Supplementary Information

Why some targets benefit from beyond rule of five drugs

Megan Egbert, Adrian Whitty[2], György M. Keserű[3]* and Sandor Vajda[1,2]**

Department of Biomedical Engineering, Boston University; 44 Cummington Mall, Boston, MA 02215

*e-mail: vajda@bu.edu

Table S1. Statistical Analysis for ligand binding profiles

Protein Name	Type	n	pBA vs. MW				pBA vs. total probes associated			
			slope	r-value	p-value	std err	slope	r-value	p-value	std err
HIV-1 Protease	Complex I	325	0.0007	0.0440	0.4297	0.0009	0.0696	0.2758	0.0000	0.0135
Heat Shock Protein 90	Complex I	155	0.0099	0.5656	1.74E-14	0.0012	0.1000	0.2992	0.0002	0.0258
Thrombin	Complex I	150	0.0096	0.5426	7.33E-13	0.0012	0.1377	0.4314	0.0000	0.0237
MAP Kinase p38	Complex I	144	0.0098	0.5649	1.65E-13	0.0012	0.0517	0.3239	0.0001	0.0127
Bromodomain BRD4	Complex I	113	0.0061	0.4624	2.53E-07	0.0011	0.0587	0.3306	0.0003	0.0159
Renin	Complex I	62	0.0125	0.7545	1.43E-12	0.0014	0.0875	0.5468	0.0000	0.0173
PPAR-γ	Complex I	52	0.0105	0.7280	9.65E-10	0.0014	0.0833	0.3058	0.0275	0.0367
MAP Kinase Kinase	Complex I	25	0.0127	0.7209	4.80E-05	0.0025	0.1120	0.5396	0.0054	0.0364
E3 ubiquitin-protein ligase X1AP	Complex I	20	0.0043	0.6479	0.0020	0.0012	0.0767	0.7756	0.0001	0.0147
Epidermal Growth Factor Receptor	Complex II	59	-0.0005	-0.0312	0.8148	0.0023	0.0420	0.2160	0.1004	0.0251
Hepatocyte Growth Factor Receptor	Complex II	54	0.0012	0.1306	0.3467	0.0012	0.0142	0.1887	0.1719	0.0102
Anaplastic Lymphoma Kinase	Complex II	32	-0.0041	-0.2366	0.1922	0.0031	0.0829	0.1724	0.3454	0.0864
Tyrosine Protein Kinase ABL1	Complex II	28	-0.0018	-0.1283	0.5151	0.0027	-0.0538	-0.4417	0.0186	0.0214
VEGFR-2	Complex II	24	0.0081	0.3254	0.1208	0.0050	0.0477	0.3163	0.1321	0.0305
Polo-like Kinase 1	Complex II	11	0.0094	0.3350	0.3139	0.0089	-0.2332	-0.8796	0.0004	0.042
Glucocorticoid Receptor	Complex II	9	0.0013	0.1416	0.7164	0.0034	-0.0368	-0.2591	0.5009	0.0518
Cyclophilin A	Complex III	8	0.0055	0.8306	0.0107	0.0015	0.8359	0.8616	0.0060	0.2011
Bcl-2	Complex III	7	0.0072	0.7173	0.0696	0.0031	0.0307	0.2305	0.6189	0.0579
DOT1-like Histone H3 Methyltransferase	Complex III	5	0.0192	0.9930	0.0007	0.0013	0.1668	0.9704	0.0061	0.024
Phosphoinositide-3 Kinase	Simple	64	0.0079	0.5629	1.29E-06	0.0015	-	-	-	-
Soluble Acetylcholine Receptor	Simple	36	0.0055	0.5078	0.0016	0.0016	-	-	-	-
Protein Farnesyltransferase	Simple	34	-0.0038	-0.3003	0.0844	0.0021	-	-	-	-
Kinesin Eg5	Simple	28	0.0072	0.4866	0.0087	0.0026	-	-	-	-
HMG-CoA Reductase	Simple	19	0.0039	0.2570	0.2882	0.0036	-	-	-	-
Hepatitis C Virus NS5b Subunit	Simple	8	0.0031	0.3434	0.4049	0.0035	-	-	-	-
Hepatitis C Virus NS34A Protease	Simple	5	0.0019	0.2692	0.6614	0.0039	-	-	-	-

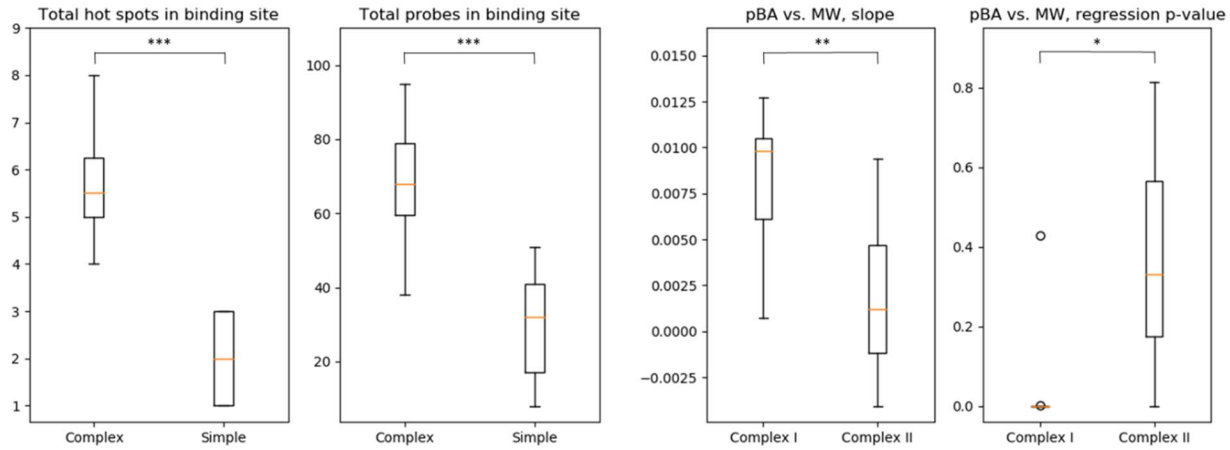

Figure S1. Statistical Analysis between **a)** "Complex" and "Simple" binding sites for the 37 protein targets. Total number of hot spots (nHS) in the binding site (left) has $nHS_{Complex}$ = 5.625 and nHS_{Simple} = 2.15, Wilcoxon-Mann-Whitney p-value = 2.66 E-7. Total number of probe clusters (PC) in the binding site (right) has $PC_{Complex}$ = 68.88, PC_{Simple} = 29.85, Wilcoxon-Mann-Whitney p-value = 9.75 E -7. **b)** "Complex I" and "Complex II" regression for pBA vs. MW ligand binding profiles. Slope (left) has $slope_{Complex\ I}$ = 0.0085, $slope_{Complex\ II}$ = 0.0019, Wilcoxon-Mann-Whitney p-value = 0.0064. Regression p-value (right) has $pvalue_{Complex\ I}$ = 0.048, $pvalue_{Complex\ II}$ = 0.377, Wilcoxon-Mann-Whitney p-value = 0.013. (HIV-1 protease is the clear outlier, with slope = 0.0007 and p-value 0.4297, as explained in the text).

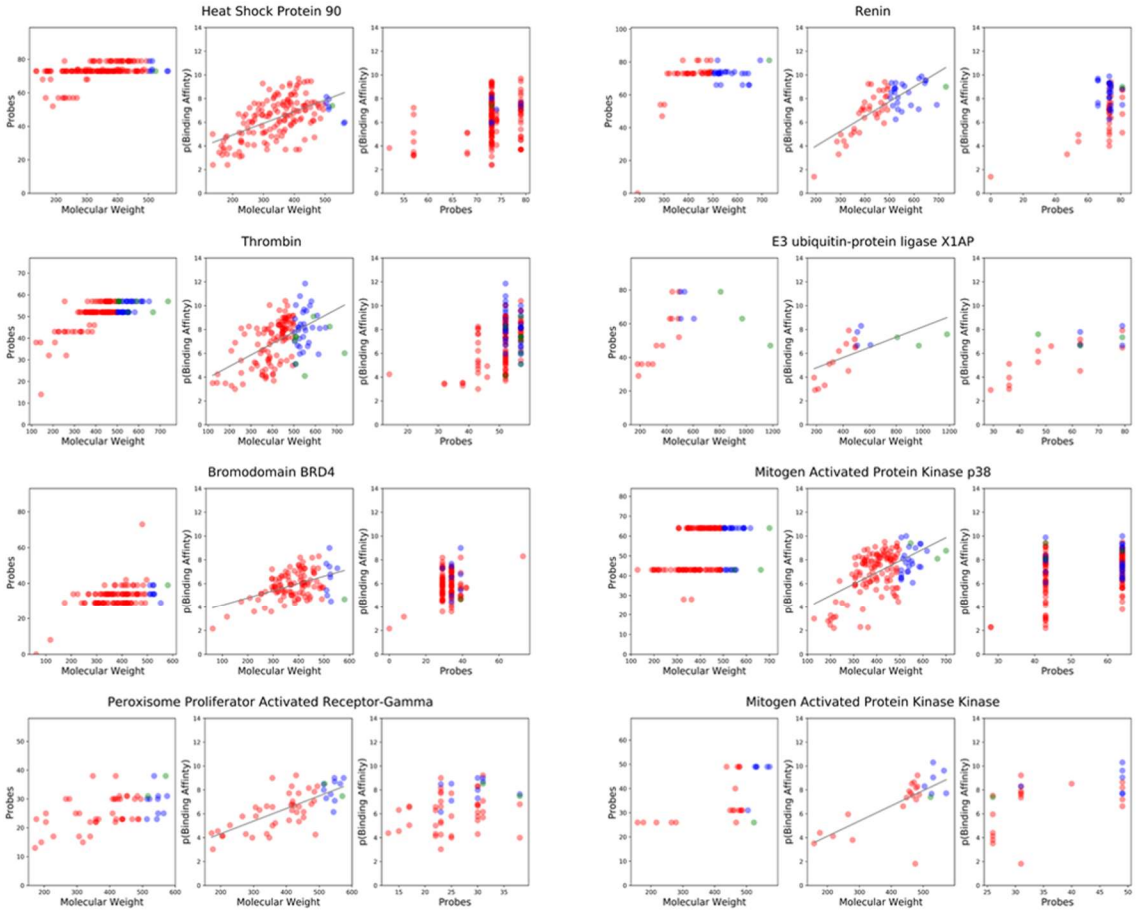

Figure S2. Structure-based ligand binding profiles of Complex I targets (excluding HIV-1 Protease, shown in Figure 4) with known binding affinity and co-crystal structures with the protein. *Probes* represents the total number of probe clusters utilized by the ligand. *p(Binding Affinity)* is the negative log of the binding affinity, expressed as K_D, K_I, or IC_{50}. *Molecular weight* is that of the ligand, measured in Daltons. Legend: Red = ligands with MW < 500 Da, Blue = eRo5 ligands, Green = bRo5 ligands (see Methods).

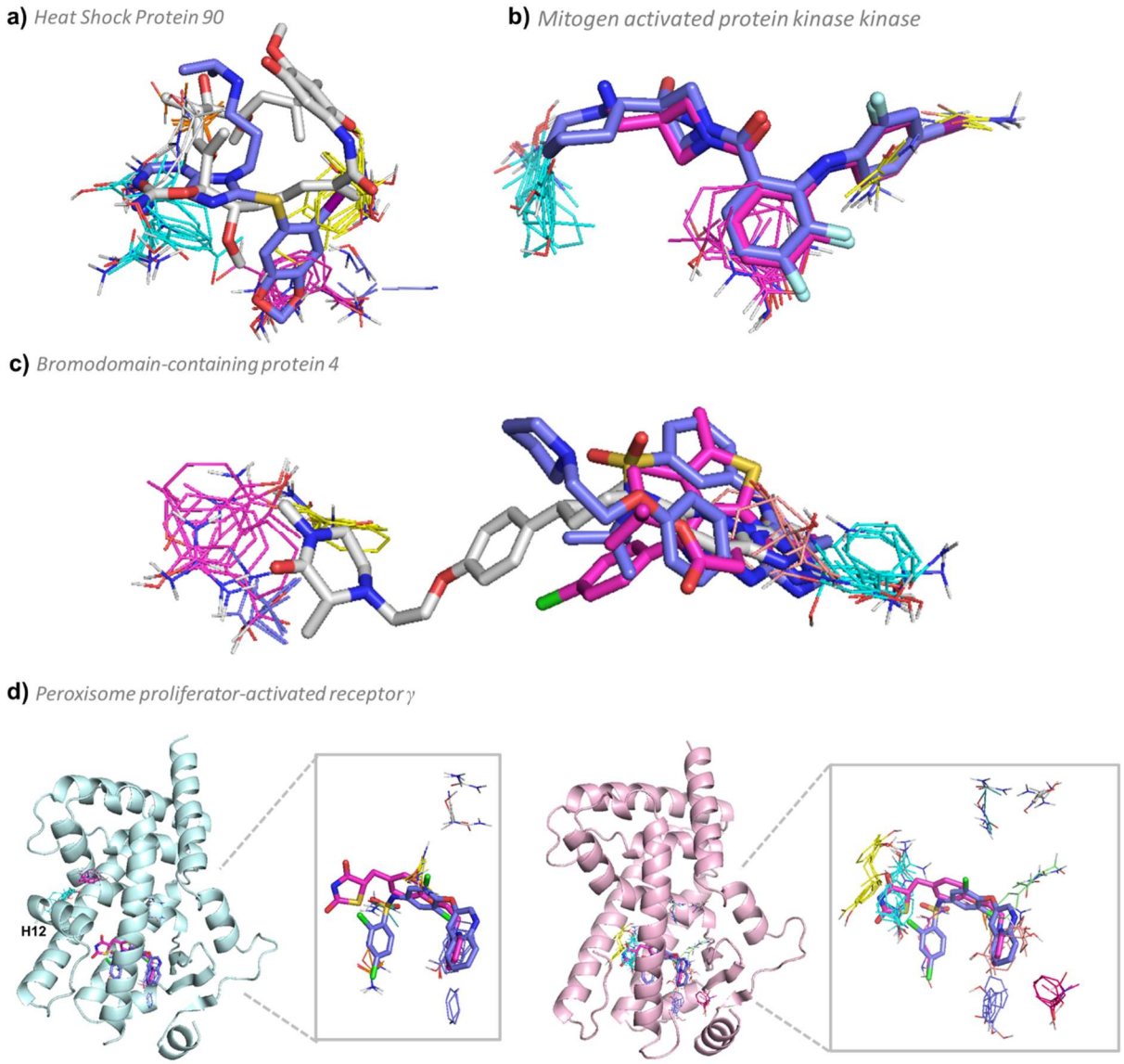

Figure S3. Ligand – binding site interactions for additional complex I targets. FTMap hot spots are colored according to standard output, described in Figure 1. **a)** Hot spot structure of heat shock protein 90. Blue = PU-H71 (2FWZ: H71, MW=512.37 Da, IC_{50}=50 nM, pBA = 7.30). White = geldanamycin (1YET: GDM, MW = 560.64 Da, K_D = 1.2 μM, pBA = 5.92). **b)** Hot spot structure of MEK1. Pink = (4ANB:YQY, MW = 477.22 Da, IC_{50} = 6.6 nM, pBA = 8.8). Blue = Cobimetinib (4AN2:EUI, MW = 531.31 Da, IC_{50} = 0.9 nM, pBA = 9.05). **c)** Hot spot structure of BRD4. Pink = JQ1 (3MXF:JQ1, MW = 458.00 Da, K_D = 49 nM, pBA = 7.31). Blue = fedratinib (4OGJ: 2TA, MW = 524.68 Da, K_D = 164 nM, pBA = 6.79). White = (5KHM: XNH, MW = 479.57 Da, K_I = 5 nM, pBA = 8.3). **d)** FTMap hot spots for PPARγ in a ligand-bound structure (3FUR, left, light blue) and a ligand-free structure (2WHR, right, light pink) overlaid with ligands: pink = rosiglitazone (1ZGY:BRL, MW = 357.43 Da, K_I = 1.0 nM, pBA = 9) and blue = INT131 (3FUR:Z12, MW = 514.21 Da, K_I = 10 nM, pBA = 8).

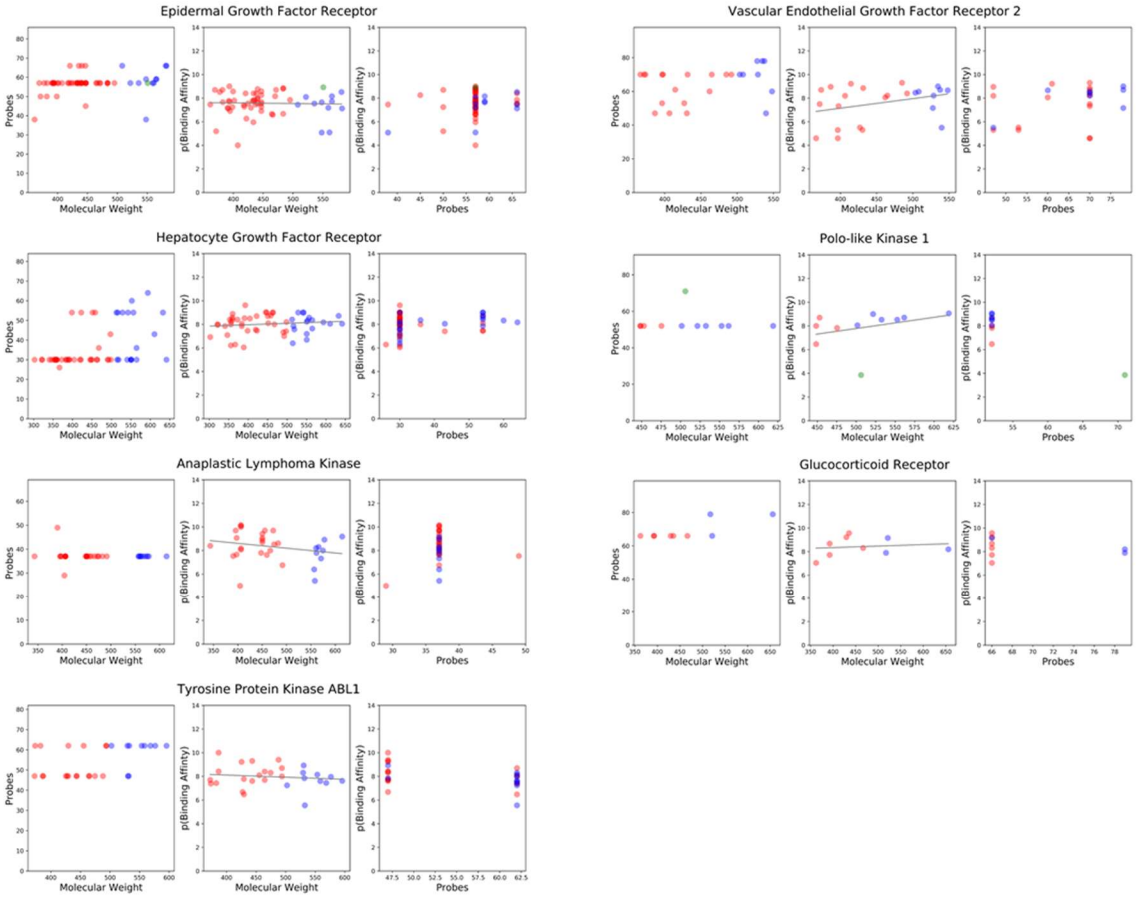

Figure S4. Structure-based ligand binding profiles of Complex II targets with known binding affinity and co-crystal structures with the protein. *Probes* represents the total number of probe clusters utilized by the ligand. *p(Binding Affinity)* is the negative log of the binding affinity, expressed as K_D, K_I, or IC_{50}. *Molecular weight* is that of the ligand, measured in Daltons. Legend: Red = ligands with MW < 500 Da, Blue = eRo5 ligands, Green = bRo5 ligands (see Methods).

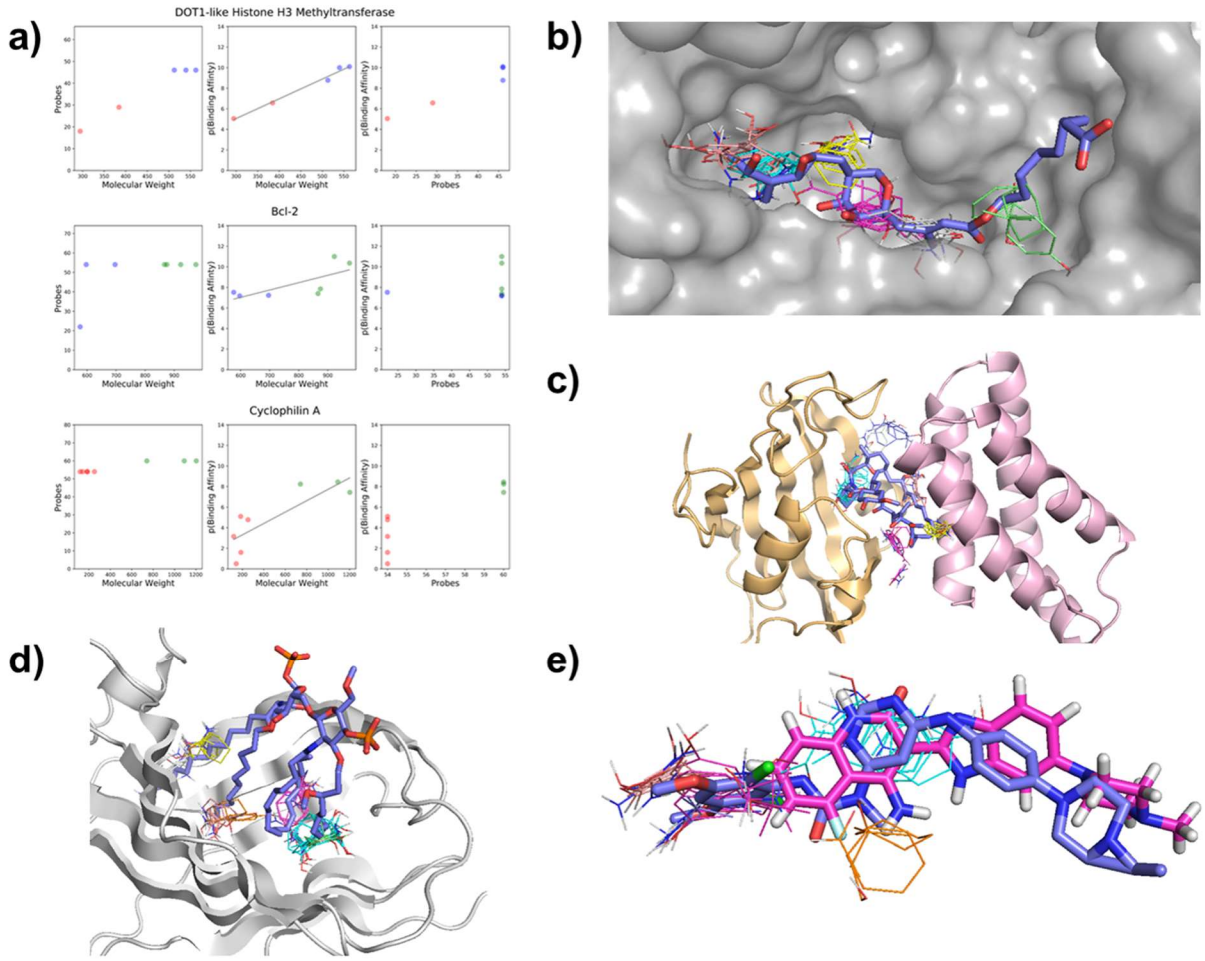

Figure S5. Binding information for additional targets in Complex III. Hot spots are identified by FTMap and colored according to standard output (Figure 1). **a)** Structure-based ligand binding profiles of DOT1-like histone H3 methyltransferase, Bcl-2, and Cyclophilin A. **b)** Binding of pseudomonic acid A (1QU2:MRC, MW = 500.6 Da, KD = 0.14 nM, pBA = 9.85) to isoleucyl-tRNA synthetase. **c)** Binding of rapamycin (4DRI:RAP, MW = 914.17 Da, K_I = 1.0 nM, pBA = 9) to FKBP12 (left, light orange) and mTOR (right, light pink). **d)** Binding of eritoran (2Z65:E55, MW = 1317.7 Da, IC_{50} = 1.5 nM, pBA = 8.82) to toll-like receptor 4. **e)** Hot spot structure of FGFR1 bound to: pink = dovitinib (5A46: 38O, MW = 392.43 Da, no binding affinity data), blue = BGJ398 (3TT0:07J, MW = 560.5 Da, K_I = 13 nM, pBA = 7.89).

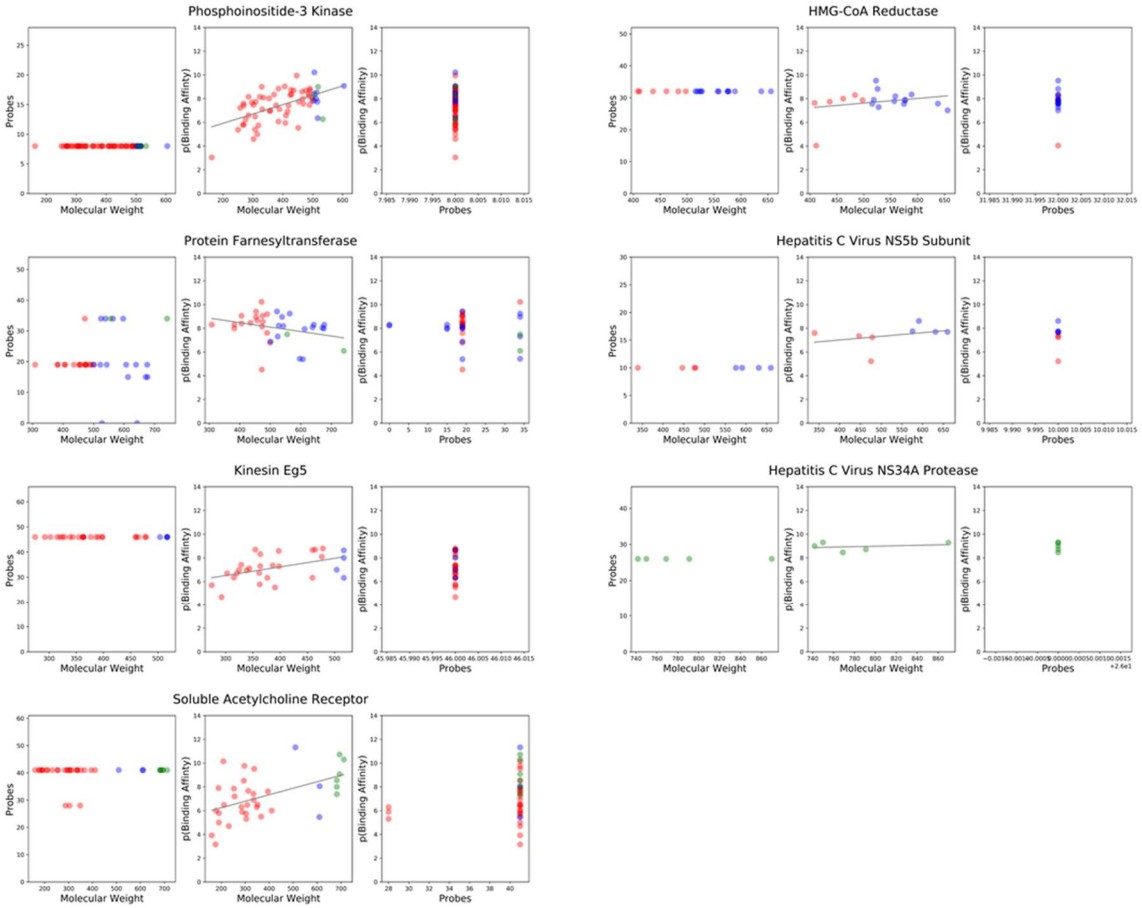

Figure S6. Structure-based ligand binding profiles of simple targets with known binding affinity and co-crystal structures with the protein. *Probes* represents the total number of probe clusters utilized by the ligand. *p(Binding Affinity)* is the negative log of the binding affinity, expressed as K_D, K_I, or IC_{50}. *Molecular weight* is that of the ligand, measured in Daltons. Legend: Red = ligands with MW < 500 Da, Blue = eRo5 ligands, Green = bRo5 ligands (see Methods).

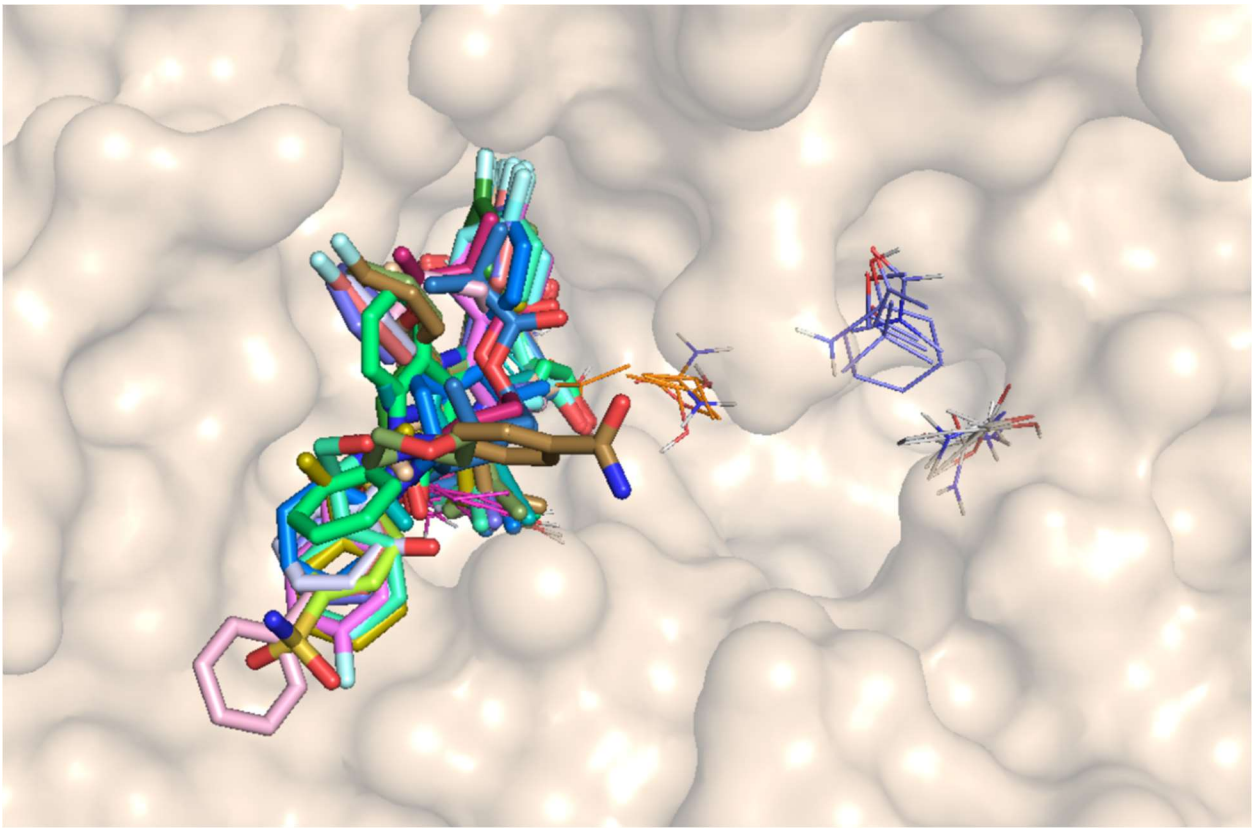

Figure S7. The binding site of HMG-CoA reductase (PDB=1HWK) with ligands overlaid. The n=19 ligands associate with three hot spots: 0(13), 1(13), and 6(6), which are mostly covered by the ligands here. Two hot spots: 4(9) and 5(7) exist to the right of the hot spot, but are not utilized by any of the existing ligands.

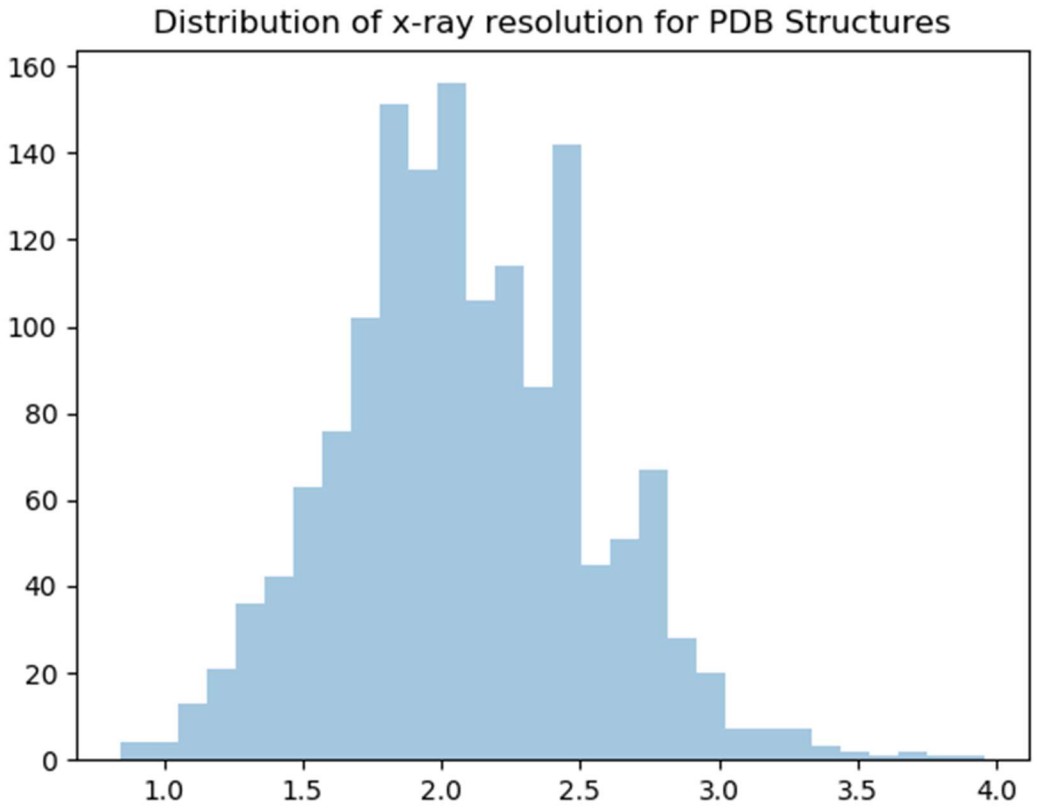

Figure S8. Distribution of x-ray resolution for PDB structures of ligands co-crystallized with a protein of interest. All but 5 ligands utilized in the structure-based ligand profiles have structures determined by x-ray crystallography and their resolutions are distributed as displayed.