JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field

Lingle Wang,[†] Yujie Wu,[†] Yuqing Deng,[†] Byungchan Kim,[†] Levi Pierce,[†] Goran Krilov,[†] Dmitry Lupyan,[†] Shaughnessy Robinson,[†] Markus K. Dahlgren,[†] Jeremy Greenwood,[†] Donna L. Romero,[‡] Craig Masse,[‡] Jennifer L. Knight,[†] Thomas Steinbrecher,[†] Thijs Beuming,[†] Wolfgang Damm,[†] Ed Harder,[†] Woody Sherman,[†] Mark Brewer,[†] Ron Wester,[‡] Mark Murcko,[†] Leah Frye,[†] Ramy Farid,[†] Teng Lin,[†] David L. Mobley,[⊥] William L. Jorgensen,[∥] Bruce J. Berne,[§] Richard A. Friesner,[§] and Robert Abel*,[†]

Bentley Wingert
7 June 2016

Outline

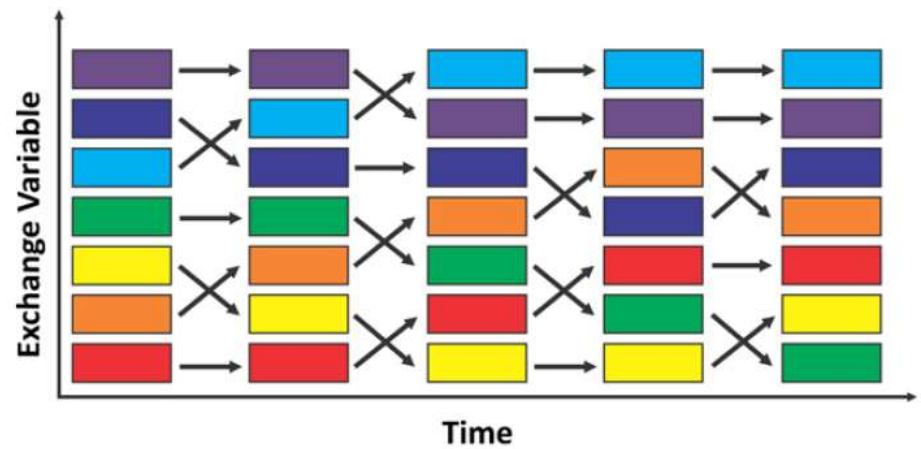
- Background
- Method
- Results
- Discussion

Background

- Free Energy Perturbation (FEP)
  - Measure difference in free energy between states by analyzing MD simulations
- Replica Exchange with Solute Tempering (REST)
  - Enhanced sampling technique
- Goal is to provide unified pipeline to allow more efficient lead compound analysis for drug design
- Do this by developing better force field and parallelization (GPUs)

$$\Delta F(A \to B) = F_B - F_A = -k_B T \ln \left\langle \exp\left(-\frac{E_B - E_A}{k_B T}\right)\right\rangle_A$$

Zwanzig, Robert W. "High-temperature equation of state by a perturbation method. I. nonpolar gases." *The Journal of Chemical Physics* 22.8 (1954): 1420-1426.



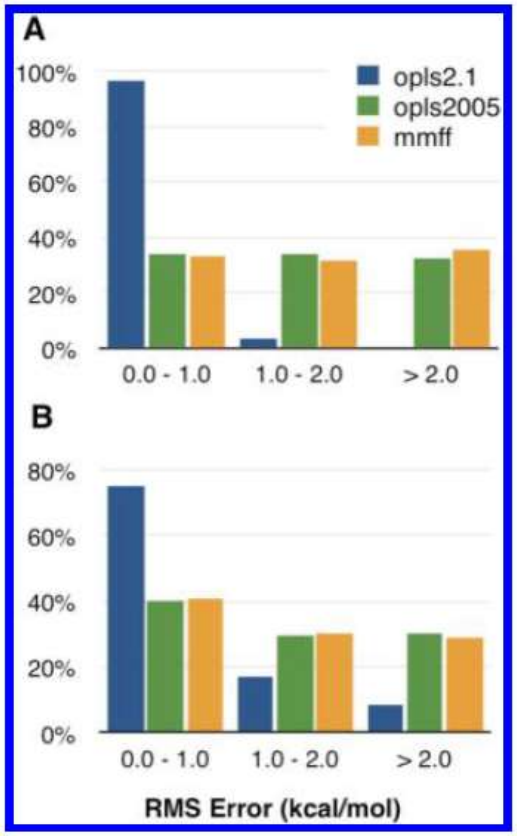
Bernardi, Rafael C., Marcelo CR Melo, and Klaus Schulten. "Enhanced sampling techniques in molecular dynamics simulations of biological systems." *Biochimica et Biophysica Acta (BBA)-General Subjects* 1850.5 (2015): 872-877.

Improved Force Field (OPLS2.1)

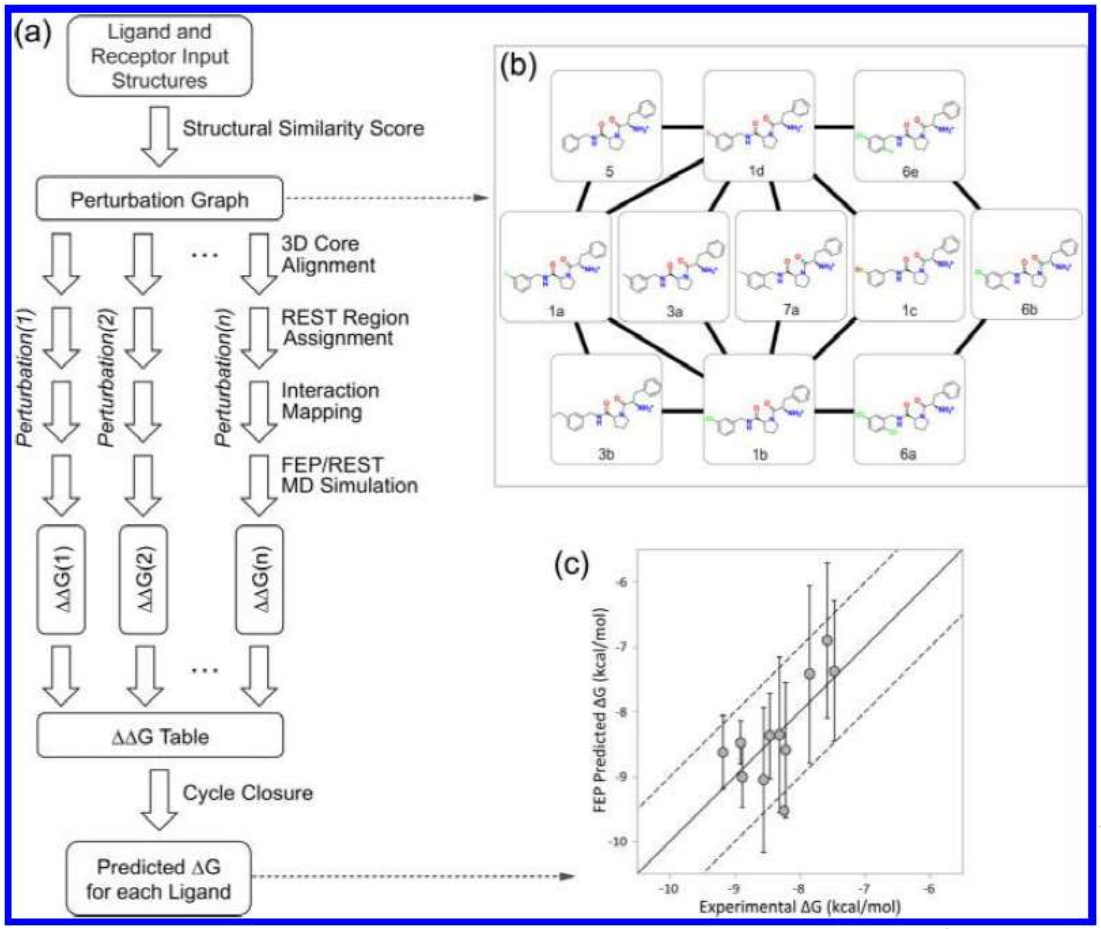
- Improve force field performance for small molecules
- Built on previous OPLS force field
- Automated generation of parameters for molecules not similar to training library
- GPU implementation

Table 1. Error Statistics for Solvation Free-Energy Results[a]

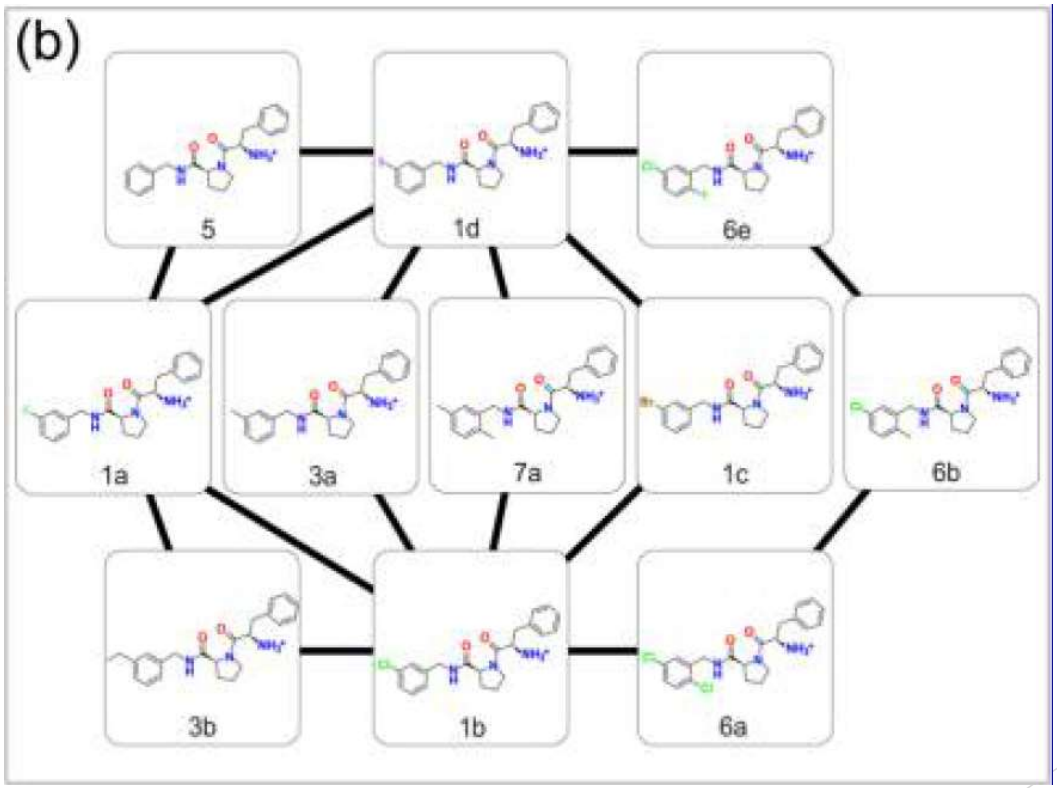
force field	MUE (kcal/mol)	RMSE (kcal/mol)	% > 2 kcal/mol error
ChelpG/CharmM	1.93	2.28	44.7
AM1-BCC/ GAFF	1.17	1.39	15.0
OPLS 2005	1.10	1.33	8.5
OPLS2.1	0.73	0.88	2.1



Method


(a)
Ligand and Receptor Input Structures
Structural Similarity Score
Perturbation Graph
3D Core Alignment
REST Region Assignment
Interaction Mapping
FEP/REST MD Simulation
Perturbation(1)
Perturbation(2)
Perturbation(n)
ΔΔG(1)
ΔΔG(2)
ΔΔG(n)
ΔΔG Table
Cycle Closure
Predicted ΔG for each Ligand
(b)
5
1d
6e
1a
3a
7a
1c
6b
3b
1b
6a
(c)
FEP Predicted ΔG (kcal/mol)
Experimental ΔG (kcal/mol)

Method



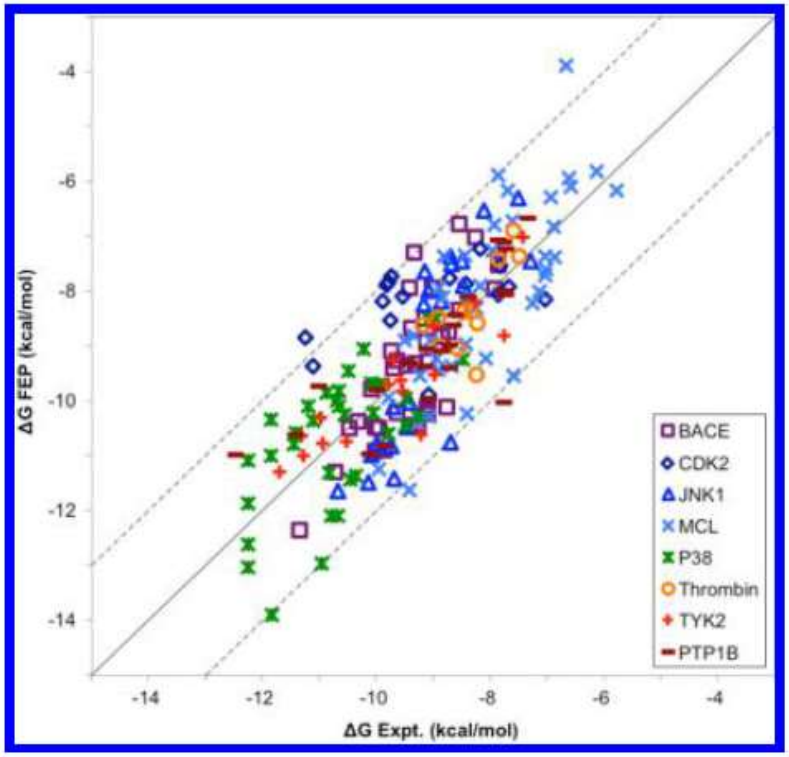
Results

	system							
	BACE	CDK2	JNK1	MCL1	p38	PTP1B	thrombin	Tyk2
no. of compds	36	16	21	42	34	23	11	16
binding affinity range (kcal/mol)	3.5	4.2	3.4	4.2	3.8	5.1	1.7	4.3
crystal structure	4DJW	1H1Q	2GMX	4HW3	3FLY	2QBS	2ZFF	4GIH
series ref	46	47	48	49	50	51	45	52,53
no. of perturbations	58	25	31	71	56	49	16	24
MUE FEP	0.84 ± 0.08	0.91 ± 0.12	0.78 ± 0.12	1.16 ± 0.10	0.80 ± 0.08	0.89 ± 0.12	0.76 ± 0.13	0.75 ± 0.11
RMSE FEP	1.03 ± 0.08	1.11 ± 0.12	1.00 ± 0.15	1.41 ± 0.12	1.03 ± 0.09	1.22 ± 0.17	0.93 ± 0.15	0.93 ± 0.12
avg σ_{cc}	0.65	0.57	0.30	0.91	0.76	0.94	0.93	0.46
obs *R*-value FEP	0.78 ± 0.07	0.48 ± 0.19	0.85 ± 0.07	0.77 ± 0.05	0.65 ± 0.09	0.80 ± 0.08	0.71 ± 0.24	0.89 ± 0.07
P-value FEP	3.9×10^{-5}	1.2×10^{-2}	7.0×10^{-8}	2.2×10^{-7}	1.6×10^{-7}	7.8×10^{-6}	1.1×10^{-2}	2.3×10^{-7}
obs *R*-value, MW	0.14	−0.48	−0.39	−0.55	−0.46	−0.84	−0.48	0.00
obs *R*-value, MM-GB/SA	−0.40	−0.53	0.65	0.42	0.66	0.67	0.93	0.79
obs *R*-value, Glide SP	0.00	−0.56	0.24	0.59	0.14	0.55	0.53	0.79

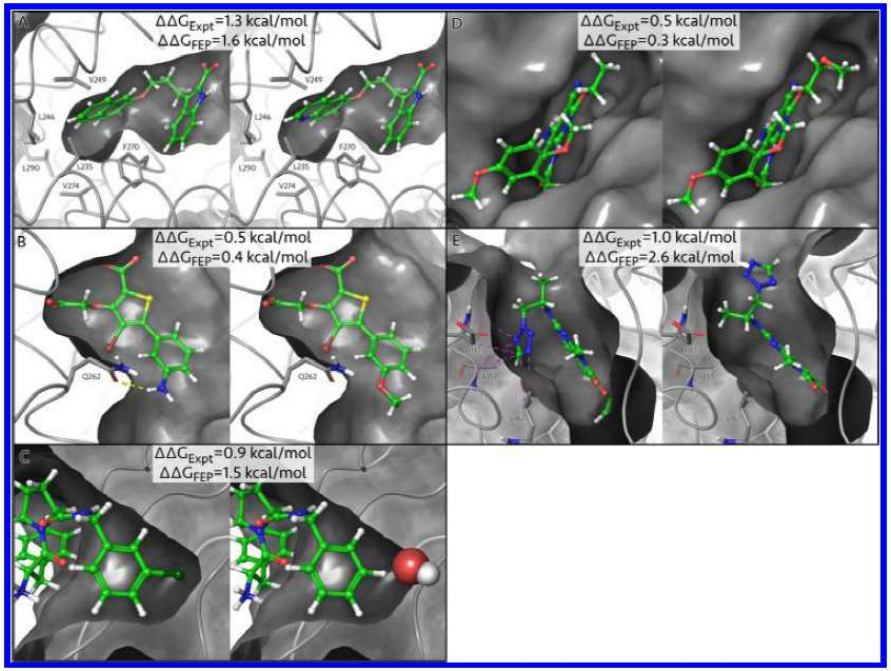
Results

	system							
	BACE	CDK2	JNK1	MCL1	p38	PTP1B	thrombin	Tyk2
no. of compds	36	16	21	42	34	23	11	16
binding affinity range (kcal/mol)	3.5	4.2	3.4	4.2	3.8	5.1	1.7	4.3
crystal structure	4DJW	1H1Q	2GMX	4HW3	3FLY	2QBS	2ZFF	4GIH
series ref	46	47	48	49	50	51	45	52,53
no. of perturbations	58	25	31	71	56	49	16	24
MUE FEP	0.84 ± 0.08	0.91 ± 0.12	0.78 ± 0.12	1.16 ± 0.10	0.80 ± 0.08	0.89 ± 0.12	0.76 ± 0.13	0.75 ± 0.11
RMSE FEP	1.03 ± 0.08	1.11 ± 0.12	1.00 ± 0.15	1.41 ± 0.12	1.03 ± 0.09	1.22 ± 0.17	0.93 ± 0.15	0.93 ± 0.12
avg σ_{cc}	0.65	0.57	0.30	0.91	0.76	0.94	0.93	0.46
obs *R*-value FEP	0.78 ± 0.07	0.48 ± 0.19	0.85 ± 0.07	0.77 ± 0.05	0.65 ± 0.09	0.80 ± 0.08	0.71 ± 0.24	0.89 ± 0.07
P-value FEP	3.9×10^{-5}	1.2×10^{-2}	7.0×10^{-8}	2.2×10^{-7}	1.6×10^{-7}	7.8×10^{-6}	1.1×10^{-2}	2.3×10^{-7}
obs *R*-value, MW	0.14	−0.48	−0.39	−0.55	−0.46	−0.84	−0.48	0.00
obs *R*-value, MM-GB/SA	−0.40	−0.53	0.65	0.42	0.66	0.67	0.93	0.79
obs *R*-value, Glide SP	0.00	−0.56	0.24	0.59	0.14	0.55	0.53	0.79

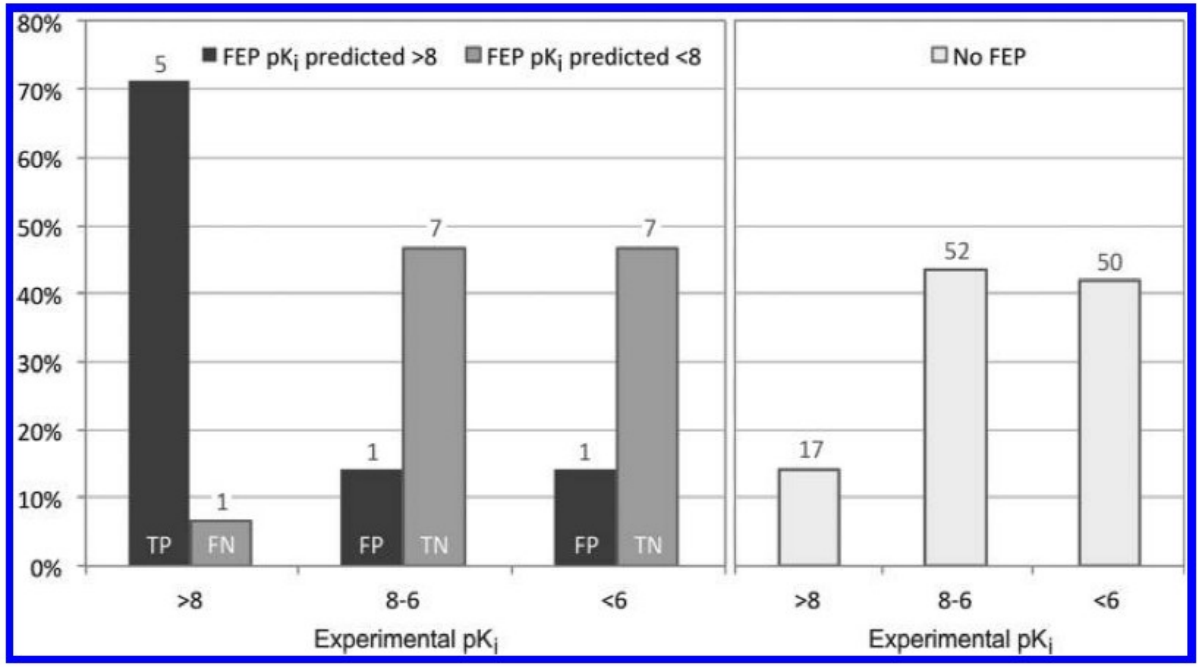
Correlation with experimental results



Variety of Interactions Captured



Prediction Accuracy



Discussion

- Results seem promising
- Only tested on small number of receptors, don't know how diverse actual test compounds were
- Applied method to eight active projects, only give results from 1.5
- Not much actually new, just put together previous work in coherent package and mostly automated workflow
- Still slow, "four perturbations per day…eight graphics cards"