Supporting Information

An Unconventional 2D Shape Similarity Method Affords Comparable Enrichment as a 3D Shape Method in Virtual Screening Experiments

*Jerry Osagie Ebalunode and Weifan Zheng**

Department of Pharmaceutical Sciences, BRITE Institute, North Carolina Central University, 1801 Fayetteville Street, Durham, NC 27707

Email: wzheng@nccu.edu

Table of Contents

1. Table **S1**

2. Table **S2**

3. Figure **S3**

Table **S1**. Comparison of the computational time costs for performing virtual screening experiments using EZSim and ROCS multi-conformer ROCS single conformer.

Target [1]	Mean Time for EZSim (mins)	Mean Time for ROCS mconf [2] (mins)	Mean Time for ROCS sconf [3] (mins)	Speedup [4]
AR	0.6	2.3	0.6	4
ER	0.4	3.2	0.6	7
HSP90	0.2	51.2	0.3	259
MR	0.1	8.3	0.1	58
NA	0.4	57.4	0.4	160
PDE5	0.5	179.4	0.6	370
PR	0.2	3.9	0.2	18
RXR	0.2	18.7	0.1	95
SAHH	0.2	24.4	0.4	116
TK-SRC	1.4	105.0	2.0	73

[1] 10 biological targets were taken for the DUD database. The targets include: Androgen receptor (AR), Estrogen receptor agonists (ER), Human heat shock protein kinase (HSP90), Mineralcorticoid receptor (MR), Neuramindase (NA), Phosphodiesterase V (PDE5), Progesterone receptor (PR), Retinoic X receptor alpha (RXR), S-adenosyl-homocysteine hydrolase (SAHH), and Tyrosine kinase SRC (TK-SRC).

[2] mconf = multiple conformers

[3] sconf = single conformers

[4] The speedup is calculated as the ratio of time required by ROCS multiple conformer over that required by EZSim. The maximum speedup is about 370 fold.

Table **S2.** Comparison of performance (average AUC) of EZsim, ROCS multi-conformer and ROCS single conformer across DUD database.

Target	Mean AUC			Mean Ligands RotorCount[3]	Mean Decoys RotorCount[4]
	EZsim	ROCS mconf[1]	ROCS sconf[2]		
ace	0.59	0.56	0.58	7	7
ache	0.63	0.66	0.61	6	5
ada	0.76	0.80	0.72	3	4
alr2	0.52	0.58	0.55	2	3
ampc	0.77	0.86	0.72	4	4
ar	0.78	0.67	0.67	1	3
cdk2	0.65	0.66	0.65	4	5
comt	0.61	0.77	0.57	3	3
cox1	0.48	0.46	0.45	3	3
cox2	0.88	0.88	0.83	4	4
dhfr	0.96	0.98	0.88	4	5
egfr	0.85	0.92	0.80	4	5
er_agonist	0.94	0.87	0.76	1	3
er_antagonist	0.78	0.84	0.80	6	7
fgfr1	0.72	0.77	0.70	5	7
fxa	0.71	0.84	0.69	7	8
gart	0.82	0.89	0.80	10	9
gpb	0.81	0.84	0.76	3	3
gr	0.78	0.72	0.71	2	3
hivpr	0.64	0.77	0.71	9	10
hivrt	0.53	0.56	0.50	4	4
hmga	0.71	0.87	0.72	7	8
hsp90	0.89	0.87	0.80	4	5
inha	0.52	0.56	0.57	4	4
mr	0.84	0.78	0.80	2	3
na	0.91	0.96	0.84	6	6
p38	0.76	0.85	0.73	4	4
parp	0.83	0.86	0.82	1	2
pde5	0.75	0.74	0.72	6	6
pdgfrb	0.65	0.69	0.66	4	5
pnp	0.88	0.94	0.85	2	3
ppar_gamma	0.78	0.91	0.81	11	10
pr	0.87	0.84	0.82	1	3
rxr_alpha	0.93	0.95	0.90	4	6
sahh	0.94	0.94	0.88	2	3
src	0.71	0.70	0.67	5	6
thrombin	0.75	0.76	0.69	7	9
tk	0.75	0.79	0.68	2	3
trypsin	0.77	0.82	0.73	8	9
vegfr2	0.58	0.57	0.58	5	6

[1] mconf = multiple conformers

[2] sconf = single conformers

[3] mean rotatable bond count for known ligands

[4] mean rotatable bond count for decoys

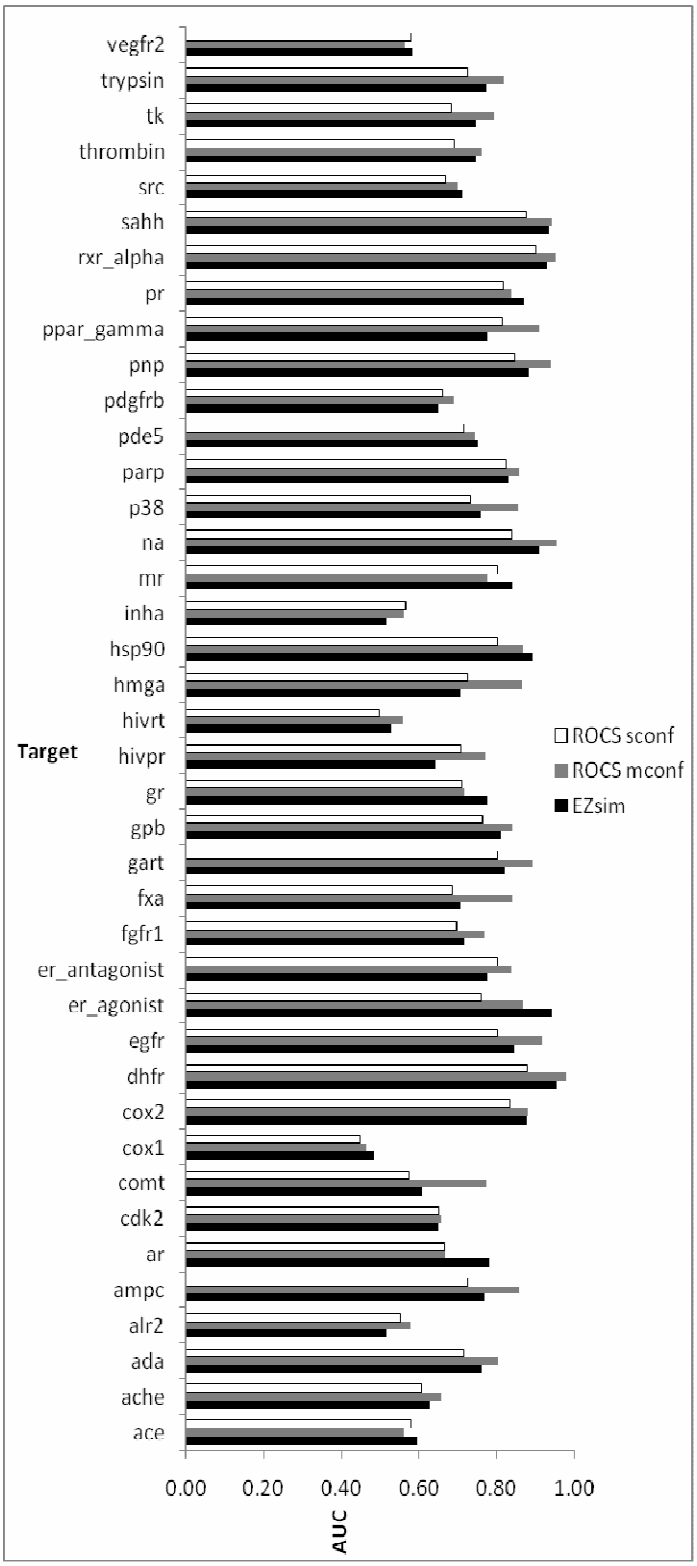
vegfr2
trypsin
tk
thrombin
src
sahh
rxr_alpha
pr
ppar_gamma
pnp
pdgfrb
pde5
parp
p38
na
mr
inha
hsp90
hmga
hivrt
hivpr
gr
gpb
gart
fxa
fgfr1
er_antagonist
er_agonist
egfr
dhfr
cox2
cox1
comt
cdk2
ar
ampc
alr2
ada
ache
ace
Target
ROCS sconf
ROCS mconf
EZsim
0.00
0.20
0.40
0.60
0.80
1.00
AUC

Figure S3