

Supporting Information to

Modeling Ordered Waters in Molecular Docking

*Niu Huang, Brian K. Shoichet**

Department of Pharmaceutical Chemistry, University of California-San Francisco,
1700 4th St., CA 94158-2330

* To whom correspondence should be addressed. B.K.S (Email) shoichet@cgl.ucsf.edu (Phone)
415-514-4126 (Fax) 415-514-4260

Protein	No. wat	No. config	EF ₁		EF ₂₀	
			ref.	wat.	ref.	wat.
CDK2	7	128	0.0	2.0	0.4	2.2
EGFr	6	64	4.0	8.8	2.0	3.4
FGFr1	3	8	0.0	0.0	0.1	0.8
HSP90	6	96	0.0	21.5	2.8	2.8
SRC	6	64	2.4	3.1	1.1	2.2
TK	5	32	54.0	45.0	5.0	5.0
VEGFr2	6	64	6.7	0.0	1.5	1.6
Fxa	5	32	14.6	37.0	3.9	4.6
Thrombin	5	32	13.7	22.8	2.9	3.0
Trypsin	5	32	20.0	31.5	2.8	4.7
ACE	6	64	46.5	32.3	3.3	3.8
ADA	7	128	0.0	0.0	2.4	3.4
COMT	2	4	8.2	41.2	2.5	2.9
PDE5	7	128	0.0	0.0	0.7	2.1
DHFR	2	6	41.4	37.9	3.8	4.1
GART	1	2	14.1	14.1	4.0	3.5
AChE	8	256	0.0	4.7	1.7	3.3
AmpC	6	216	4.7	28.3	4.7	4.5
GPB	8	256	0.0	11.4	1.1	2.4
HIVPR	4	16	11.2	11.2	2.2	2.7
HMGR	3	8	36.8	39.6	2.1	2.0
NA	4	16	40.4	64.7	2.8	3.6
PNP	3	8	27.7	27.7	4.4	4.0
SAHH	1	2	81.0	75.0	5.0	5.0

Table S1. Docking enrichment factors for 24 protein targets using the displaceable water method (wat.) and for a reference set where displaceable waters are not modeled (ref.). The PDB structure used were as described,¹⁴ except for CDK2, EGFr, SRC and VEGFr2, where good binding site waters could not be identified and a different structure was used (PDB IDs 1di8, 1xkk, 1y57 and 1ywn, respectively).

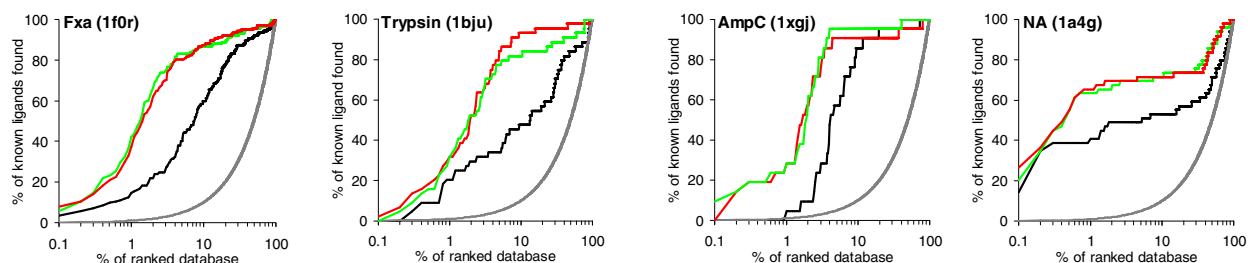
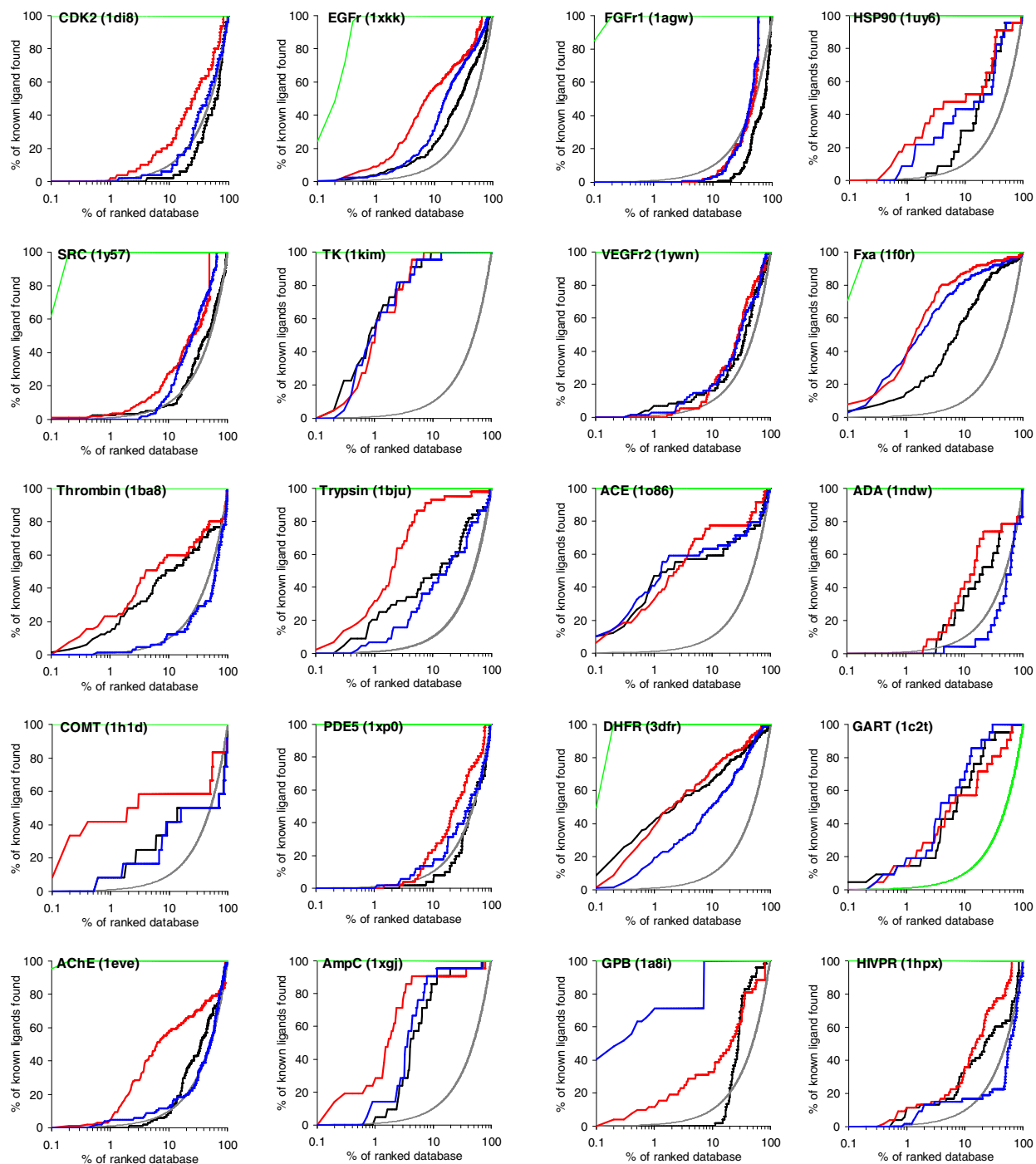


Figure S1. Docking enrichment plots where the black line is docking enrichment calculated by neglecting the water molecules and the red line is the enrichment calculated using the flexible-water sampling algorithm during docking, and the green line represents the enrichment calculated by adding the known “replaceable” water molecules that extracted from the protein apo crystal structure to occupy the ligand binding region, except the crystallographic waters selected from the protein-ligand holo complex structure.



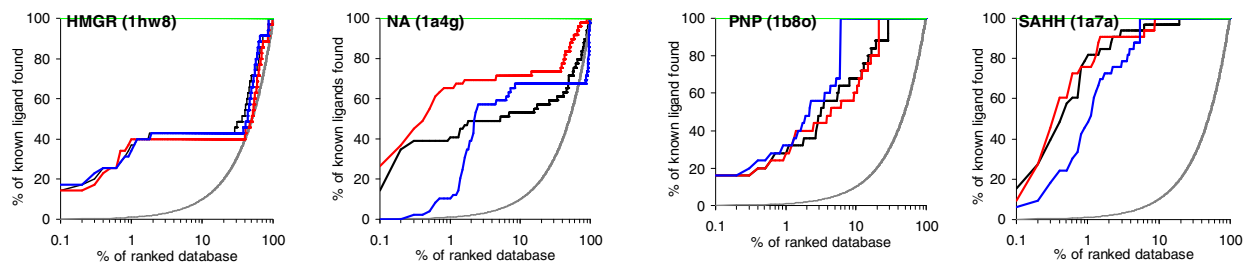


Figure S2. Docking enrichment plots. The gray line is random selection, the green line is the perfect selection, the black line is docking enrichment calculated by neglecting the water molecules, the blue line represents the enrichment calculated by fixing water molecules during docking and the red line is the enrichment calculated using the flexible-water sampling algorithm during docking. Targets are listed in same order as in Table 1 and Figure 2.