the Gini coefficient of the weakly inhibited kinases is due to virtually all low percentages of inhibition becoming zero after rounding with a very few kinases showing inhibition at a measurable level. Weak inhibition is also associated with higher sensitivity of the Gini coefficient to changes in single point residual activities (Supporting Information). For instance, the largest relative effect of about −1.11% on the Gini coefficient is observed when the residual activity of AMPK(r) inhibited by 10 μM PD184352 is decreased from 100% to 90% (i.e., 10% relative). For strongly inhibited kinases, the relevant effect due to 10% change in residual activities is much smaller (0.02−0.20%).

The precision at which Gini coefficients are established benefits as well from stronger inhibition. The coefficients become relatively insensitive to experimental accuracy r (Table 6) only at concentrations between 10^2 and 10^3 nM ($\log c = 2-3$), which correspond to 10−100 times the IC_{50} value for the most potently inhibited kinase. Therefore, selectivity assessment using the Gini coefficient should be performed by profiling a compound at a concentration within such a range. Hence, the maximum value of the measured inhibition should fall within the range between 91 and 99%.

As discussed above, the selectivity of compounds having the same maximum potency a can be compared by profiling them at the same concentration. If compounds have different potencies a, another approach is needed. Inspection of the Gini coefficient values calculated for $a = 10$ and $a = 100$ to assess the effects of concentration (Supporting Information) demonstrates that compounds of different potency a and the same selectivity modulator b would offer the same Gini coefficients only if the ratio between IC_{50} corresponding to the highest potency a and concentration c is the same. Under such conditions, the maximum inhibition is also the same (cf. the Hill equation above; $IC_{50} = a$). For instance, the Gini coefficient G equals 0.140 both when $a = 10$, $b = 0.8$, $r = 2$, $c = 1000$ (ratio $a/c = 0.01$) and when $a = 100$, $b = 0.8$, $r = 2$, and $c = 10000$ (ratio $a/c = 0.01$). In practice, the selectivity of compounds having different potencies a can be compared directly if the maximum observed inhibition is similar. This can be achieved by adjusting the compound concentration c in the assay accordingly.

High selectivity as revealed by Gini raises the probability that an inhibitor safety profile will be sufficient for the compound to enter the clinic, Indeed, PD184352 (CI-1040) was the first MEK-targeted compound to enter clinical trials. Its safety profile allows continuous dosing.[26] Nevertheless, when inhibition of certain individual kinases is known to result in undesired side effects, such kinases need to be included in the profiling set and their inhibition studied in detail by establishing the relevant IC_{50} or K_i values. At the same time, however, a kinase-selective compound may still exhibit additional non-

Table 6. Gini Coefficient Values Calculated Using the Mathematical Model of a Full Kinome Profiling for a Compound with $a = 10$ nM and Selectivity Modulator $b = 0.5$, 0.8, 1.0, 1.2, and 1.5 Assuming Experimental Accuracy r of 0%, 1%, and 2%

	$b = 0.5$			$b = 0.8$			$b = 1.0$			$b = 1.2$			$b = 1.5$		
r (%)	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2
log c															
−4	0.313	*a*	*a*	0.555	*a*	*a*	0.709	*a*	*a*	0.830	*a*	*a*	0.935	*a*	*a*
−3	0.313	*a*	*a*	0.555	*a*	*a*	0.709	*a*	*a*	0.830	*a*	*a*	0.935	*a*	*a*
−2	0.312	0.994	*a*	0.555	*a*	*a*	0.709	*a*	*a*	0.830	*a*	*a*	0.934	*a*	*a*
−1	0.312	0.348	*a*	0.554	0.996	*a*	0.708	0.998	*a*	0.830	0.998	*a*	0.934	0.998	*a*
0	0.307	0.270	0.834	0.548	0.944	0.976	0.703	0.978	0.988	0.825	0.987	0.992	0.932	0.993	0.995
1	0.267	0.138	0.281	0.506	0.472	0.624	0.663	0.821	0.900	0.792	0.938	0.959	0.913	0.978	0.983
2	0.138	0.027	0.138	0.364	0.363	0.366	0.530	0.533	0.541	0.681	0.649	0.747	0.844	0.910	0.930
3	0.027	0.003	0.027	0.139	0.139	0.140	0.287	0.287	0.288	0.461	0.461	0.458	0.691	0.693	0.683
4	0.003	0.003	0.003	0.022	0.022	0.022	0.069	0.069	0.069	0.175	0.175	0.176	0.424	0.425	0.425
$1/IC_{50}$	0.313			0.555			0.709			0.830			0.935		

[a] Not determined due to arithmetic error (division by zero).

kinase-related activities. For instance, the PKC inhibitor[27] Hispidin behaves also as a BACE1 inhibitor.[28] Roscovitin, a cdk5 inhibitor,[29] binds to ion channels.[30] Lavendustin A may affect tubulin polymerization.[31] These cross-reactivities are not considered in the Gini coefficient calculation.

Conclusions

In summary, although the chemical and physical meaning of the Gini coefficient is not as clear as that of an IC_{50} or K_i value, it can provide a very useful tool to quickly estimate the selectivity of kinase inhibitors against a large set of kinases. It works for ATP-competitive and noncompetitive inhibitors. The optimal concentration range for determination of the Gini coefficient is between 10 and 100 times the IC_{50} value for the most potently inhibited kinase. The Gini coefficient allows objective ranking of compounds in terms of their overall selectivity at a pharmacologically relevant concentration. It uses single-point experimental data generated economically at the same, single ATP concentration for all kinases. It is population-independent so that, for instance, the Gini coefficient generated with a 50 kinase panel can be compared with the Gini coefficient obtained by screening against the panel of 100 kinases. Furthermore, the Gini coefficient can give quantitative meaning to the terms "selective" and "nonselective". Most probably the Gini coefficient can also be used to estimate selectivity of inhibitors against other families of protein targets.

Acknowledgment. I would like to thank Upstate/Millipore for providing the inhibition data used in this work.

Supporting Information Available: Calculation of Gini coefficient and hit rate for 85 kinase inhibitors, calculation of example Lorenz curves for PD184352, AG1024, and staurosporine, calculation of Lorenz curves for clusters of Gini 0.417−0.460 and 0.553−0.595 at 10 μM ATP, correlation between Gini values at 10 and 100 μM ATP, calculation of Gini coefficient for subsets of 10, 20, 40, 60, and 80 kinases for AG1024 (10 μM at 10 μM ATP and staurosporine, 1 μM at 100 μM ATP and PD184352, 10 μM at 10 μM ATP), calculation of Gini coefficient for subsets of 50, 100, 200, 300, and 400 kinases for AG1024 (10 μM at 10 μM ATP and staurosporine, 1 μM at 100 μM ATP and PD184352, 10 μM at 10 μM ATP), Gini normal distribution testing based on a subset of 50 kinases inhibited by 10 μM AG1024 at 10 μM ATP, sensitivity testing using AG1024 (10 μM, 10 μM ATP), effect of concentration on the Gini coefficient for $a = 10$ and $a = 100$, example of calculation of the Gini coefficient using $1/IC_{50}$, sensitivity of the Gini coefficient to changes in the value of slope S for $a = 10$, example of calculation of the Gini coefficient for testing its sensitivity to changes in the value of slope S using $a = 10$, $b = 0.5$, $S = 0.9$, $r = 2$, and $c = 100$, and Excel tool to calculate the Gini coefficient. 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) Gill, A. Protein kinases in drug discovery and development. *Drug Discovery Today* **2004**, *9*, 16−7.
(3) de Carcer, G.; de Castro, I. P.; Malumbres, M. Targeting cell cycle kinases for cancer therapy. *Curr. Med. Chem.* **2007**, *14*, 969−85.
(4) Martin, M. W.; Newcomb, J.; Nunes, J. J.; McGowan, D. C.; Armistead, D. M.; Boucher, C.; Buchanan, J. L.; Buckner, W.; Chai, L.; Elbaum, D.; Epstein, L. F.; Faust, T.; Flynn, S.; Gallant, P.; Gore, A.; Gu, Y.; Hsieh, F.; Huang, X.; Lee, J. H.; Metz, D.; Middleton, S.; Mohn, D.; Morgenstern, K.; Morrison, M. J.; Novak, P. M.; Oliveira-dos-Santos, A.; Powers, D.; Rose, P.; Schneider, S.; Sell, S.; Tudor, Y.; Turci, S. M.; Welcher, A. A.; White, R. D.; Zack, D.; Zhao, H.; Zhu, L.; Zhu, X.; Ghiron, C.; Amouzegh, P.; Ermann, M.; Jenkins, J.; Johnston, D.; Napier, S.; Power, E. Novel 2-aminopyrimidine carbamates as potent and orally active inhibitors of Lck: synthesis, SAR, and in vivo antiinflammatory activity. *J. Med. Chem.* **2006**, *49*, 4981−91.
(5) Kulkarni, R. G.; Achaiah, G.; Sastry, G. N. Novel targets for antiinflammatory and antiarthritic agents. *Curr. Pharm. Des.* **2006**, *12*, 2437−54.
(6) Graczyk, P. P.; Khan, A.; Bhatia, G. S.; Palmer, V.; Medland, D.; Numata, H.; Oinuma, H.; Catchick, J.; Dunne, A.; Ellis, M.; Smales, C.; Whitfield, J.; Neame, S. J.; Shah, B.; Wilton, D.; Morgan, L.; Patel, T.; Chung, R.; Desmond, H.; Staddon, J. M.; Sato, N.; Inoue, A. The neuroprotective action of JNK3 inhibitors based on the 6,7-dihydro-5H-pyrrolo[1,2-a]imidazole scaffold. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 4666−70.
(7) Sun, X.; Minohara, M.; Kikuchi, H.; Ishizu, T.; Tanaka, M.; Piao, H.; Osoegawa, M.; Ohyagi, Y.; Shimokawa, H.; Kira, J. The selective Rho-kinase inhibitor Fasudil is protective and therapeutic in experimental autoimmune encephalomyelitis. *J. Neuroimmunol.* **2006**, *180*, 126−34.
(8) Paniagua, R. T.; Sharpe, O.; Ho, P. P.; Chan, S. M.; Chang, A.; Higgins, J. P.; Tomooka, B. H.; Thomas, F. M.; Song, J. J.; Goodman, S. B.; Lee, D. M.; Genovese, M. C.; Utz, P. J.; Steinman, L.; Robinson, W. H. Selective tyrosine kinase inhibition by imatinib mesylate for the treatment of autoimmune arthritis. *J. Clin. Invest.* **2006**, *116*, 2633−42.
(9) Birault, V.; Harris, C. J.; Le, J.; Lipkin, M.; Nerella, R.; Stevens, A. Bringing kinases into focus: efficient drug design through the use of chemogenomic toolkits. *Curr. Med. Chem.* **2006**, *13*, 1735−48.
(10) (a) Liao, J. J.-L. Molecular Recognition of Protein Kinase Binding Pockets for Design of Potent and Selective Kinase Inhibitors *J. Med. Chem.* **2007**, *50*, 409−424. (b) Scapin, G. Protein kinase inhibition: different approaches to selective inhibitor design. *Curr. Drug Targets* **2006**, *7*, 1443−54. (c) Ortiz, A. R.; Gomez-Puertas, P.; Leo-Macias, A.; Lopez-Romero, P.; Lopez-Vinas, E.; Morreale, A.; Murcia, M.; Wang, K. Computational approaches to model ligand selectivity in drug design. *Curr. Top. Med. Chem.* **2006**, *6*, 41−55. (d) Pratt, D. J.; Bentley, J.; Jewsbury, P.; Boyle, F. T.; Endicott, J. A.; Noble, M. E. Dissecting the determinants of cyclin-dependent kinase 2 and cyclin-dependent kinase 4 inhibitor selectivity. *J. Med. Chem.* **2006**, *49*, 5470−7. (e) Fernandez, A.; Maddipati, S. A priori inference of cross reactivity for drug-targeted kinases. *J. Med. Chem.* **2006**, *49*, 3092−100. (f) Luo, Y. Selectivity assessment of kinase inhibitors: strategies and challenges. *Curr. Opin. Mol. Ther.* **2005**, *7*, 251−255.
(11) Chu, X.-J.; DePinto, W.; Bartkovitz, D.; So, S.-S.; Vu, B. T.; Packman, K.; Lukacs, C.; Ding, Q.; Jiang, N.; Wang, K.; Goelzer, P.; Yin, X.; Smith, M. A.; Higgins, B. X.; Chen, Y.; Xiang, Q.; Moliterni, J.; Kaplan, G.; Graves, B.; Lovey, A.; Fotouhi, N. Discovery of [4-Amino-2-(1-methanesulfonylpiperidin-4-ylamino)-pyrimidin-5-yl](2,3-difluoro-6- methoxyphenyl)methanone (R547), a potent and selective cyclin-dependent kinase inhibitor with significant in vivo antitumor activity. *J. Med. Chem.* **2006**, *49*, 6549−60.
(12) Boschelli, D. H.; Wu, B.; Ye, F.; Wang, Y.; Golas, J. M.; Lucas, J.; Boschelli, F. Synthesis and Src kinase inhibitory activity of a series of 4-[(2,4-dichloro-5-methoxyphenyl)amino]-7-furyl-3-quinolinecarbonitriles. *J. Med. Chem.* **2006**, *49*, 7868−76.
(13) 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.; Lelias, 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*, 329−36.
(14) (a) Gini, C. *Variabilitá e mutabilita, 1912* (reprinted in *Memorie di metodologica statistica*); Pizetti, E., Salvemini, T., Eds.; Libreria Eredi Virgilio Veschi: Rome, 1955. (b) Dorfman, R. A Formula for the Gini Coefficient. *Rev. Econ. Stat.* **1979**, *61*, 146−149. (c) Kleiber, C.; Kotz, S. A characterization of income distributions in terms of generalized Gini coefficients. *Soc. Choice Welfare* **2002**, *19*, 789−794.
(15) (a) Haidich, A.; Ioannidis, J. The Gini coefficient as a measure for understanding accrual inequalities in multicenter clinical studies. *J. Clin. Epidemiol.* **2004**, *57*, 341−348. (b) Harcha, B. D.; Correll, R. L.; Meech, W.; Kirkby, C. A.; Pankhurst C. E. Using the Gini coefficient with BIOLOG substrate utilisation data to provide an alternative quantitative measure for comparing bacterial soil communities *J. Microbiol. Methods* **1997**, *30*, 91−101. (c) Kaufmann, T.; Schupfer, G.; Bauer, M. The Gini coefficient. A numerical grading for the degree of standardization of surgical subspecialities. *Der Anaesthesist* **2006**, *55*, 791−6.
(16) Lorenz, M. C. Methods for measuring the concentration of wealth. *J. Am. Stat. Assoc.* **1905**, *9*, 209−219.

(17) Upstate/Millipore's in-house data obtained using direct radiometric KinaseProfiler assay platform and shared with the author for the purposes of this study (see Supporting Information).
(18) Koh, J. Y.; Wie, M. B.; Gwag, B. J.; Sensi, S. L.; Canzoniero, L. M.; Demaro, J.; Csernansky, C.; Choi, D. W. Staurosporine-induced neuronal apoptosis. *Exp. Neurol.* **1995**, *135*, 153−9.
(19) (a) Delaney, A. M.; Printen, J. A.; Chen, H.; Fauman, E. B.; Dudley, D. T. Identification of a novel mitogen-activated protein kinase kinase activation domain recognized by the inhibitor PD 184352. *Mol. Cell Biol.* **2002**, *22*, 7593−602. (b) Davies, S. P.; Reddy, H.; Caivano, M.; Cohen, P. Specificity and mechanism of action of some commonly used protein kinase inhibitors. *Biochem. J.* **2000**, *351*, 95−105.
(20) Ohren, J. F.; Chen, H.; Pavlovsky, A.; Whitehead, C.; Zhang, E.; Kuffa, P.; Yan, C.; McConnell, P.; Spessard, C.; Banotai, C.; Mueller, W. T.; Delaney, A.; Omer, C.; Sebolt-Leopold, J.; Dudley, D. T.; Leung, I. K.; Flamme, C.; Warmus, J.; Kaufman, M.; Barrett, S.; Tecle, H.; Hasemann, C. A. Structures of human MAP kinase kinase 1 (MEK1) and MEK2 describe novel noncompetitive kinase inhibition. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1192−7. Erratum in *Nat. Struct. Mol. Biol.* **2005**, *12*, 278.
(21) The details will be reported elsewhere.
(22) The distribution of the Gini coefficient values is close to normal as demonstrated by repeated calculations of the coefficient using the data for 10 μM AG1024 at 10 μM ATP. Samples of 50 kinases were selected 350 times out of the "virtual kinome" of 510 kinases (skewness: −0.08; kurtosis: −0.21; mean 0.5582 SE 0.0019; standard deviation 0.0357 SE 0.0013; see Supporting Information). About 59% of the Gini values fell within mean ± 1 SD and 97% within mean ± 2 SD. The relevant percentage values expected for the normal distribution are 68% and 95%, respectively.
(23) The bottom response is assumed to be 0% and top response 100%.
(24) The value $S = 1$ assumes an ideal binding curve. For instance, slope S for several PKC isoforms ranges between 0.7 and 1.1 (http://www.invitrogen.com/downloads/O13473_Pherastar_Application_Note_FINAL.pdf accessed on 10 July 2007). For PI3 kinase, it can vary between 0.9 and 1.2 (http://www.upstate.com/img/coa/33-016-manual.pdf accessed on 10 July 2007). The Gini coefficient is not very sensitive to changes in S (Supporting Information).
(25) This is valid only for $S = 1$.
(26) Allen, L. F.; Sebolt-Leopold, J.; Meyer, M. B. CI-1040 (PD184352), a targeted signal transduction inhibitor of MEK (MAPKK). *Semin. Oncol.* **2003**, *30*, 105−16.
(27) Gonindard, C.; Bergonzi, C.; Denier, C.; Sergheraert, C.; Klaebe, A.; Chavant, L.; Hollande, E. Synthetic hispidin, a PKC inhibitor, is more cytotoxic toward cancer cells than normal cells in vitro. *Cell Biol Toxicol.* **1997**, *13*, 141−53.
(28) Park, I. H.; Jeon, S. Y.; Lee H. J.; Kim, S. I.; Song, K. S. A beta-secretase (BACE1) inhibitor hispidin from the mycelial cultures of *Phellinus linteus*. *Planta Med.* **2004**, *70*, 143−6.
(29) Goodyear, S.; Sharma, M. C. Roscovitine regulates invasive breast cancer cell (MDA-MB231) proliferation and survival through cell cycle regulatory protein cdk5. *Exp. Mol. Pathol.* **2007**, *82*, 25−32.
(30) Buraei, Z.; Schofield, G.; Elmslie, K. S. Roscovitine differentially affects CaV2 and Kv channels by binding to the open state. *Neuropharmacology* **2007**, *52*, 883−94.
(31) Mu, F.; Coffing, S. L.; Riese, D. J., II; Geahlen, R. L.; Verdier-Pinard, P.; Hamel, T. E.; Johnson, J.; Cushman, M. Design, synthesis, and biological evaluation of a series of lavendustin A analogues that inhibit EGFR and Syk tyrosine kinases, as well as tubulin polymerization. *J. Med. Chem.* **2001**, *44*, 441−52.

JM070562U