Supplementary Information

Ensemble Docking in Drug Discovery: How Many Protein Configurations from Molecular Dynamics Simulations Are Needed to Reproduce Known Ligand Binding?

Wilfredo Evangelista Falcon
Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, TN;
and
UT/ORNL Center for Molecular Biophysics, Oak Ridge, TN;

Sally R. Ellingson
College of Medicine,
University of Kentucky, Lexington, KY

Jeremy C. Smith*
Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, TN;
and
UT/ORNL Center for Molecular Biophysics, Oak Ridge, TN

Jerome Baudry*
Department of Biological Sciences,
The University of Alabama in Huntsville, Huntsville, AL

Table S1. Thresholds to determine outliers for ADORA2A, ADRB2, OPRD1, and OPRK1 in random selection distributions.

ADORA2A

Actives:	844	**Decoys:**	10899	**Total:**	11743
#Percentage%	0.5 %	1.0 %	5.0 %	10.0 %	
Expected sample size	59	117	587	1174	
Expected Actives.	4	8	42	84	
#Upper Value ITQ	9	17	59	107	

ADRB2

Actives:	447	**Decoys:**	15255	**Total:**	15702
#Percentage%	0.5 %	1.0 %	5.0 %	10.0 %	
Expected sample size	79	157	785	1570	
Expected Actives.	2	4	22	45	
#Upper Value ITQ	7	11	35	63	

OPRD1

Actives:	377	**Decoys:**	14703	**Total:**	15080

#Percentage%	0.5 %	1.0 %	5.0 %	10.0 %
Expected sample size	75	151	754	1508
Expected Actives.	2	4	19	38
#Upper Value ITQ	7	10	32	55

OPRK1

Actives:	307	**Decoys:**	11973	**Total:**	12280

#Percentage%	0.5 %	1.0 %	5.0 %	10.0 %
Expected sample size	61	123	614	1228
Expected Actives.	2	3	15	31
#Upper Value ITQ	4	8	26	45