Supplemental Tables

Protein target	PDB id	Inhibitor molecular weight (Da)	Potency (μM)	Ligand efficiency (kcal/mol per heavy atom)	θ_{lig}
βII tryptase	2bm2	402	0.015	0.35	0.56
penicillin G acylase	1gm8	350	16	0.27	0.46
urokinase	1owe	291	0.63	0.38	0.41
ADAM33	1r55	331	0.16	0.40	0.39
β-lactamase	1l2s	318	26	0.33	0.39
factor Xa	1lpz	467	0.025	0.32	0.38
Chk1	2br1	391	15.4	0.23	0.38
NS5B polymerase	1yvf	438	0.1	0.34	0.38
factor VIIa	1ygc	548	0.00035	0.34	0.37
vascular endothelial growth factor receptor 2	1y6b	479	0.038	0.30	0.37
tRNA-guanine transglycosylase	1n2v	208	83	0.37	0.36
tryptophan synthase	1k3u	290	-NR-	-NR-	0.36
cyclin-dependent kinase 2	1ke5	329	0.56	0.37	0.35
thymidine kinase	1ofl	252	4.1	0.41	0.34
thyroid hormone receptor α1	1nav	355	0.025	0.45	0.34
neuraminidase B	1vcj	351	26	0.25	0.34
thymidine phosphorylase	1uou	245	0.02	0.65	0.34
neuraminidase A	1l7f	330	0.0008	0.54	0.33
c-Abl tyrosine kinase	1opk	427	0.00015	0.46	0.33
protein kinase 5	1v0p	433	0.13	0.31	0.33
progesterone receptor	1sqn	298	0.0004	0.58	0.33
p38 kinase	1ywr	475	0.032	0.29	0.33
androgen receptor	1z95	430	0.076	0.33	0.33
thyroid hormone receptor β1	1n46	367	0.00003	0.53	0.32
purine nucleoside phosphorylase	1v48	337	0.0069	0.50	0.32
estrogen receptor α	1sj0	466	0.0008	0.37	0.32
thymidylate kinase	1w2g	242	27	0.37	0.32
c-Jun terminal kinase 3	1pmn	487	0.007	0.34	0.32
glucokinase	1v4s	349	1000	0.18	0.32
phosphodiesterase 5A	1xoz	389	0.0012	0.42	0.31
myosin II	1yv3	292	4.9	0.33	0.31
c-kit tyrosine kinase	1t46	496	0.413	0.23	0.30
heat shock protein 90	2bsm	388	0.14	0.34	0.30
thrombin	1oyt	408	0.057	0.33	0.30
cyclin-dependent kinase 5	1unl	354	0.2	0.35	0.30
glucocorticoid receptor	1m2z	392	0.06	0.35	0.29

glutamate receptor 6	1tt1	215	64.6	0.38	0.29
vitamin D nuclear receptor	1s19	413	0.0017	0.40	0.28
glycogen synthase kinase 3β	1q41	277	0.022	0.50	0.28
activated Cdc42 kinase 1	1u4d	245	-NR-	-NR-	0.28
transthyretin	1tz8	268	-NR-	-NR-	0.27
adipocyte fatty acid-binding protein	1tow	253	0.57	0.45	0.27
dihydrofolate reductase	1s3v	375	0.038	0.37	0.27
acetylcholinesterase	1gpk	244	4.3	0.41	0.27
HIV-1 reverse transcriptase	1jla	364	-NR-	-NR-	0.26
HIV-1 protease	1kzk	576	0.00004	0.34	0.24

Table S1: Inhibitors bound to traditional targets, a subset of the Astex set. Potency is taken from reported K_d or K_i values where available; if unavailable, IC_{50} values were used instead. In several cases ("-NR-"), no measure of potency has been reported. θ_{lig} indicates the fraction of ligand SASA exposed in the complex, as defined in Equation 1.

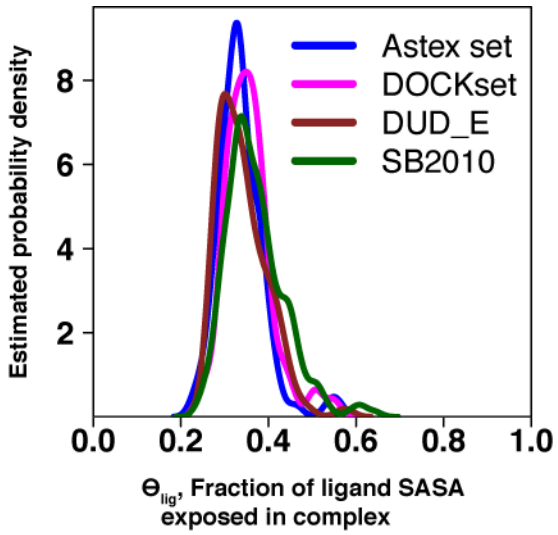

Figure S1 (complements Figure 1) The distribution of the extent of inhibitor solvent exposure (θ_{lig}) is similar across a number of drug-like sets: Astex [27] (*blue*), DOCK [54] (*magenta*), DUD-E [29] (*brown*), and SB2010 [30] (*green*).

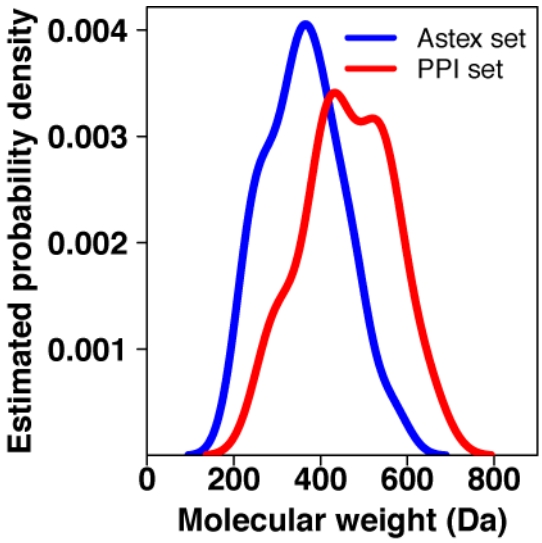

Figure S2 (complements Figure 2) The distribution of molecular weights for the inhibitors in each set underlies the observed difference in ligand efficiencies. Inhibitors binding at protein interaction sites (*red, median value 475 Da*) are typically larger than their drug-like counterparts (*blue, median value 355 Da*), and the difference in the means is statistically significant ($p < 10^{-4}$).

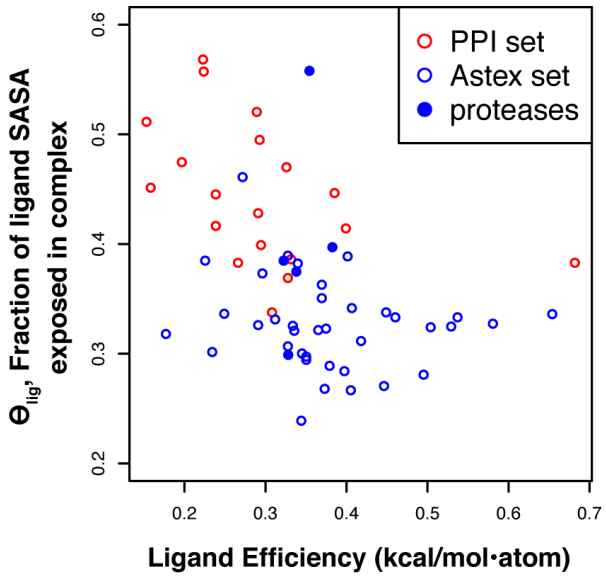

Figure S3 (complements Figure 2) The relationship between θ_{lig} and ligand efficiency for both the PPI set and the Astex set. While there is a statistically significant negative correlation between these properties for the PPI set (as noted in **Figure 2c**), no statistically significant correlation exists for the Astex set ($p = 0.27$).

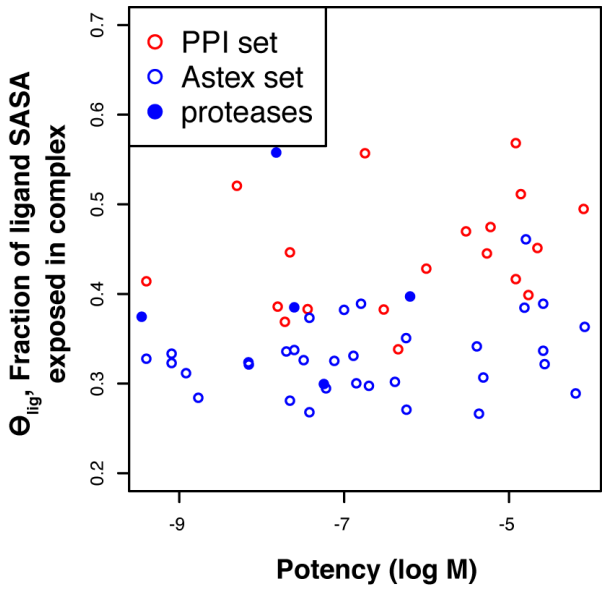

Figure S4 (complements Figure 2) The relationship between θ_{lig} and potency for both the PPI set and the Astex set. No statistically significant correlation exists between these properties for either set ($p = 0.33$ for the PPI set, $p = 0.50$ for the Astex set).

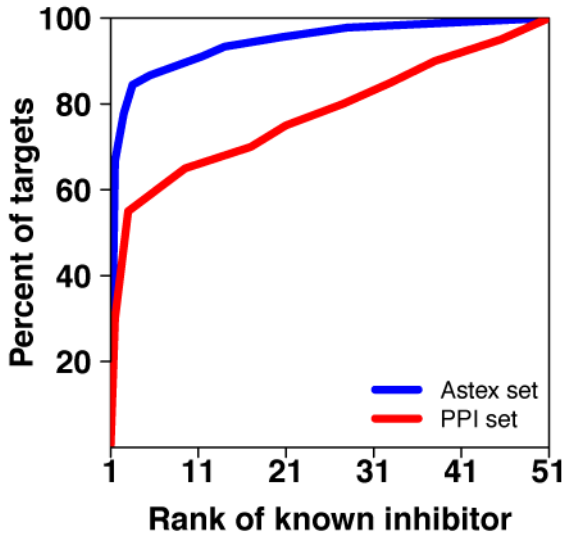

Figure S5 (complements Figure 3) A known inhibitor was embedded in a custom set of 50 "decoy" compounds selected by the DUD-E server to match the physical properties of the known inhibitor. FRED exhibits superior ability to identify the known drug-like inhibitors from the decoy compounds (*blue*), relative to inhibitors that bind at protein interaction sites (*red*), and the difference in the means is statistically significant ($p < 0.002$).

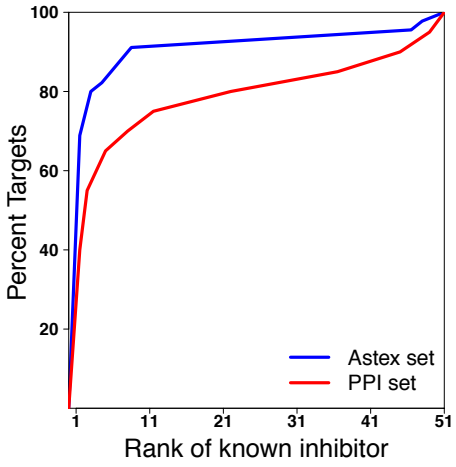

Figure S6 (complements Figure 3) The observation that virtual screening at protein interaction sites performs less well than for drug-like compounds holds for other docking software as well. DOCK 6.6 was used to identify a known inhibitor embedded in a custom set of 50 "decoy" compounds selected by the DUD-E server to match the physical properties of the known inhibitor. DOCK 6.6 exhibits superior ability to identify the known drug-like inhibitors from the decoy compounds (*blue*), relative to inhibitors that bind at protein interaction sites (*red*), though this difference is not statistically significant ($p = 0.11$).