Supplementary Information

QSAR via Multisite λ-Dynamics in the Orphaned TSSK1B Kinase

Xiaorong Liu[1], Pui Ki Tsang[1], Matthew B Soellner[1], and Charles L Brooks III[1, 2, *]

[1]Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109, USA

[2]Biophysics Program, University of Michigan, Ann Arbor, Michigan 48109, USA

[*]To whom correspondence should be addressed. E-mail: brookscl@umich.edu Phone: (734) 647-6682

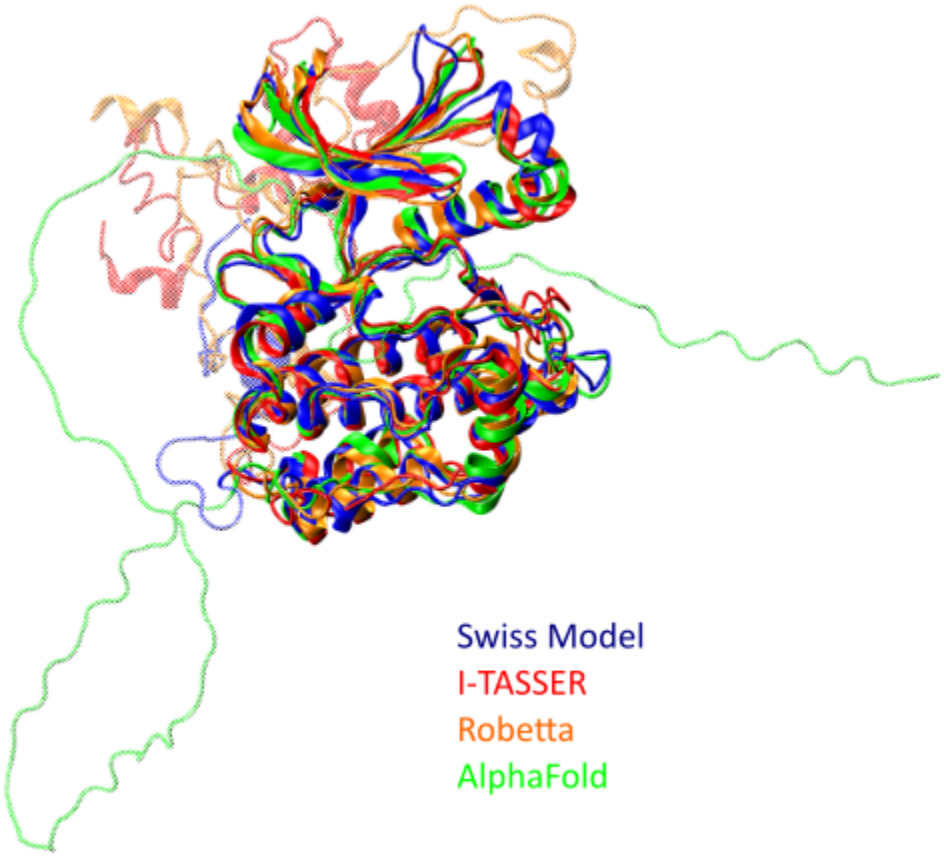


Figure S1. Predicted structures of TSSK1B from Swiss Model (blue), I-TASSER (red), Robetta (yellow), and AlphaFold 2 (green). Residues 1-275 are shown in solid color, while the disordered tail at C-terminus (residues 276-367) is transparent. These four structures are superimposed based on the well-folded region (residues 1-275).

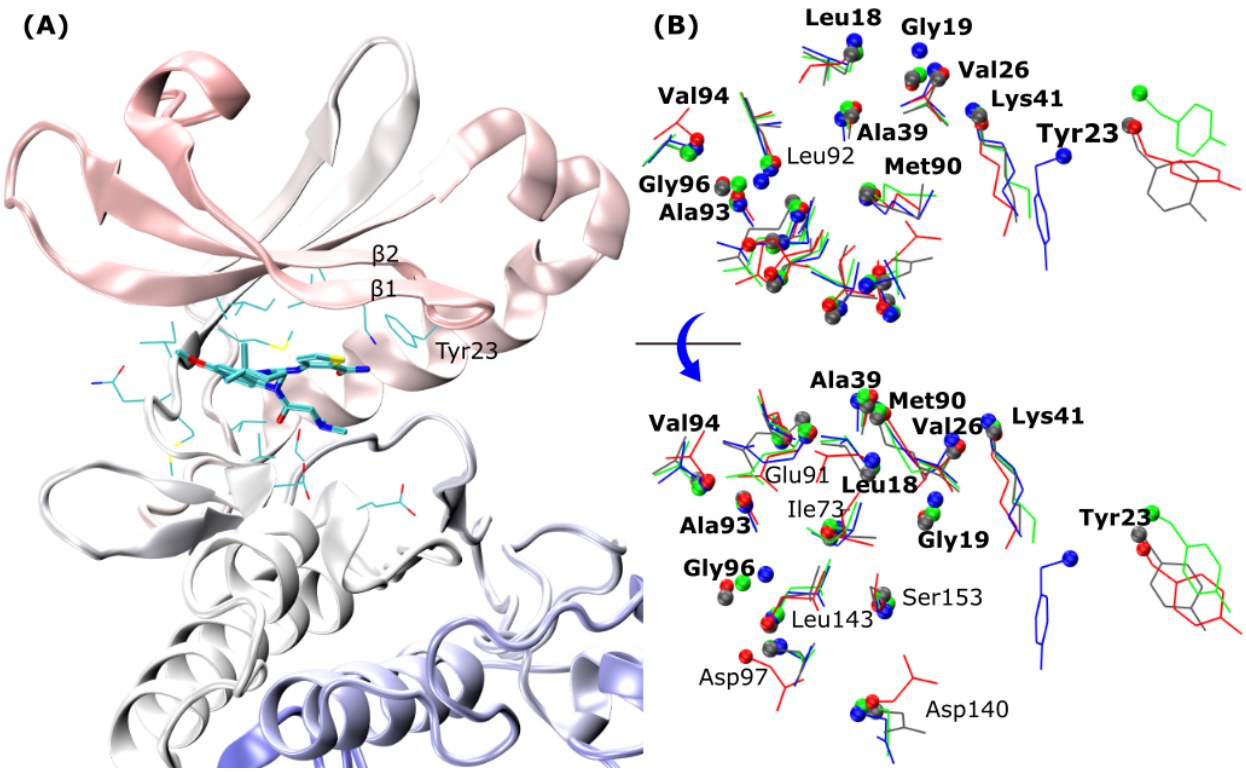


Figure S2. (A) Experimental structure of bovine G protein-coupled receptor kinase 1 in complex with the lead compound KQQ (PDB ID: 4PNI). Protein is shown in New Cartoon, with its color changing from red at the N terminus to blue at the C terminus. Ligand KQQ is shown in Licorice, and residues in the binding pocket are shown in line. A residue is considered in the binding pocket when any of its heavy atoms are within 4.5 Å of the heavy atoms of ligand KQQ. A total of 17 residues were identified based on this criterion. (B) Residues in the binding pocket of predicted TSSK1B structures from Swiss Model (blue), I-TASSER (red), Robetta (yellow), and AlphaFold 2 (green). Cα atoms of these 17 residues are shown as spheres. Structures of these residues are highly similar, except for Tyr23, which is located at the loop between β1 and β2 and quite flexible in our MD simulations.

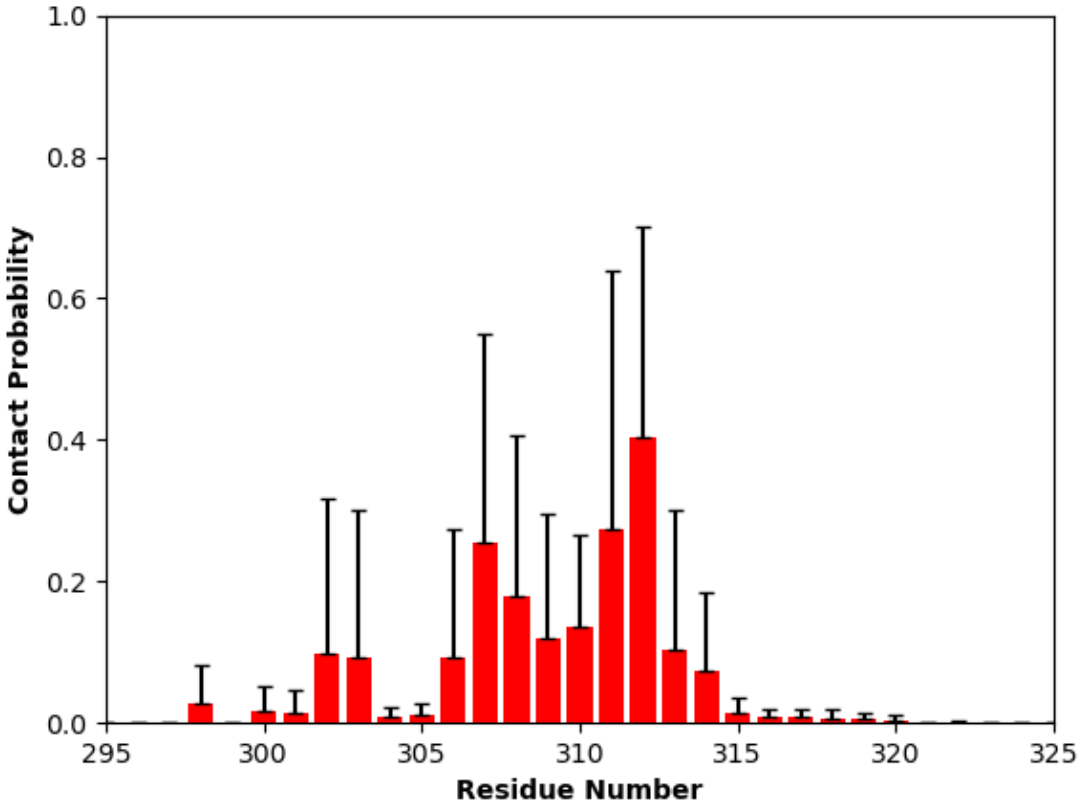


Figure S3. Probabilities of forming contacts between molecule KQQ and the C-terminal tail of TSSK1B (residues 276-367). The X-axis only ranges from 295 to 325, since the rest of the C-terminal tail didn't form contacts with molecule KQQ at all in our simulations. A contact is considered formed when the minimum distance of heavy atoms between KQQ and a residue is less than or equal to 4.2 Å. Five independent MD simulations were performed at a temperature of 298.15 K and a pressure of 1 atm, starting with five different structures built using Swiss Model. Each simulation lasted for more than 1 μs. Here averaged values over these five independent simulations were reported. The error bars indicate standard deviation. Although some residues show average contact probabilities greater than 0.2, we didn't observe any contacts that are consistently present in all simulations, as suggested by their large error bars. This again implies the lack of specific contact between molecule KQQ and the C-terminal tail of TSSK1B.

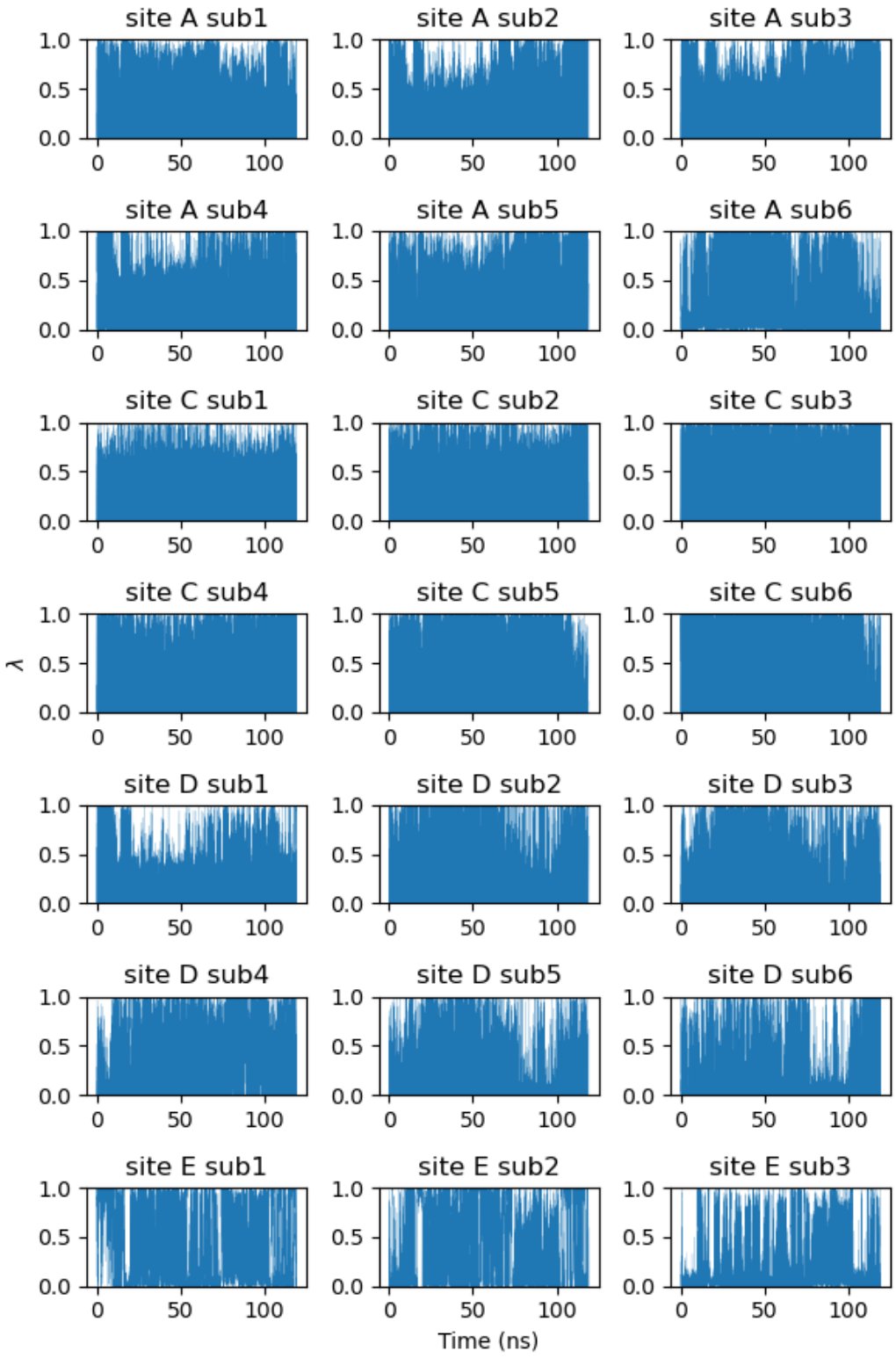
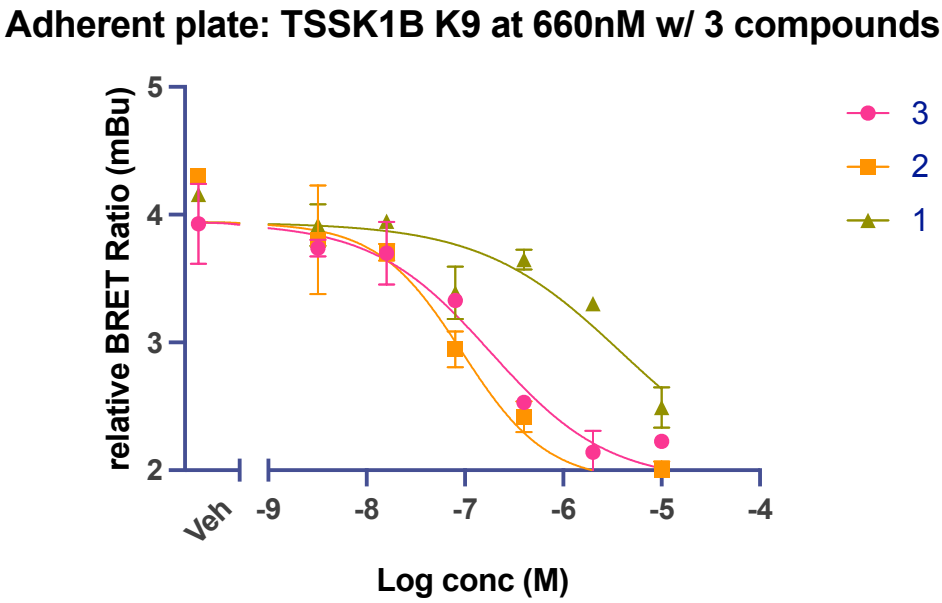


Figure S4. λ values as a function of simulation time in one production simulation of protein bound state in simulation set II.



	3	2	1	Global (shared)
IC50	1.821e-007	9.356e-008	3.872e-006	

Figure S5. Experimental measurements of IC_{50} for three ligands where sites A, B, C and D are the same as those in molecule KQQ and site E is one of the three substituents in Figure 2. The TSSK1B NanoBRET TE assay kit (NV4471) was purchased from Promega and carried out as described in the assay kit. HEK293 cells (ATCC) were used for transfection. Commercial nanoBRET tracer K9 (N2632) was used for the tracer at 660 nM. The adherent cell format was used as this led to optimal BRET. BRET ratios were calculated from the donor signal (415 nm) and acceptor signal (610 nm).

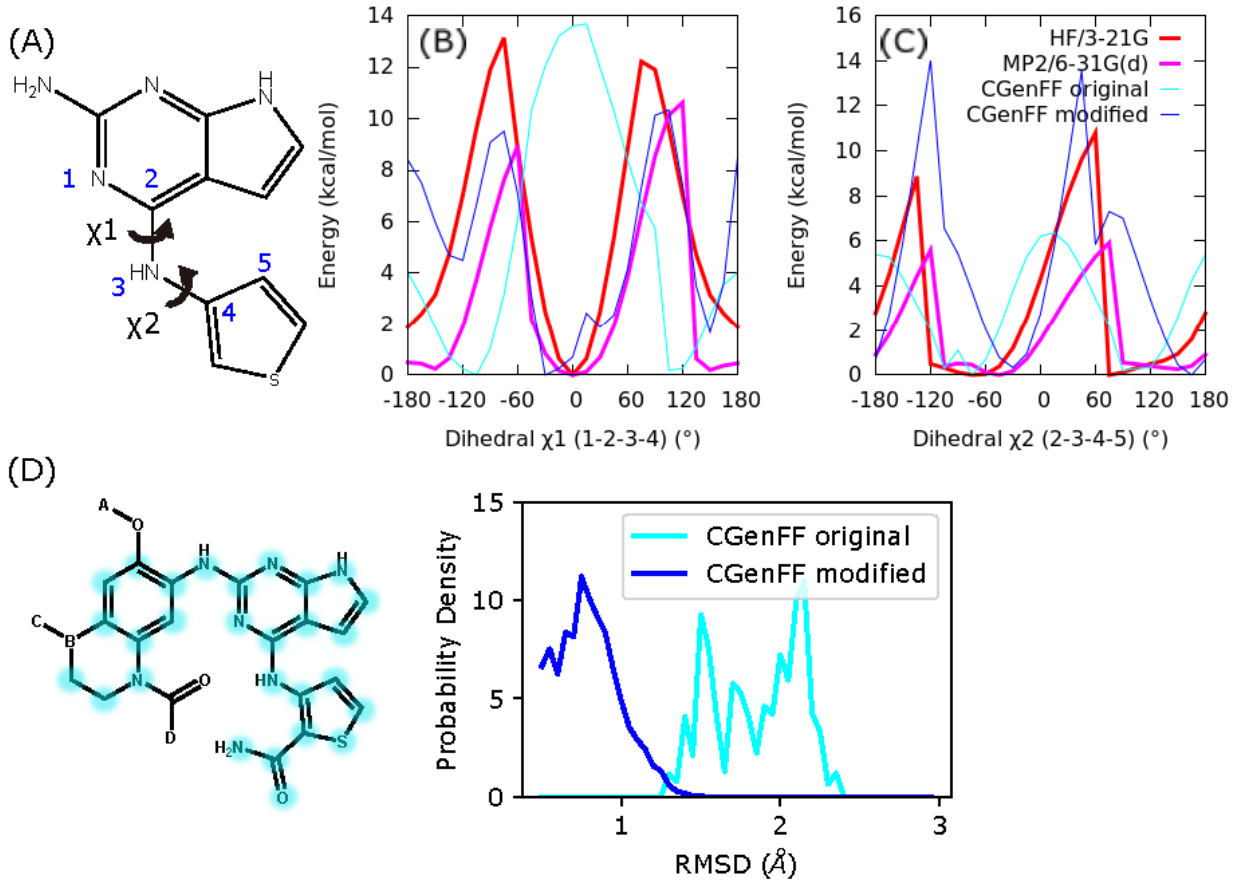


Figure S6. Dihedral scan of rotatable bonds connecting site E substituent with the maximum common substructure. (A) Model compound used in dihedral scans for χ1 and χ2. All dihedrals are defined using heavy atoms that are labeled in blue. (B) Dihedral scan for χ1. (C) Dihedral scan for χ2. All quantum mechanics calculations were performed using ORCA [1]. Note that although χ1 and χ2 are coupled, we only sequentially optimized χ1 and χ2 dihedral potentials using 1D cosine functions. This appears to be sufficient as the internal conformation of ligand in our MSλD simulations are highly consistent with that in PDB 4PNI, as shown in panel D. (D) Probability distribution of ligand heavy atom RMSD in our MSλD simulations with respect to that in PDB 4PNI. Atoms that are selected for RMSD calculations are highlighted in blue. Here we extracted snapshots from MSλD simulations where λ of substituent 3 at site E is greater than 0.8, without considering what substituents are at sites A, B, C or D.

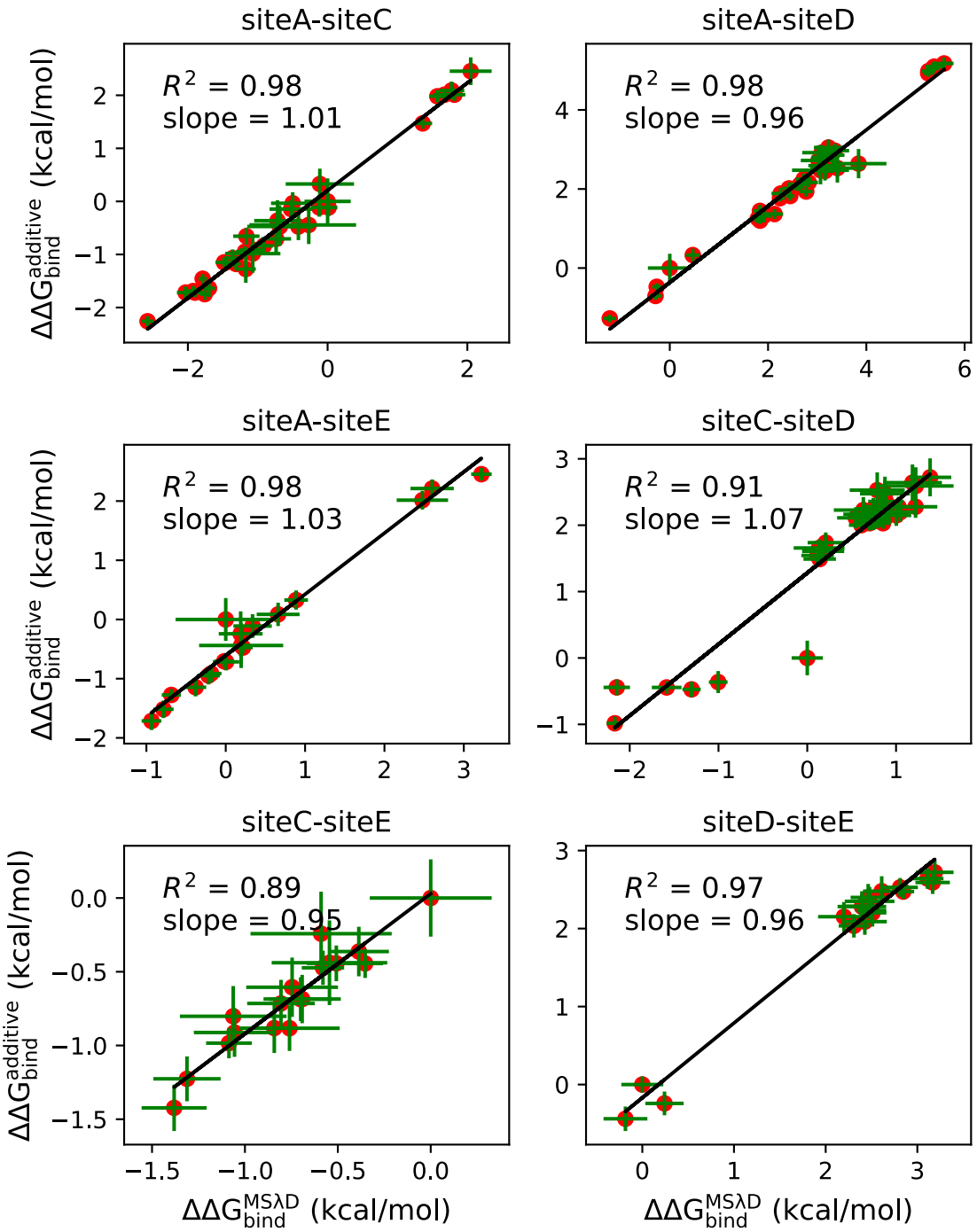


Figure S7. Comparison of relative binding free energies computed from the standard estimator in **Eq. 8** and the additive model in **Eq. 9** in the main text. Results were computed from simulation set II.

Table S1. Relative binding free energy values estimated using the additive model when substituent at a given site is varied.

Simulation Set II	Site A	Site C	Site D		Site E
H	0.00 ± 0.35	0.00 ± 0.25	0.00 ± 0.07	**1**	0.00 ± 0.08
Me	-0.71 ± 0.07	-0.36 ± 0.15	2.47 ± 0.05	**2**	-0.24 ± 0.13
Et	-0.47 ± 0.05	-0.47 ± 0.09	2.53 ± 0.09	**3**	-0.44 ± 0.14
Bu	-1.28 ± 0.06	-0.98 ± 0.07	2.59 ± 0.13		
iPr	0.33 ± 0.14	-0.44 ± 0.06	2.64 ± 0.11		
tBu	2.45 ± 0.07	-0.44 ± 0.10	2.72 ± 0.14		
Simulation Set III	**Site A**	**Site C**	**Site D**		**Site E**
C-H		0.00 ± 0.11			
H	0.00 ± 0.16	0.26 ± 0.11	0.00 ± 0.22	**1**	0.00 ± 0.26
Me	1.48 ± 0.05	0.01 ± 0.11	0.41 ± 0.13	**2**	-0.28 ± 0.07
Et	1.70 ± 0.05	-0.30 ± 0.07	1.48 ± 0.10	**3**	-0.89 ± 0.15
Bu	0.94 ± 0.08	-0.68 ± 0.09	1.72 ± 0.14		
iPr	2.88 ± 0.19	-0.79 ± 0.07	1.90 ± 0.11		
tBu	4.68 ± 0.39	-1.26 ± 0.13	2.58 ± 0.19		

References

1. Neese, F., et al., *The ORCA quantum chemistry program package.* Journal of Chemical Physics, 2020. **152**(22).