Gibbs Sampler Based λ-Dynamics and Rao-Blackwell Estimator for Alchemical Free Energy Calculation Supporting Information

Xinqiang Ding,[†] Jonah Z. Vilseck,[‡] Ryan L. Hayes,[‡] and Charles L. Brooks III[*,†,‡,¶]

†*Department of Computational Medicine & Bioinformatics, University of Michigan*

‡*Department of Chemistry, University of Michigan*

¶*Biophysics Program, University of Michigan, Ann Arbor, Michigan 48109, United States*

E-mail: brookscl@umich.edu

Table S1. Scaling Factors Used for Electrostatic Interaction and Lennard-Jones Interaction for Discrete λ in Pairwise GSLD Simulation With T4L Lysozyme

λ	l_1	l_2	l_3	l_4	l_5	l_6	l_7	l_8	l_9	l_{10}	l_{11}	l_{12}
benzene λ_{elec}	1.00	0.66	0.33	0	0	0	0	0	0	0	0	0
benzene λ_{LJ}	1	1	1	1	0.9	0.8	0.7	0.55	0.45	0.3	0.2	0.1
p-xylene λ_{elec}	0	0	0	0	0	0	0	0	0	0	0	0
p-xylene λ_{LJ}	0	0	0	0	0.1	0.2	0.3	0.45	0.55	0.7	0.8	0.9

λ	l_{13}	l_{14}	l_{15}	l_{16}
benzene λ_{elec}	0	0	0	0
benzene λ_{LJ}	0	0	0	0
p-xylene λ_{elec}	0	0.33	0.66	1
p-xylene λ_{LJ}	1	1	1	1

The formula used for the λ-dependent nonbonded interaction energy is:

$$V_{\text{nb}}(\lambda_{\text{elec}}, \lambda_{\text{LJ}}) = \lambda_{\text{elec}} \frac{q_i q_j}{r_{ij}} + \lambda_{\text{LJ}} 4\epsilon_{ij} \left(\frac{1}{[\alpha_{\text{LJ}}(1-\lambda_{\text{LJ}}) + (r_{ij}/\sigma_{ij})^6]^2} - \frac{1}{\alpha_{\text{LJ}}(1-\lambda_{\text{LJ}}) + (r_{ij}/\sigma_{ij})^6} \right)$$

Table S2. Free Energy Differences (ΔG in kcal/mol) for The Harmonic System Calculated Using Rao-Blackwell Estimator and Empirical Estimator with Cutoff of 0.9 and 0.99 with 10 ns of Sampling.

model system	$\Delta G_{analytic}$	GSLD/Rao-Blackwellization ΔG	Empirical(0.90) ΔG	Empirical(0.99) ΔG
symmetrical	0.00	0.00±0.03	0.00±0.03	-0.00±0.07
asymmetrical	-0.56	-0.56±0.02	-0.41±0.03	-0.50±0.06

Table S3. Free Energy Estimation (kcal/mol) for The Three Benzene Derivatives Sampled Using Pairwise GSLD and Calculated Using RBE and Empirical Estimators with Different Cutoffs in Vacuum and Water.

substituent	ΔG_{vac} (2 ns)			ΔG_{solv} (3 ns)		
change	RBE	Empirical(0.90)	Empirical(0.99)	RBE	Empirical(0.90)	Empirical(0.99)
benzene → phenol	-10.27±0.01	-10.34±0.03	-10.30±0.06	-14.73±0.07	-14.96±0.06	-14.87±0.08
benzene → benzaldehyde	6.15±0.02	6.18±0.04	6.18±0.09	3.04±0.09	2.20±0.08	2.60±0.08
phenol → benzaldehyde	16.42±0.01	16.50±0.03	16.52±0.08	17.81±0.16	16.99±0.09	17.43±0.11

Table S4. Biasing Potential Energies (kcal/mol) Used for The Three Benzene Derivatives Sampled in Pairwise GSLD in Vacuum and Water.

	benzene → phenol	benzene → benzaldehyde	phenol → benzaldehyde
vacuum	10.23, 10.41, 10.66, 9.95, 10.37, 10.40, 10.51, 10.04, 10.00, 10.27	-6.17, -6.00, -6.06, -5.99, -6.15, -6.32, -6.04, -6.22, -6.33, -6.11	-16.24, -16.32, -16.50, -16.28, -16.46, -15.98, -16.33, -16.32, -16.11, -16.45
water	14.82, 14.94, 14.99, 14.58, 14.99, 14.66, 14.75, 15.08, 14.78, 14.66	-2.34, -1.79, -1.97, -2.36, -2.01, -2.03, -2.34, -2.26, -2.45, -2.20	-16.99, -16.92, -17.21, -16.29, -16.51, -16.90, -16.63, -17.02, -17.07, -16.94

Table S5. Free Energy (kcal/mol) Estimation for The Three Benzene Derivatives Using FEP/MBAR in Vacuum and Water.

substituent change	ΔG_{vac} (2.2 ns)	ΔG_{solv} (11 ns)
benzene → phenol	-10.28±0.01	-14.74±0.02
benzene → benzaldehyde	6.14±0.01	3.01±0.02
phenol → benzaldehyde	16.42±0.01	17.76±0.13

Table S6. Free Energy (kcal/mol) Estimation for The Three Benzene Derivatives Sampled Using GSLD for Multiple Ligands and Calculated Using RBE and Empirical Estimators with Different Cutoffs in Vacuum and Water.

substituent	ΔG_{vac} (2 ns)			ΔG_{solv} (3 ns)		
change	RBE	Empirical(0.90)	Empirical(0.95)	RBE	Empirical(0.90)	Empirical(0.99)
benzene → phenol	-10.28±0.01	-10.29±0.05	-10.32±0.16	-14.81±0.14	-14.70±0.07	-14.79±0.14
benzene → benzaldehyde	6.14±0.02	6.14±0.05	6.12±0.14	2.92±0.10	2.32±0.08	2.68±0.07
phenol → benzaldehyde	16.42±0.01	16.43±0.04	16.44±0.15	17.73±0.09	17.02±0.06	17.47±0.14

Table S7. Biasing Potential Energies (kcal/mol) Used for The Three Benzene Derivatives Sampled in GSLD for Multiple Ligands in Vacuum and Water.

	$(G^b_{\text{benzene}}, G^b_{\text{phenol}}, G^b_{\text{benzaldehyde}})$
vacuum	(-1.26, 8.84, -7.58), (-1.35, 8.93, -7.59), (-1.41, 8.99, -7.58), (-1.33, 8.93, -7.61), (-1.41, 8.91, -7.50), (-1.37, 8.89, -7.52), (-1.19, 8.85, -7.66), (-1.37, 9.03, -7.66), (-1.42, 8.92, -7.50), (-1.41, 9.03, -7.62)
water	(-4.40, 10.86, -6.47), (-3.96, 10.57, -6.61), (-4.01, 10.51, -6.51), (-4.33, 10.55, -6.22), (-4.16, 10.61, -6.45), (-4.10, 10.64, -6.54), (-3.92, 10.56, -6.64), (-3.90, 10.60, -6.70), (-4.23, 10.69, -6.47), (-4.11, 10.57, -6.46)

Table S8. Biasing Potential Energies (kcal/mol) Used for The Alchemical Change: benzene to p-xylene in Water.

	benzene → p-xylene
water	10.20, 11.00, 10.91, 11.44, 11.08, 10.79, 11.22, 11.05, 10.90, 11.51

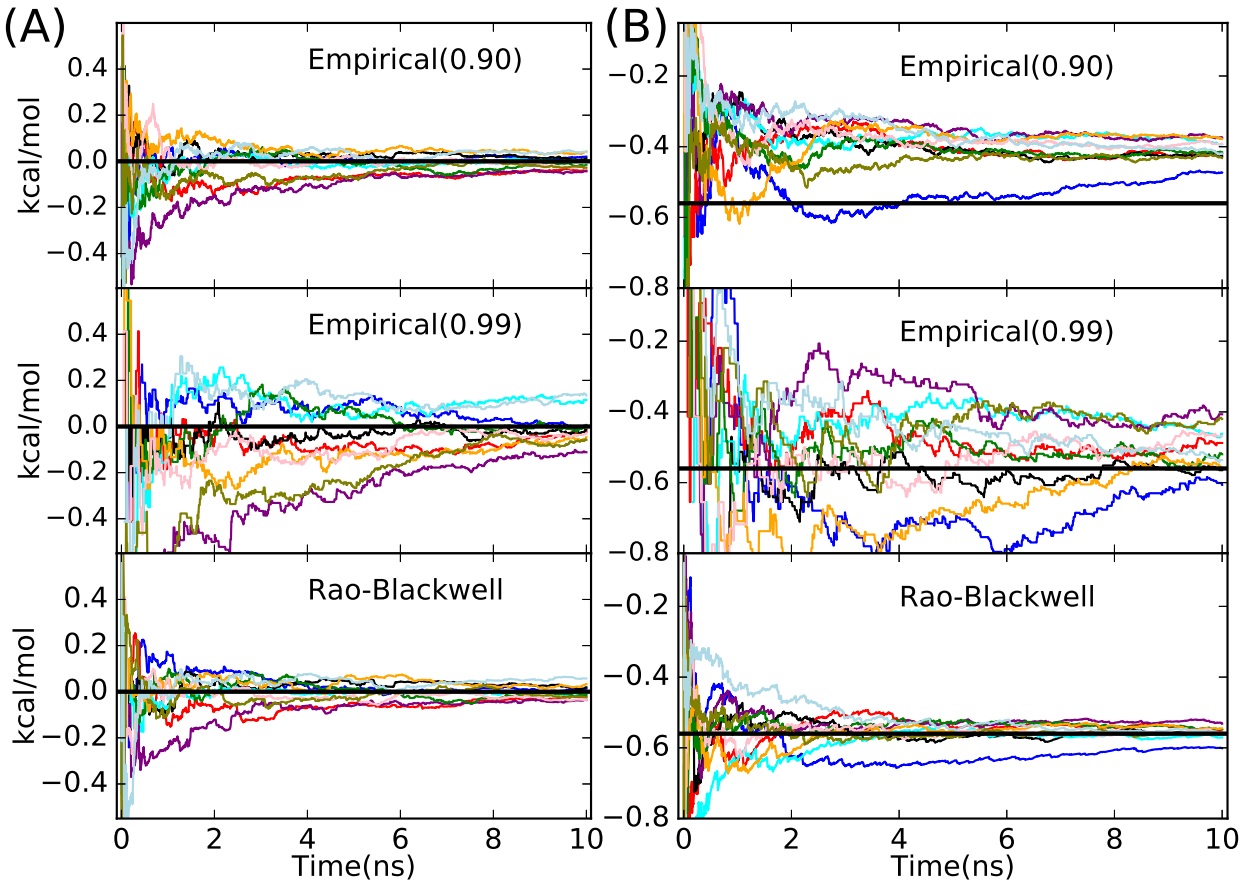

Figure S1. **(A)** Estimated free energies from 10 individual repeat for the symmetrical harmonic system using empirical estimators with a cutoff of 0.9 (up) and 0.99 (middle) and the Rao-Blackwell estimator (bottom). **(B)** Estimated free energies from each repeat for the asymmetrical harmonic system using empirical estimators with a cutoff of 0.9 (up) and 0.99 (middle) and the Rao-Blackwell estimator (bottom). The horizontal black line is the calculated free energy changes using numerical integration.

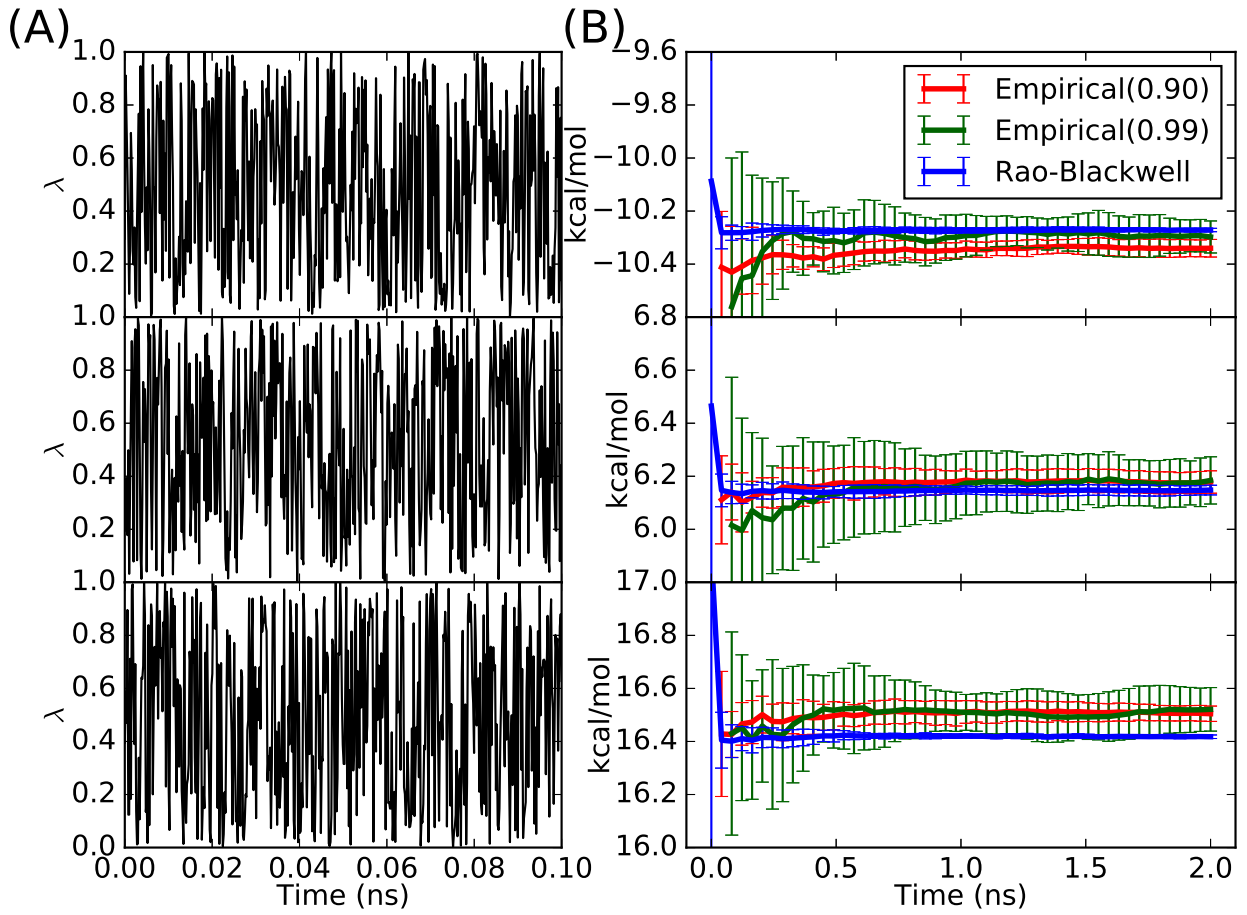

Figure S2. **(A)** λ trajectories from the simulation in vacuum using pairwise GSLD for alchemical changes benzene to phenol (top), benzene to benzaldehyde (middle), and phenol to benzaldehyde (bottom). **(B)** Estimated alchemical free energy changes in vacuum using empirical estimators with different cutoff values and Rao-Blackwell estimator for alchemical changes benzene to phenol (top), benzene to benzaldehyde (middle), and phenol to benzaldehyde (bottom).

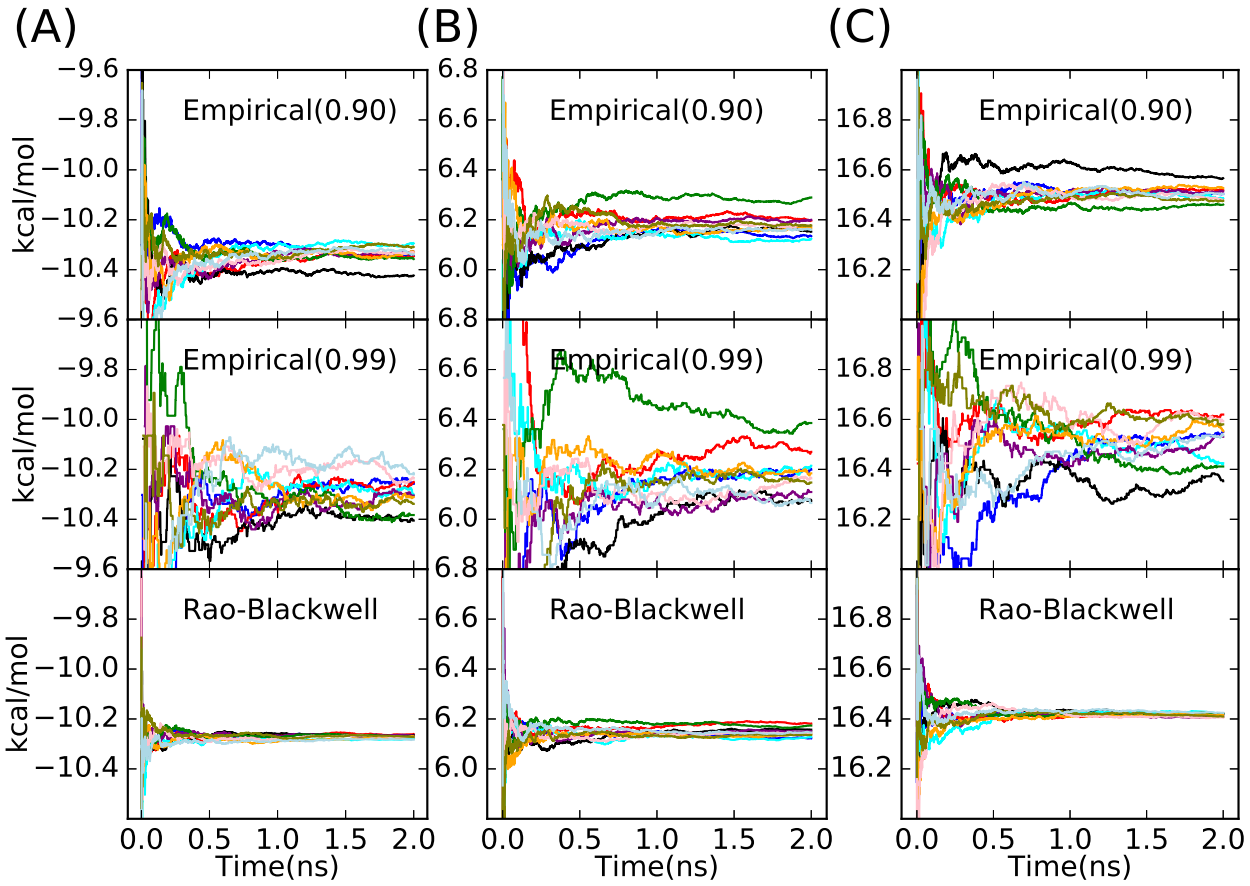

Figure S3. **(A)** Estimated alchemical free energy changes in vaccum from 10 individual repeat for alchemical changes benzene to phenol using empirical estimators with a cutoff of 0.9 (up) and 0.99 (middle) and the Rao-Blackwell estimator (bottom). **(B)** and **(C)** show the same data for alchemial change benzene to benzaldehyde and alchemical change phenol to benzaldehyde, respectively.

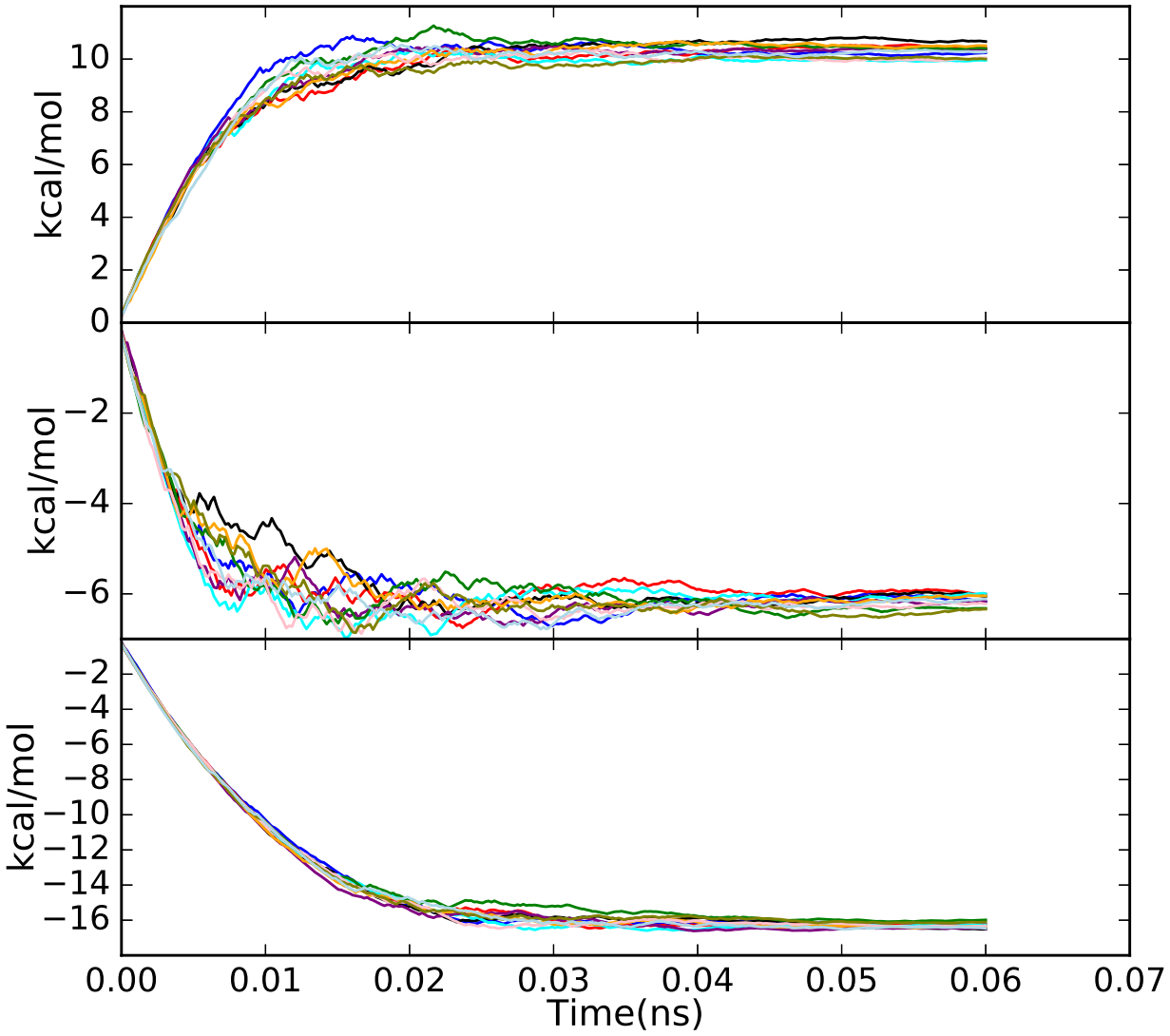

Figure S4. Biasing potential energies calculated using the Wang-Landau like algorithm for the alchemical changes: benzene to phenol (top), benzene to benzaldehyde (middle) and phenol to benzaldehyde (bottom) in vacuum.

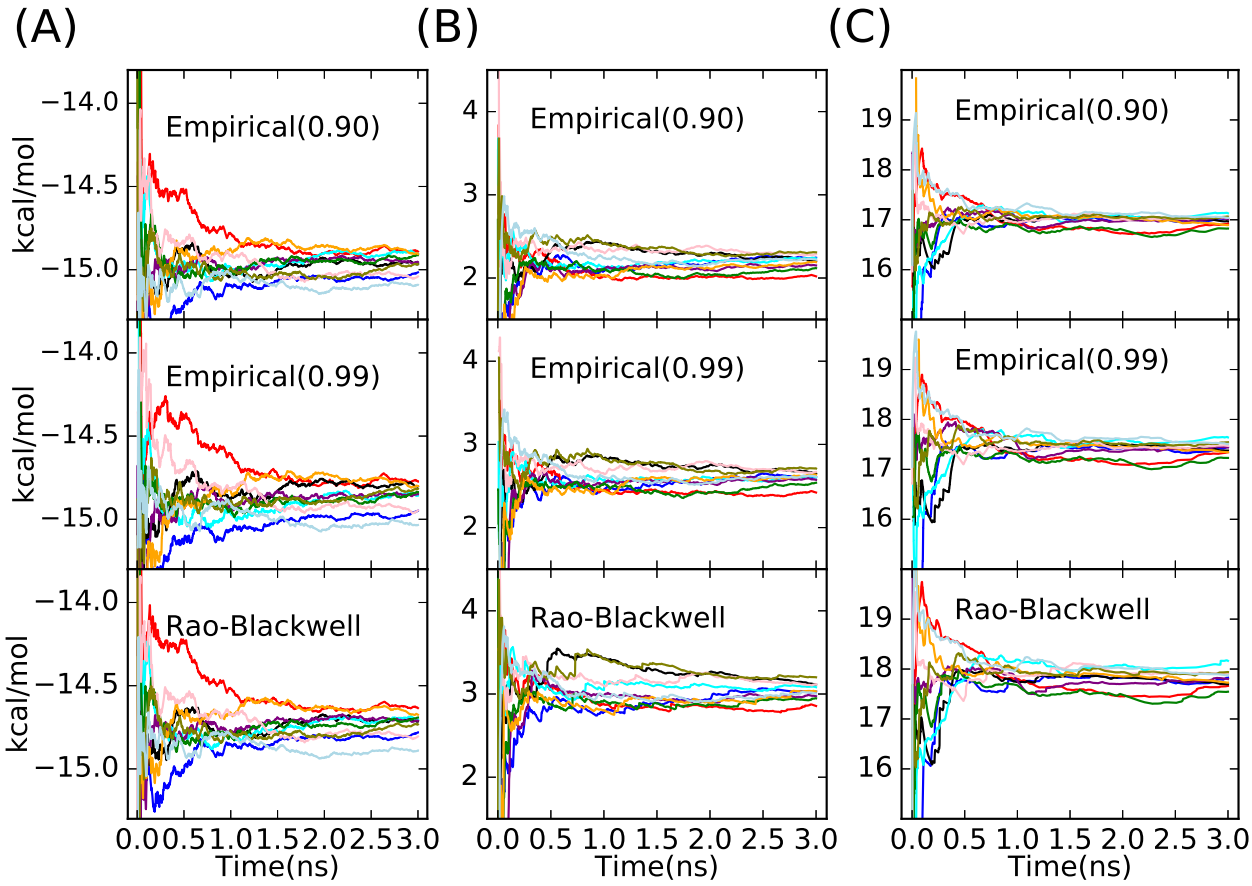

Figure S5. **(A)** Estimated alchemical free energy changes in water from 10 individual repeat for alchemical changes benzene to phenol using empirical estimators with a cutoff of 0.9 (up) and 0.99 (middle) and the Rao-Blackwell estimator (bottom). **(B)** and **(C)** show the same data for alchemial change benzene to benzaldehyde and alchemical change phenol to benzaldehyde, respectively.

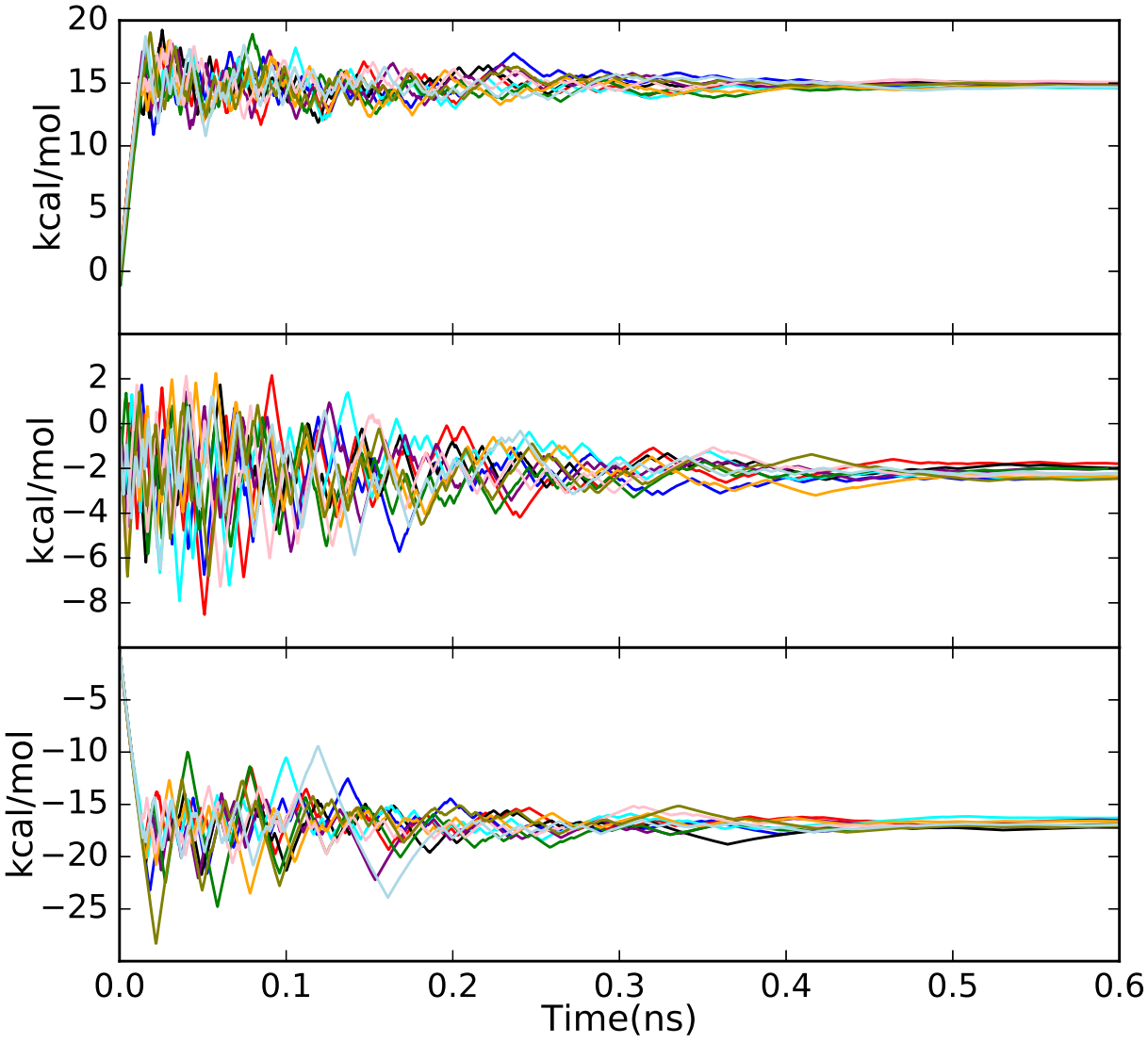

Figure S6. Biasing potential energies calculated using the Wang-Landau like algorithm for the alchemical changes: benzene to phenol (top), benzene to benzaldehyde (middle) and phenol to benzaldehyde (bottom) in water.

Table S9. Biasing Potential Energies (kcal/mol) Used for The Alchemical Change: benzene to p-xylene in T4 Lysozyme.

	benzene → p-xylene
$\chi = -180°$	(0.00, 0.28, 0.61, 0.75, 1.22, 1.36, 0.95, -0.39, -1.48, -2.56, -3.08, -3.51, -3.61, 0.21, 4.24, 8.18) (0.00, 0.12, 0.35, 0.50, 0.92, 0.85, 0.33, -1.37, -2.09, -2.31, -2.02, -2.15, -1.78, 1.99, 6.00, 9.75) (0.00, 0.19, 0.47, 0.70, 1.11, 1.09, 0.63, -0.63, -1.52, -2.29, -2.63, -2.83, -2.96, 0.86, 4.94, 8.81) (0.00, 0.25, 0.33, 0.59, 0.98, 1.08, 0.67, -0.71, -1.87, -3.44, -4.17, -4.55, -4.79, -0.92, 3.05, 7.01) (0.00, 0.16, 0.39, 0.62, 0.90, 0.89, 0.62, -0.83, -1.89, -3.23, -3.88, -4.31, -4.42, -0.57, 3.53, 7.57) (0.00, 0.25, 0.39, 0.57, 0.98, 1.07, 0.70, -0.81, -1.63, -2.76, -3.22, -3.52, -3.68, 0.26, 4.22, 8.25) (0.00, 0.12, 0.30, 0.73, 1.04, 0.96, 0.80, -0.35, -0.48, -0.75, -0.75, -0.72, -0.59, 3.14, 7.05, 11.01) (0.00, 0.24, 0.45, 0.59, 1.12, 1.15, 0.77, -0.78, -1.90, -3.10, -3.60, -4.01, -4.44, -0.47, 3.52, 7.52) (0.00, 0.26, 0.43, 0.72, 1.12, 1.19, 0.88, -0.31, -1.38, -2.54, -2.98, -3.30, -3.57, 0.42, 4.29, 8.32) (0.00, 0.14, 0.43, 0.68, 0.91, 0.79, 0.57, -0.76, -1.53, -2.21, -2.41, -2.38, -2.42, 1.56, 5.59, 9.55)
$\chi = -60°$	(0.00, 0.28, 0.53, 0.73, 0.89, 1.00, 0.62, -0.36, -1.00, -1.97, -2.34, -2.64, -2.85, 1.02, 5.04, 8.91) (0.00, 0.37, 0.60, 0.80, 1.07, 0.95, 0.61, -0.45, -1.24, -2.05, -2.26, -2.35, -2.28, 1.49, 5.46, 9.34) (0.00, 0.26, 0.51, 0.84, 1.07, 1.01, 0.56, -0.13, -0.48, -0.86, -1.04, -1.00, -0.94, 2.99, 6.98, 11.04) (0.00, 0.21, 0.58, 0.89, 1.11, 0.98, 0.62, -0.40, -1.12, -1.85, -2.13, -2.20, -2.18, 1.57, 5.50, 9.36) (0.00, 0.25, 0.42, 0.67, 0.93, 0.88, 0.67, 0.04, -0.17, -0.32, -0.48, -0.55, -0.45, 3.44, 7.33, 11.34) (0.00, 0.29, 0.47, 0.88, 1.06, 1.17, 1.14, 0.82, 0.61, 0.35, 0.29, 0.20, 0.14, 4.01, 7.97, 11.81) (0.00, 0.24, 0.51, 0.89, 1.10, 1.09, 0.75, -0.42, -1.07, -1.94, -2.10, -2.20, -2.29, 1.69, 5.54, 9.51) (0.00, 0.23, 0.48, 0.78, 0.97, 1.09, 0.84, -0.18, -0.57, -0.95, -1.11, -1.33, -1.44, 2.54, 6.45, 10.35) (0.00, 0.21, 0.44, 0.81, 1.05, 1.01, 0.62, -0.73, -1.53, -2.27, -2.49, -2.55, -2.76, 1.30, 5.06, 9.10) (0.00, 0.30, 0.53, 0.76, 0.92, 1.04, 0.74, 0.08, -0.53, -0.73, -0.88, -0.93, -0.93, 2.93, 7.00, 10.89)

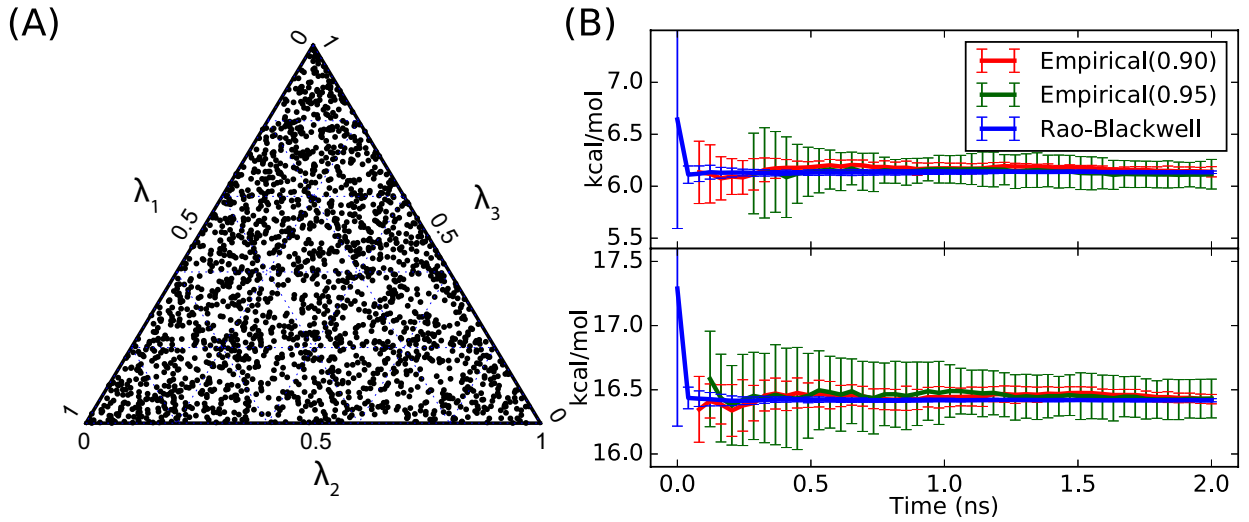

Figure S7. **(A)** $(\lambda_1, \lambda_2, \lambda_3)$ trajectory from the simulation in vacuum using GSLD for multiple ligands **(B)** Estimated alchemical free energy changes in vacuum using empirical estimator with different cutoff values and RBE for alchemical changes: benzene to benzaldehyde (top) and phenol to benzaldehyde (bottom).

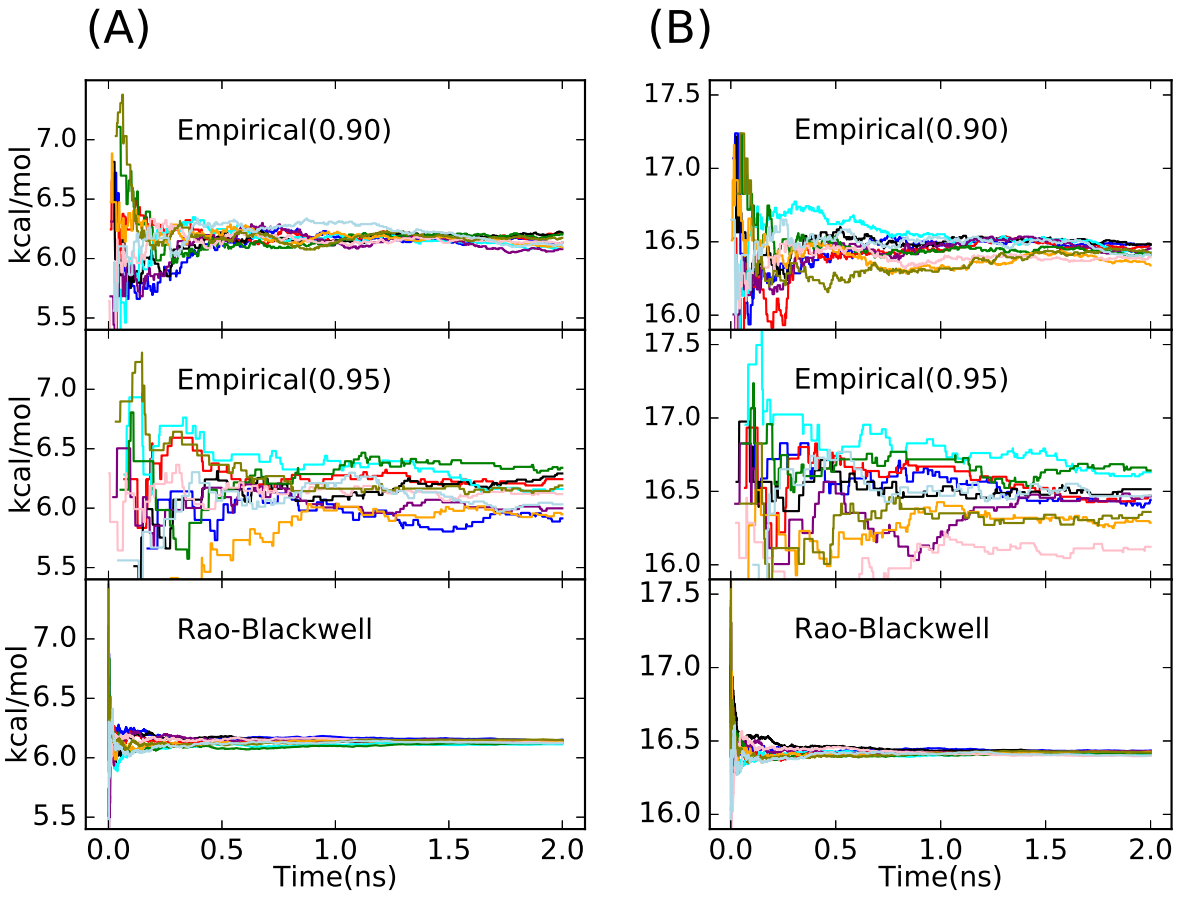

Figure S8. Estimated alchemical free energy changes in vacuum using GSLD for multiple ligands from 10 individual repeat using empirical estimators with a cutoff of 0.9 (up) and 0.99 (middle) and the Rao-Blackwell estimator (bottom) for alchemical changess: benzene to benzaldehyde **(A)** and phenol to benzaldehyde **(B)** .

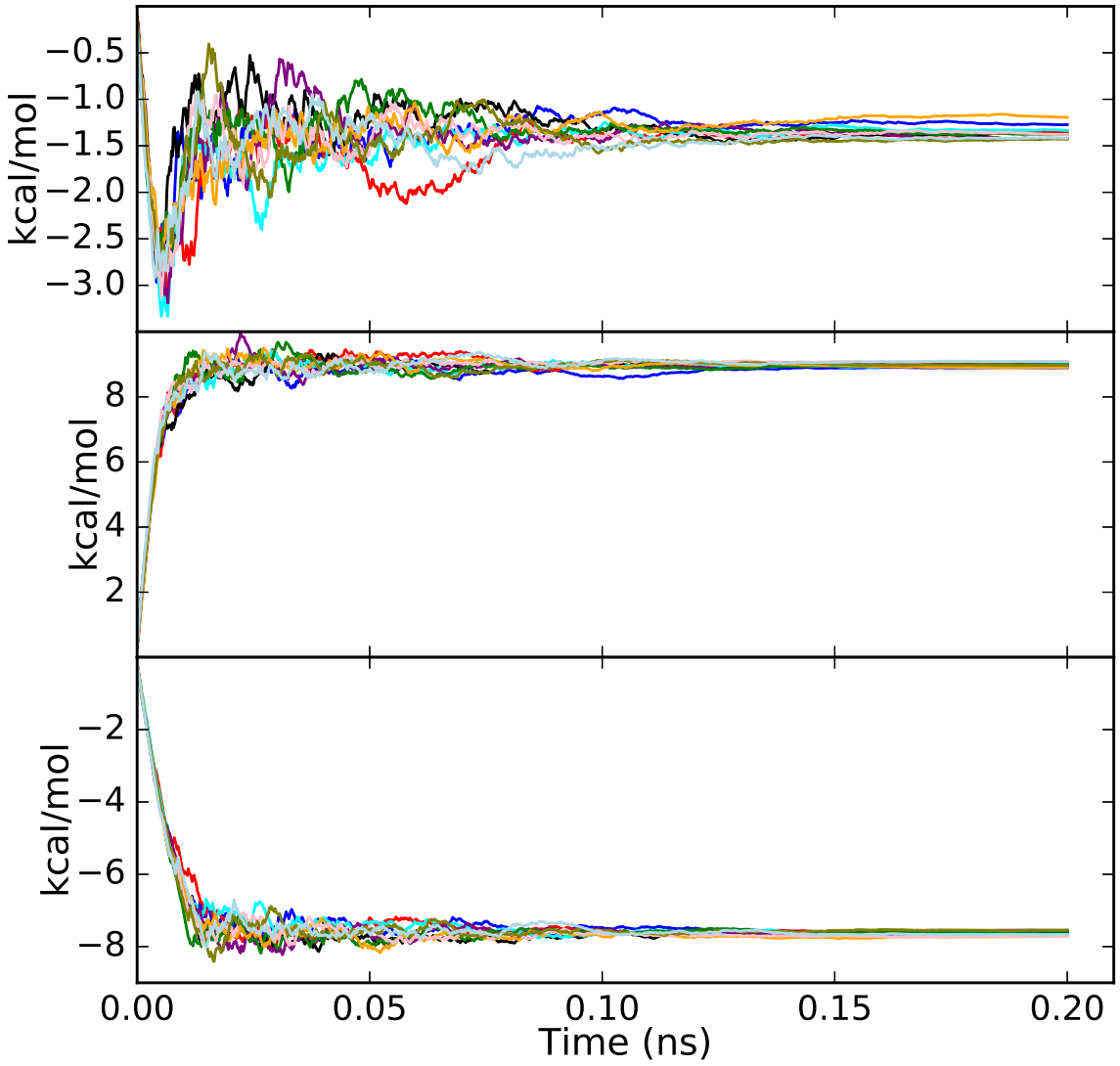

Figure S9. Biasing potential energies for benzene (top), phenol (middle) and benzaldehyde (bottom) calculated using the Wang-Landau like algorithm in GSLD in multiple ligands in vacuum. Each line represents the convergence of the biasing potential from one of 10 repreats.

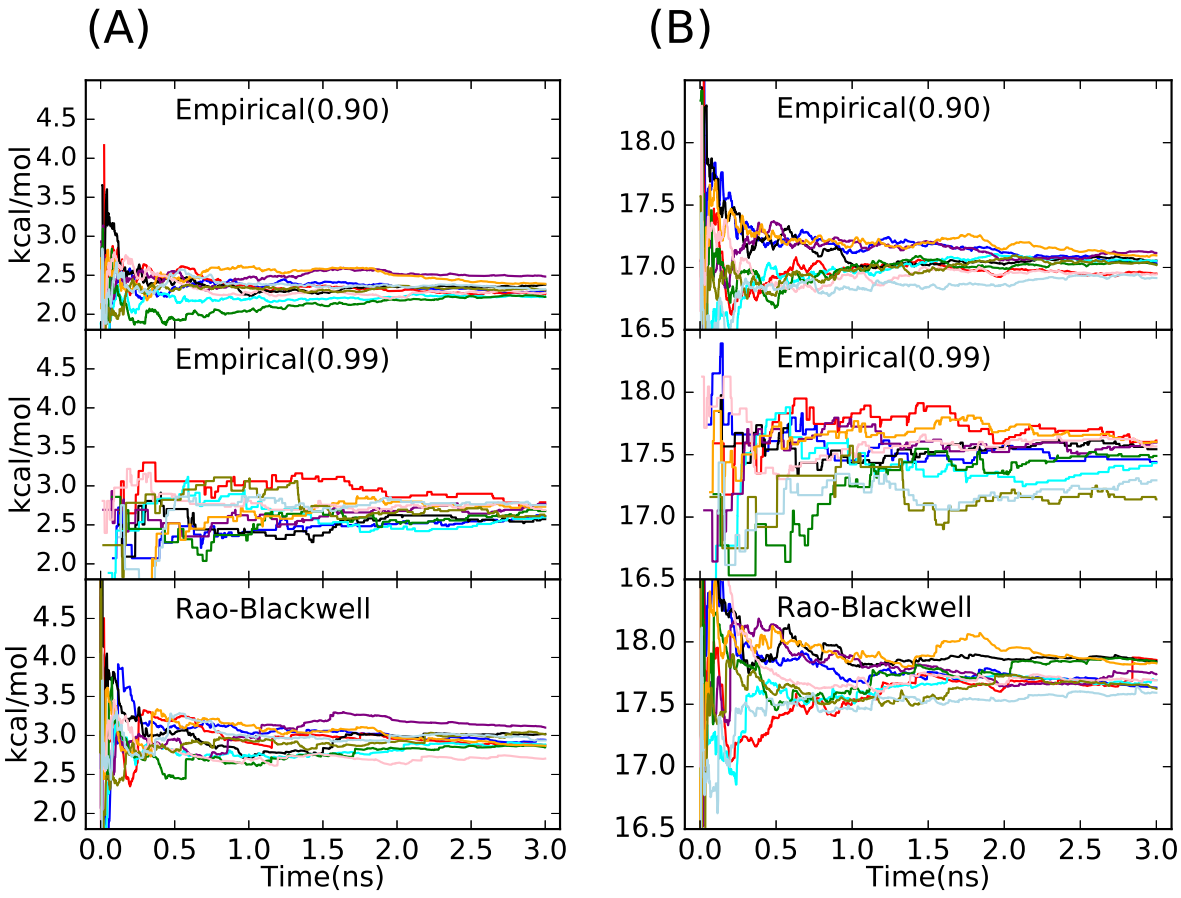

Figure S10. Estimated alchemical free energy changes in water using GSLD for multiple ligands from 10 individual repeat using empirical estimators with a cutoff of 0.9 (up) and 0.99 (middle) and the Rao-Blackwell estimator (bottom) for alchemical changess: **(A)** benzene to benzaldehyde and **(B)** phenol to benzaldehyde.

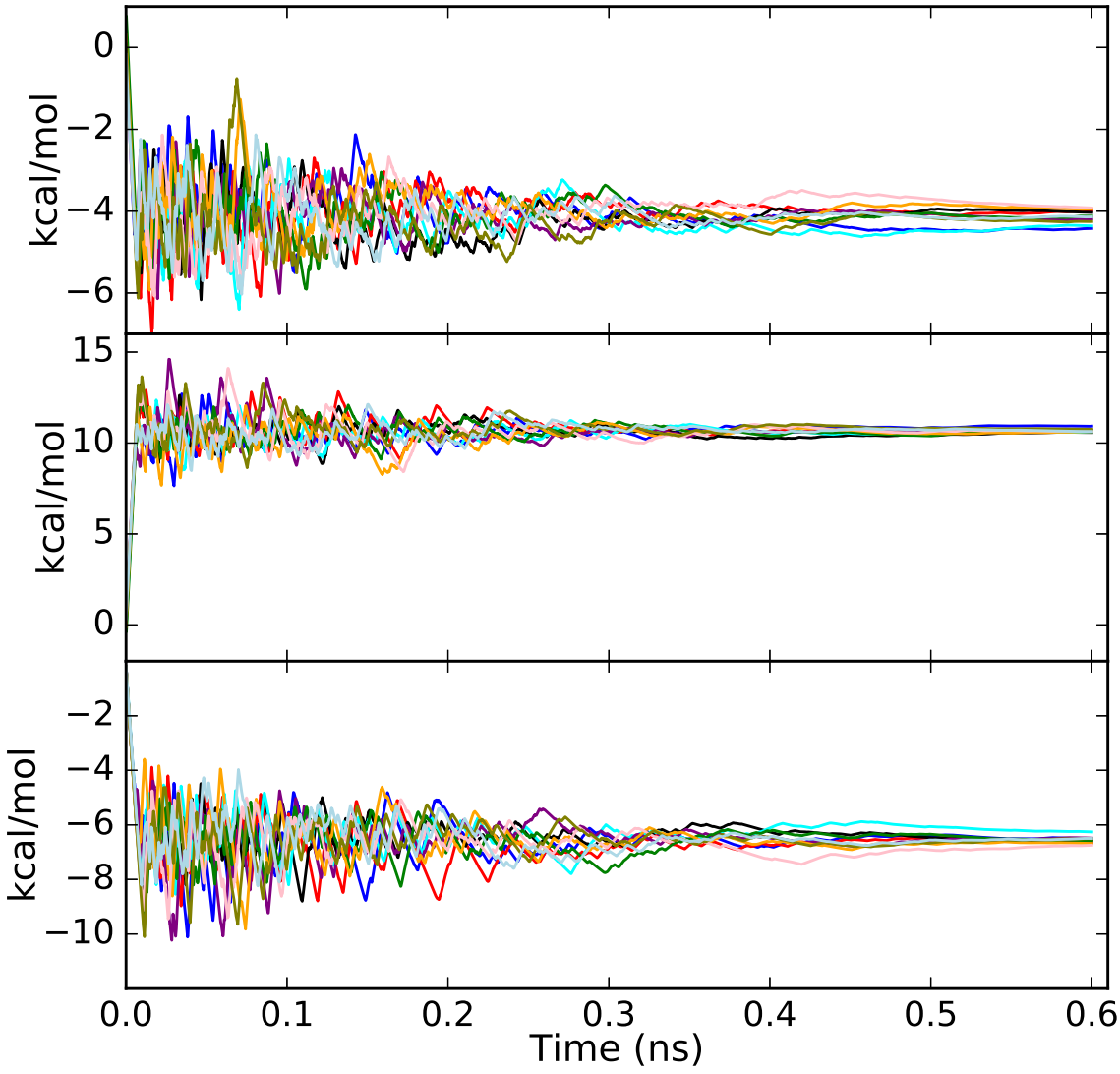

Figure S11. Biasing potential energies for benzene (top), phenol (middle) and benzaldehyde (bottom) calculated using the Wang-Landau like algorithm in GSLD in multiple ligands in water. Each line represents the convergence of the biasing potential from one of 10 repeats.

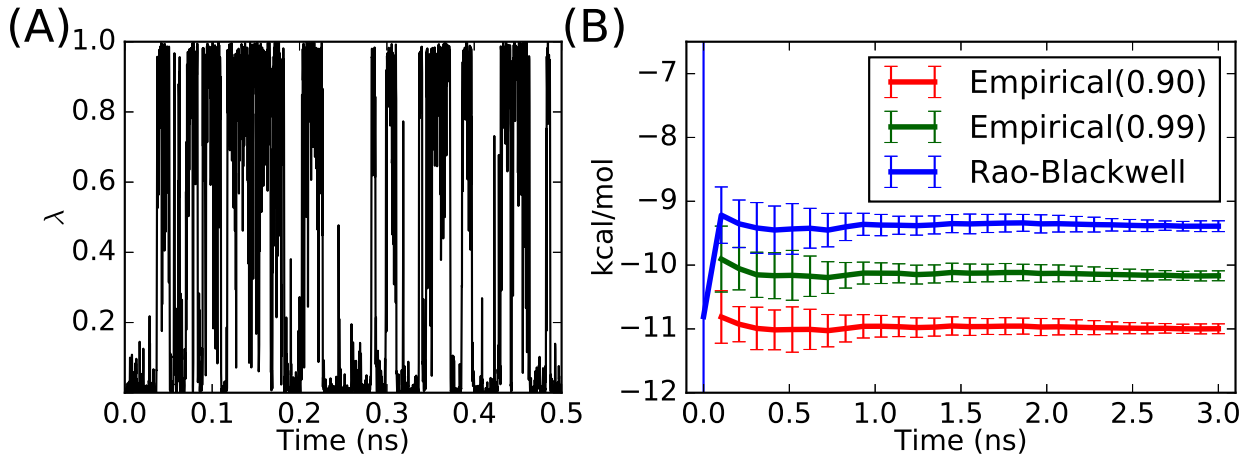

Figure S12. **(A)** λ trajectory of the simulation in water using GSLD for alchemical change from benzene to p-xylene. **(B)** Estimated alchemical free energy changes in water using RBE and empirical estimators with different cutoff values.

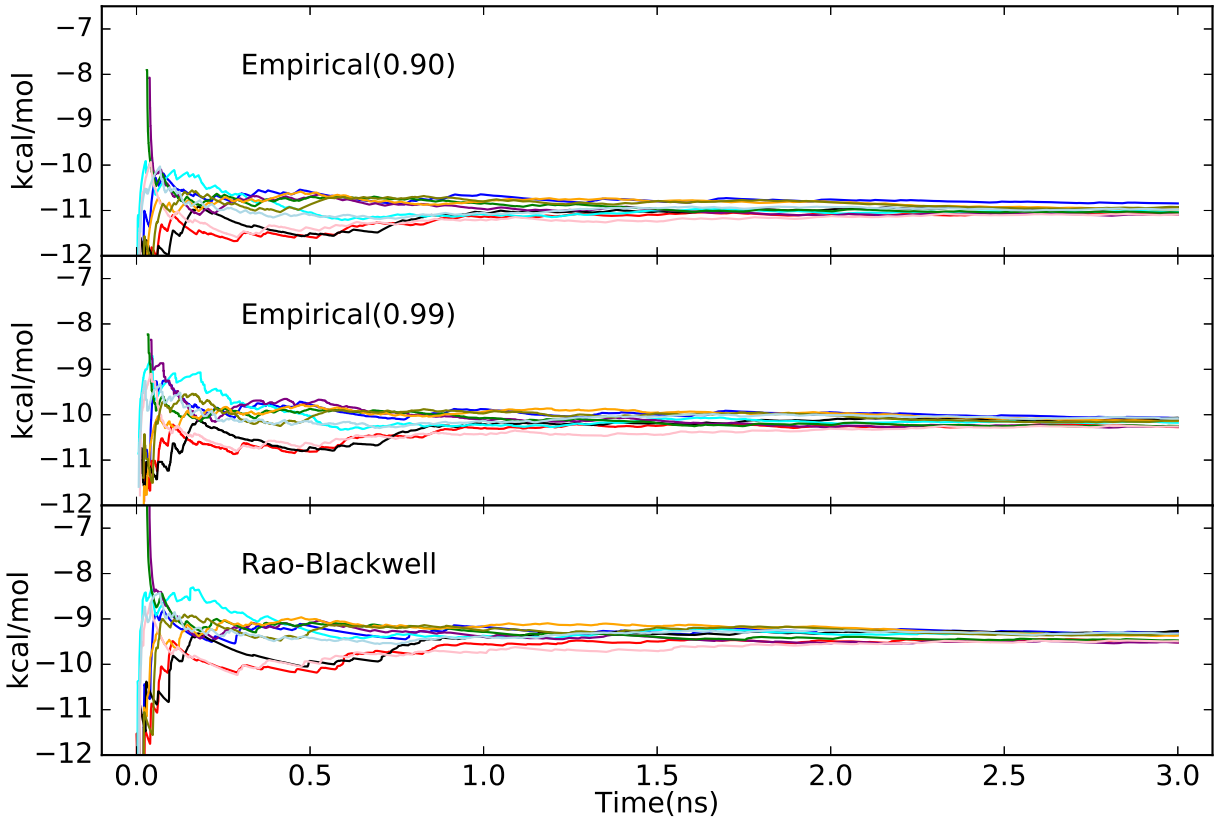

Figure S13. Estimated alchemical free energy changes in water using GSLD for multiple ligands from 10 individual repeat using empirical estimators with a cutoff of 0.9 (up) and 0.99 (middle) and the Rao-Blackwell estimator (bottom) for alchemical changess: benzene to p-xylene.

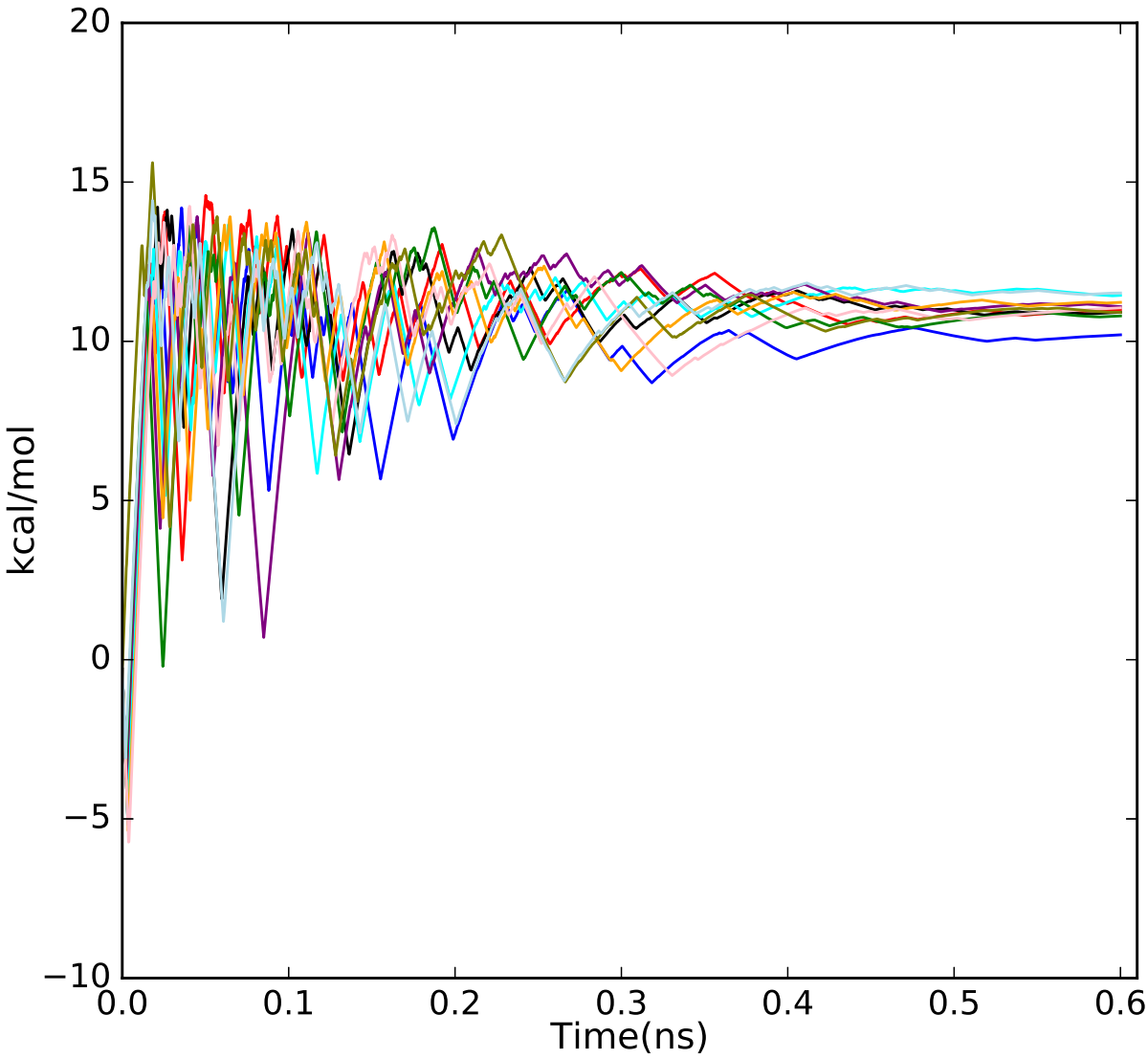

Figure S14. Biasing potential energies calculated using the Wang-Landau like algorithm for the alchemical changes: benzene to p-xylene in water.

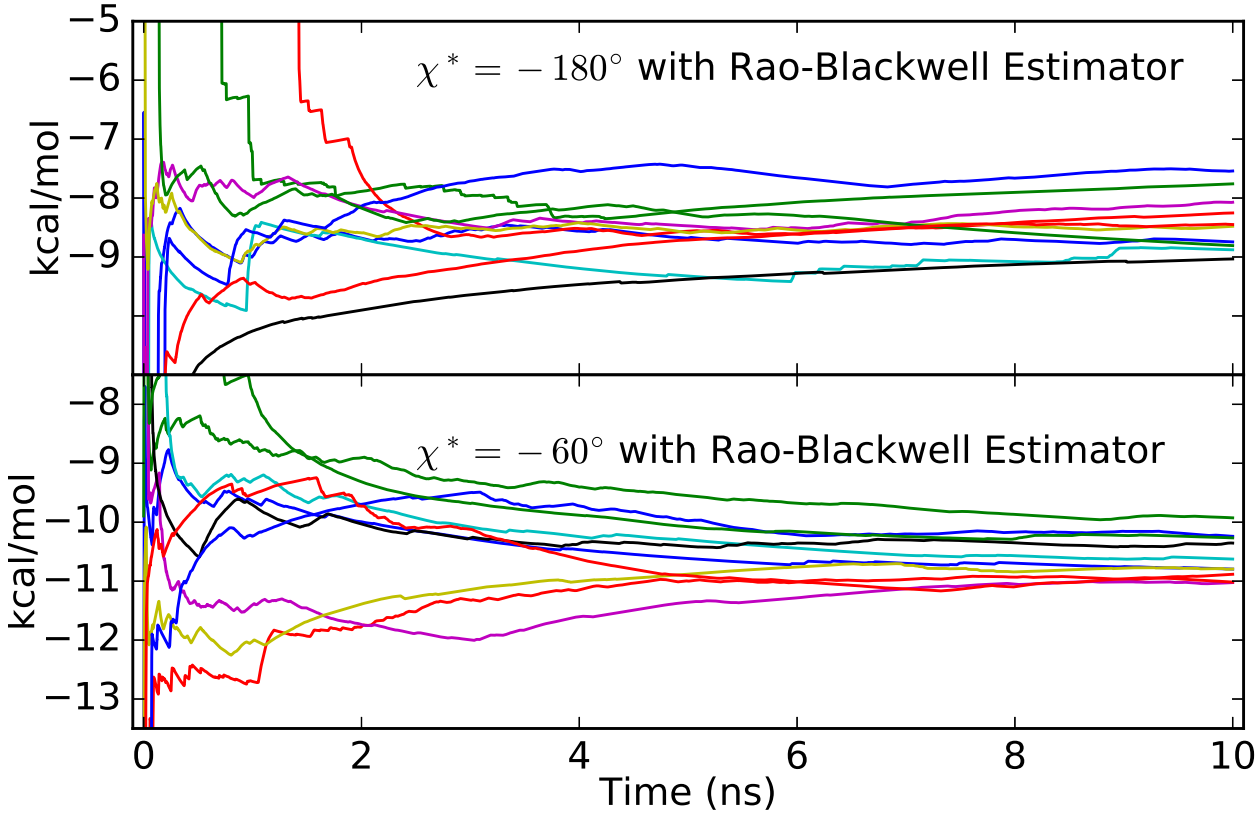

Figure S15. Estimated alchemical binding free energy changes from 10 individual repeat using GSLD and Rao-Blackwell estimator for the alchemical change: benzene to p-xylene binding with protein T4L when the side-chain dihedral angle χ (N-CA-CB-CG1) is restricted at $\chi^* = -180°$ (top) and $\chi^* = -60°$.

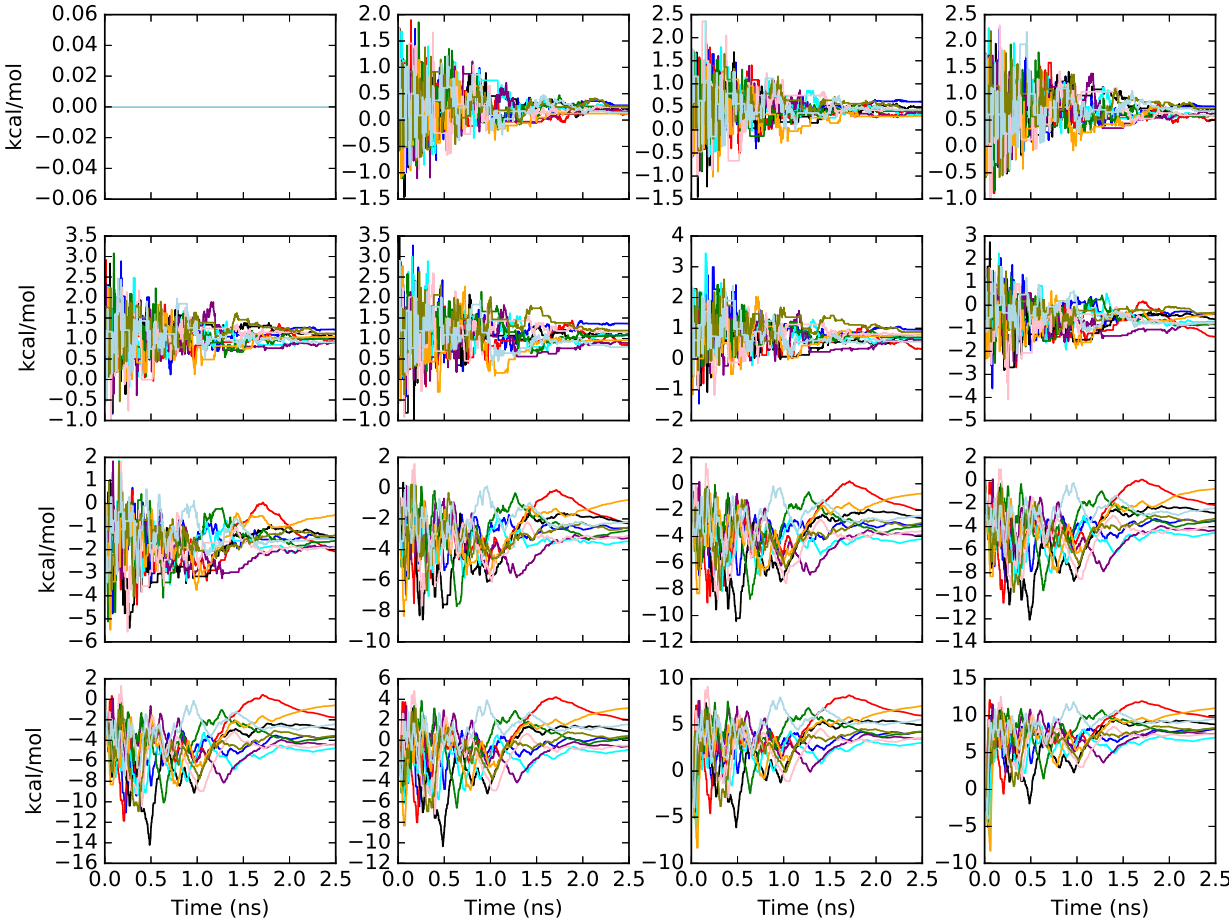

Figure S16. Relative biasing potential energies with respect to the state $\lambda = l_1$ calculated using the Wang-Landau like algorithm for the alchemical changes: benzene to p-xylene in T4 Lysozyme when the side chain dihedral angle χ is restricted around $-180°$.

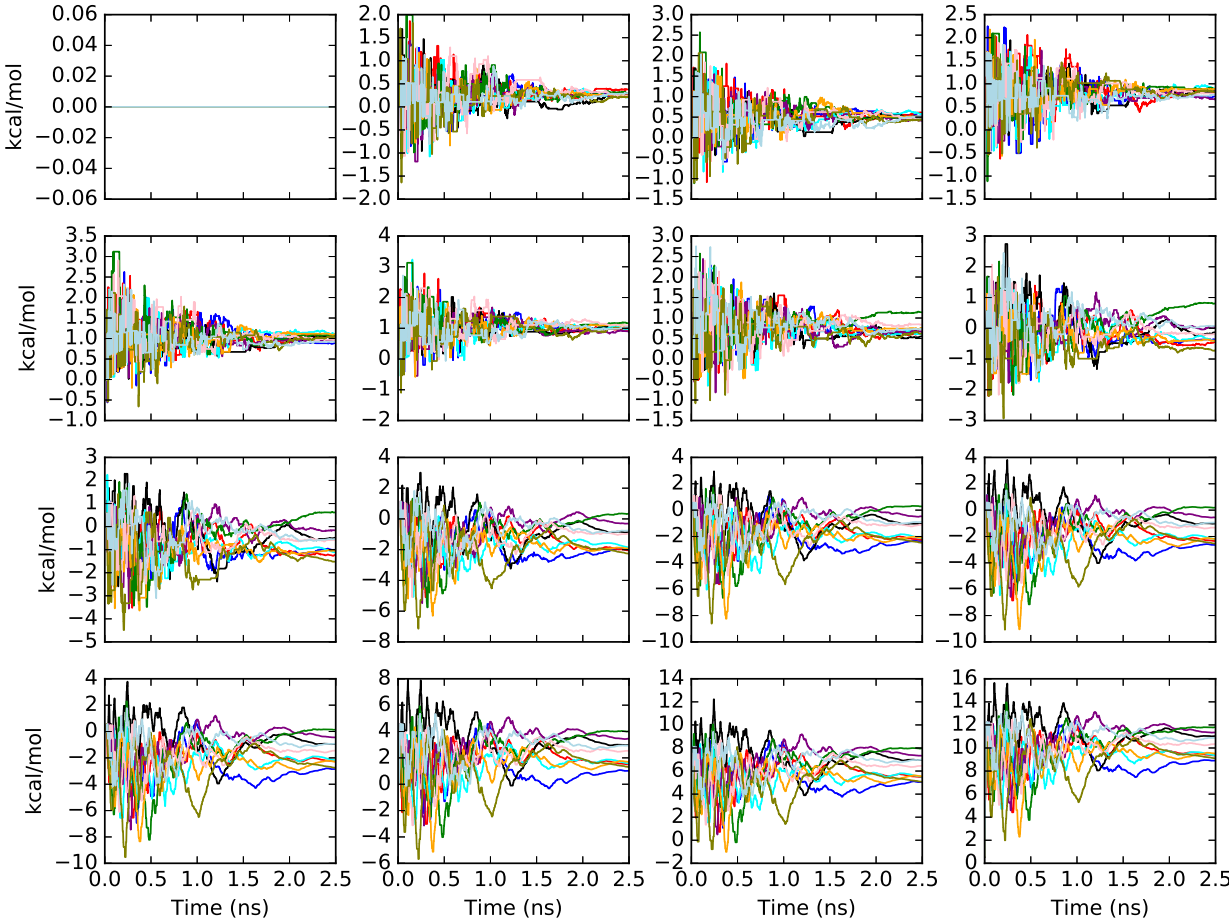

Figure S17. Relative biasing potential energies with respect to the state $\lambda = l_1$ calculated using the Wang-Landau like algorithm for the alchemical changes: benzene to p-xylene in T4 Lysozyme when the side chain dihedral angle χ is restricted around $-60°$

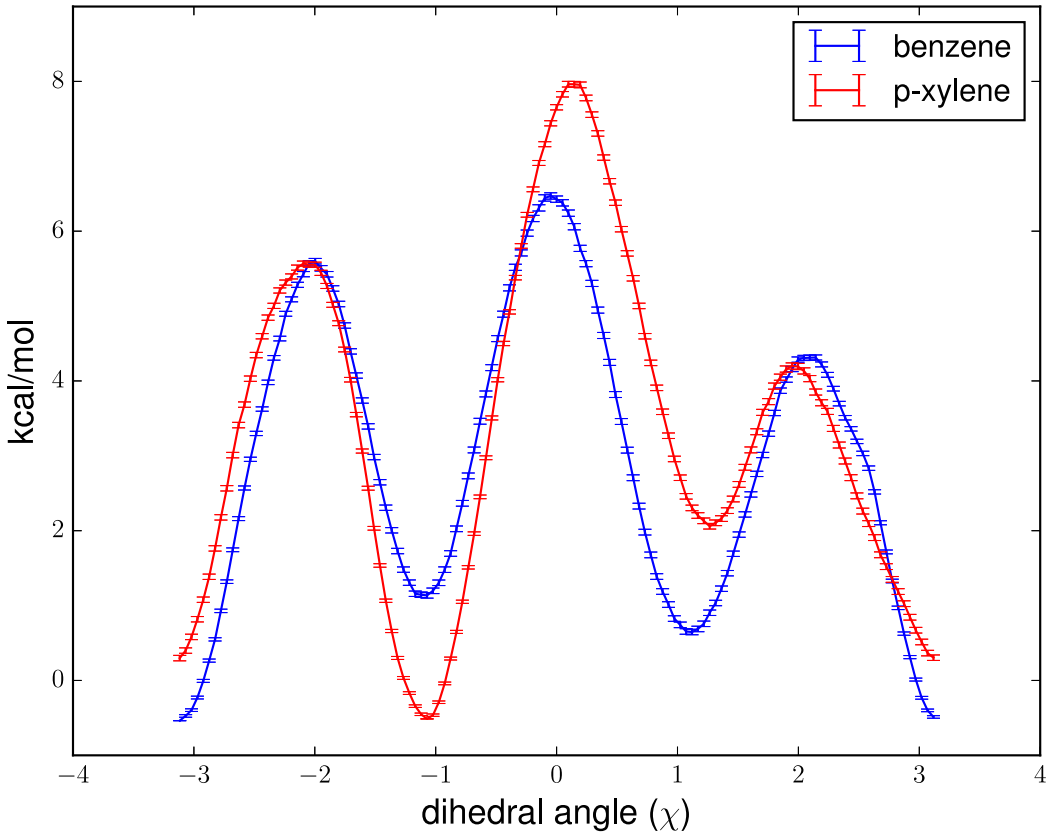

Figure S18. Potential of mean force (PMF) on the side chain dihedral angle N-CA-CB-CG1(χ) in the residue Val11 in T4 lysozyme. **Blue:** the PMF when T4 lysozyme binds with benzene. **Red:** the PMF when T4 lysozyme binds with p-xylene.