Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Information: A Robust Fragment-Based Method of Calculating Hydrogen Atom Transfer Activation Barrier in Complex Molecules

Yizhou Liu,[a]* Frank C. Pickard, IV,[bc] Gregory W. Sluggett,[a] Iasson G. Mustakis[d]

a. Analytical Research and Development, Pfizer Worldwide Research and Development, 445 Eastern Point Road, Groton, CT 06340, USA. * Email: Yizhou.Liu@pfizer.com
b. Pharmaceutical Sciences, Pfizer Worldwide Research & Development, Groton, CT 06340, United States
c. Current address:
d. Chemical Research & Development, Pfizer Worldwide Research & Development, Groton, CT 06340, United State

Table of Contents

Energy Correlations Between Successive Steps of Calculation

As the geometry optimization process progresses from low levels of theory to higher ones with some conformers eliminated by energy filters between some steps, it is important to evaluate energy correlations between successive steps of optimization to choose appropriate energy cutoffs. This evaluation avoids premature elimination of a conformer that has a high energy at a low level of theory but later optimizes to a low energy at a high level of theory. We performed this evaluation with reaction 1A and applied the energy cutoffs based on its results to the other reactions.

Figure S1a shows the energy correlation between GFN-xTB with restrained reaction core and B3LYP-GD3/6-31G(d,p) with frozen reaction core. The most stable conformation in each case is set to 0 kcal/mol. The scattering in the plot is due to several reasons: 1) The geometry can change as the conformers from GFN-xTB are reoptimized by B3LYP-GD3. 2) Solvation free energy, calculated by ALPB in GFN-XTB and IEFPCM in B3LYP-GD3, is included in the comparison in addition to electronic energies. The different solvation models can contribute to scattering in the correlation. 3) Relative electronic energy is different between the two levels of theory even in single-point (same geometry) calculations. Based on Figure S1a, we adopt a conservative cutoff of 10 kcal/mol going from GFN-xTB to B3LYP-GD3, because no conformer above 10 kcal/mol in GFN-xTB optimizes to a low energy in B3LYP-GD3. In fact, a more aggressive cutoff of 5 kcal/mol still seems reasonable because only two conformers between 5 and 10 kcal/mol in GFN-xTB optimized to energies of 4–5 kcal/mol in B3LYP-GD3.

Figure S1b shows the energy correlation between conformers from B3LYP-GD3/6-31G(d,p) optimized with frozen reaction core and the corresponding TS guesses from relaxed potential energy scans. Both include electronic energies and solvation free energies by IEFPCM. Overall, the correlation appears tight except for one large and a few smaller outliers. This relatively tight correlation is not automatically expected without this evaluation, because the TS guess from the relaxed scan can have a different geometry and therefore energy from the initial conformation. The relatively tight correlation indicates that the initial geometry resulting from the frozen-core B3LYP-GD3 optimization is already close to the TS guess. In this work we adopt a conservative cutoff of 10 kcal/mol, although a 5 kcal/mol cutoff is likely sufficient based on the result.

Figure S1c plots the energy (electronic energy + solvation free energy) correlation between TS guesses and the final optimized TS geometries, which matches almost exactly. This result suggests that the TS guess is almost identical to the optimized TS geometry at the same level of theory, *i.e.*, the guesses from the relaxed scan are very good. Most of the TS guesses successfully optimize to HAT TS's as confirmed by IRC, except for a small number that drift to torsional TS's.

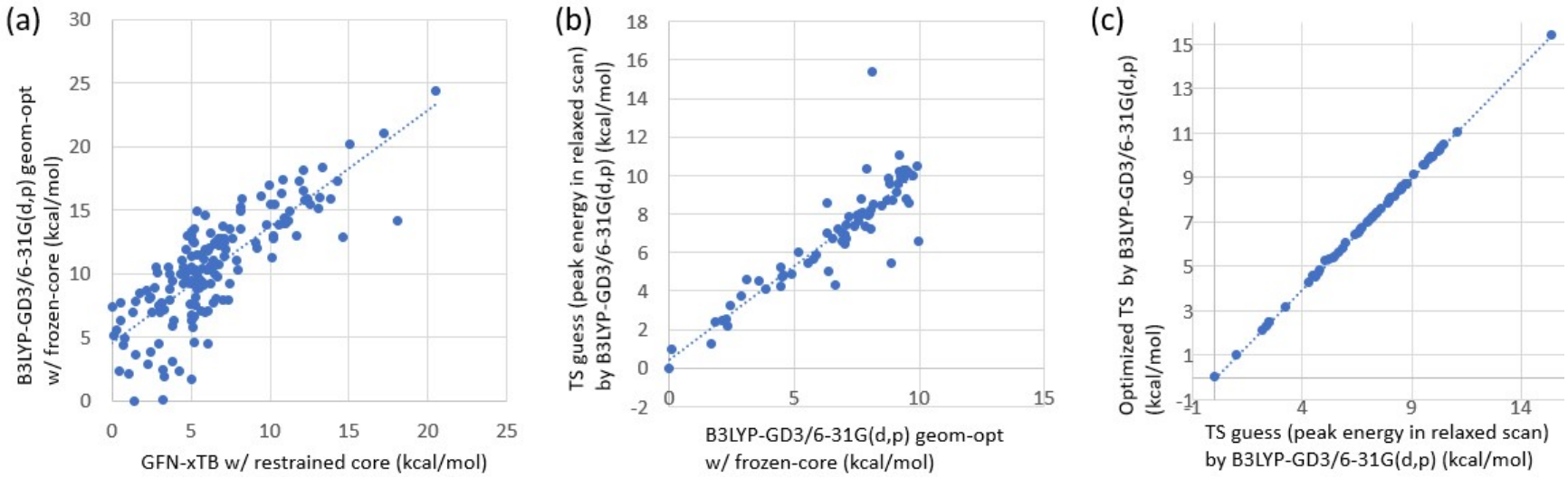

Figure S1. Energy correlations between successive steps of geometry optimization.

In this work, we obtain the electronic energy of the TS by single-point calculation with DLPNO-CCSD(T)/cc-pVTZ. Figure S2 shows the correlation between the TS electronic energies by DLPNO-CCSD(T)/cc-pVTZ and B3LYP-GD3/Def2TZVP. The MAD (mean absolute difference) is 1.2 kcal/mol with some differences exceeding 2 kcal/mol. DLPNO-CCSD(T)/cc-pVTZ is expected to yield more accurate results, especially for the unusual bond geometries in the TS.

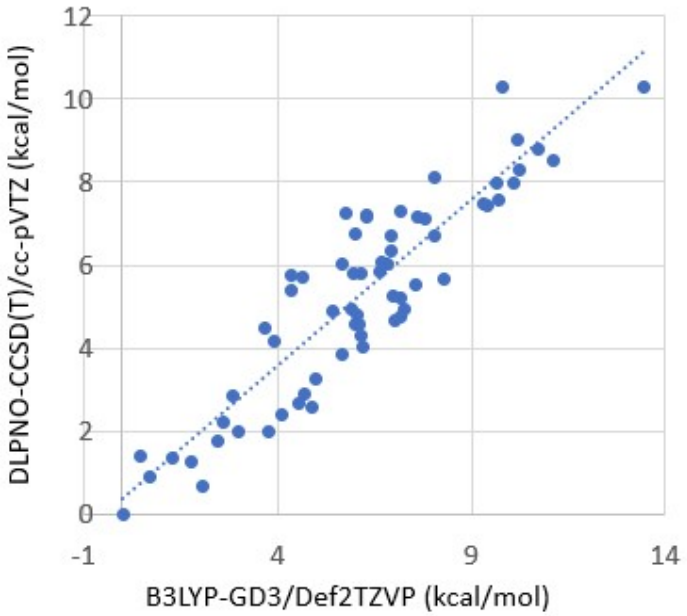

Figure S2. Correlation between TS electronic energies from single-point calculations by B3LYP-GD3/Def2TZVP and DLPNO-CCSD(T)/cc-pVTZ.

Contributions of ZPVE and Thermo Correction Terms

Besides electronic energy and solvation free energy, ZPVE and thermo correction terms including G_{trans}^{o}, G_{rot} and G_{vib} also contribute to free energy. Figure S3 shows that the combined contribution of ZPVE and thermo corrections varies significantly between different TS conformations. Only the differences relative to the conformer of the lowest value (#43) are displayed. Among the thermo terms, G_{trans}^{o} is the same for all conformers and G_{rot} also has very small variations. Therefore, the differences observed are largely due to ZPVE and G_{vib}, both of which are calculated by mRRHO in this work.

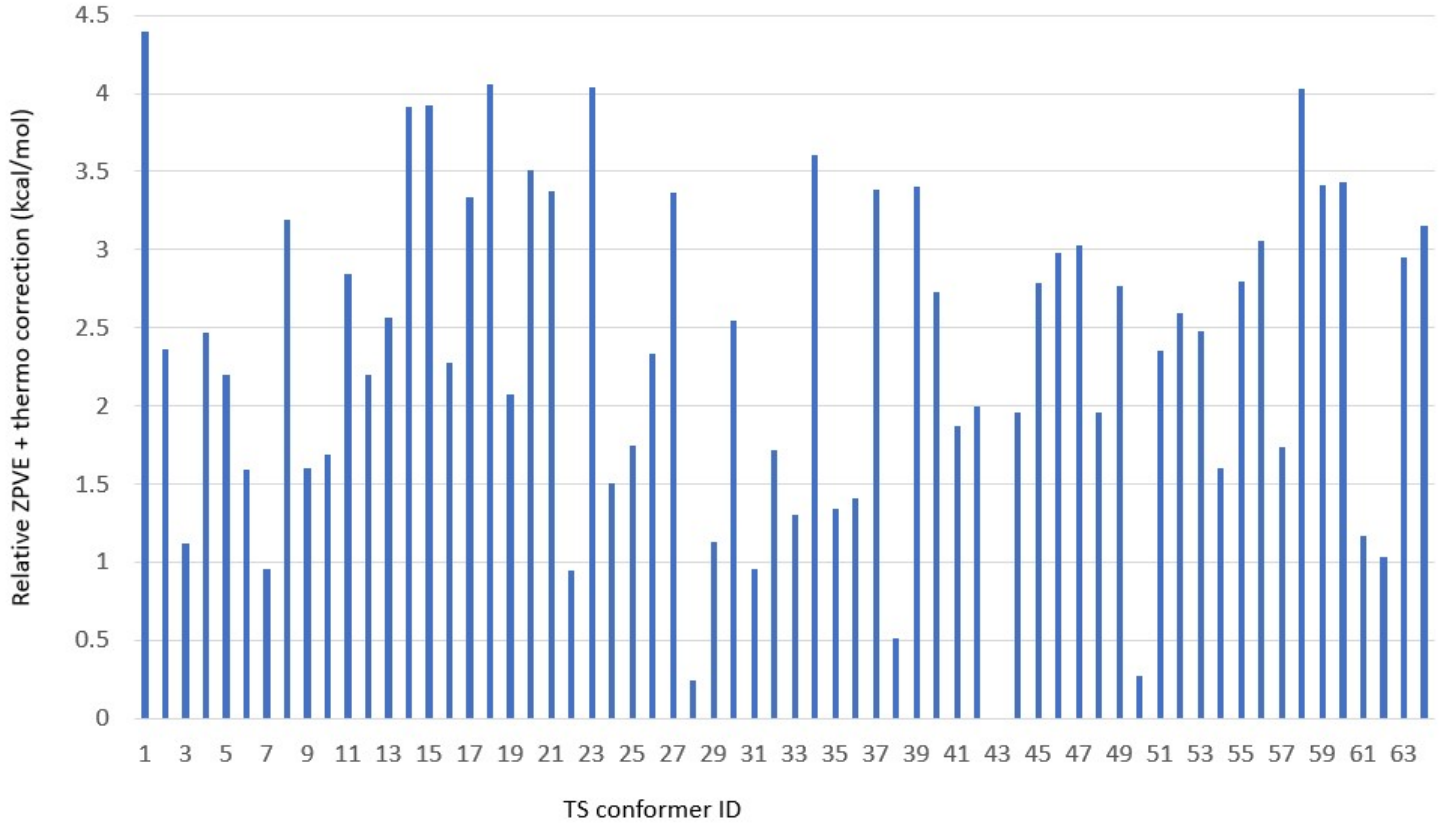

Figure S3. Relative ZPVE + thermo corrections for different TS conformations of 1A. The lowest value (# 43) is set to zero.

Compensation Effect Among Different Free Energy Components

In the main text, we mentioned that several components of the free activation energy, namely, $\Delta^{\ddagger}\langle H_0 + \Delta G_{solv}\rangle$, $\Delta^{\ddagger}\langle G_{vib}\rangle$, and $-T\Delta^{\ddagger}S_{mix}$, all display significant differences between whole molecules and corresponding fragments, but adding them together leads to significantly reduced differences in $\Delta^{\ddagger}G^{o}_{MC}$. This compensation effect is clearly seen in the multi-conformational single-step reaction path analysis (Table 2). It is also clearly observed in both the reactant binding step (Table 3) and the HAT step (Table 4) of the multi-conformational multi-step reaction path analysis.

Figure S4 shows the itemized (free) energy terms in Table 3 and 4 as stacked columns. There is a clear trend that a large magnitude of $D\left(\Delta\left\langle H^{bind}_{0} + \Delta G^{bind}_{solv}\right\rangle\right)$ tends to be compensated by a large $D\left(\Delta\left\langle G^{bind}_{vib}\right\rangle - T\Delta S^{bind}_{mix}\right)$ with the opposite sign (Figure S4a). The same trend is also observed in the HAT step (Figure S4b). The compensation between $D\left(\Delta\langle H_0 + \Delta G_{solv}\rangle\right)$ and $D\left(-T\Delta S_{mix}\right)$ is due to conformational averaging and can be understood by comparing to the LC-TST. Assuming that the lowest-G conformation in the LC-TST also has the lowest enthalpy H_{LC}, then based on Eq (8), $\langle H\rangle$ in MC-TST cannot be more negative than H_{LC} because $\langle H\rangle$ receives contributions from higher-energy conformations. Therefore, their difference $\langle H\rangle - H_{LC}$ is a positive quantity that represents the contribution of CF to enthalpy. Meanwhile, MC-TST also receives a negative contribution $-TS_{mix}$ that is absent in TST-LC, which represents the contribution of CF to entropy. The self-cancelation of these two opposite contributions of CT to enthalpy and entropy can be propagated to a reaction as correlated changes in enthalpy and mixing entropy, as well as correlated differences between different reactions, as observed for $D\left(\Delta\langle H_0 + \Delta G_{solv}\rangle\right)$ and $D\left(-T\Delta S_{mix}\right)$. The compensation effect between $D\left(\Delta\langle H_0 + \Delta G_{solv}\rangle\right)$ and $D\left(\Delta\langle G_{vib}\rangle\right)$ is less pronounced (Figure S4) and unrelated to conformational exchange because this effect is also present on a single-conformational basis between $D\left(\Delta^{\ddagger}H_0 + \Delta^{\ddagger}\Delta G_{solv}\right)$ and $D\left(\Delta^{\ddagger}G_{vib}\right)$ based on LCSS-TST calculations(See Figure S5). This compensation presumably arises due to weakening of intermolecular interactions by low-frequency vibrations of the reactants.[1 2, 3]

While it tends to reduce free activation energy differences between whole molecules and fragments, any compensation effect is more likely a trend than a rule. For example, the outlier 1D/1d in the conventional fragment-based method exhibits a large $D\left(\Delta^{\ddagger}G^{o}_{MC}\right)$ of -5.0 kcal/mol (Table 2), mostly because of a large $D\left(\Delta G^{o,bind}_{MC}\right)$ of -4.3 kcal/mol (Table 3) that arises from an uncompensated mixing entropy difference in the reaction binding step(Figure S4a). Incomplete compensation like this undermines the reliability of the conventional FB method. In contrast, the binding-corrected FB method explicitly takes the difference in reactant binding free energies into account and therefore does not reply on any compensation effect in the binding step. However, the compensation effect in the covalent activation step, as shown in Figure S4b, can benefit the binding-corrected FB method which assumes that $\Delta^{\ddagger}G^{HAT}_{whole} \approx \Delta^{\ddagger}G^{HAT}_{frag}$.

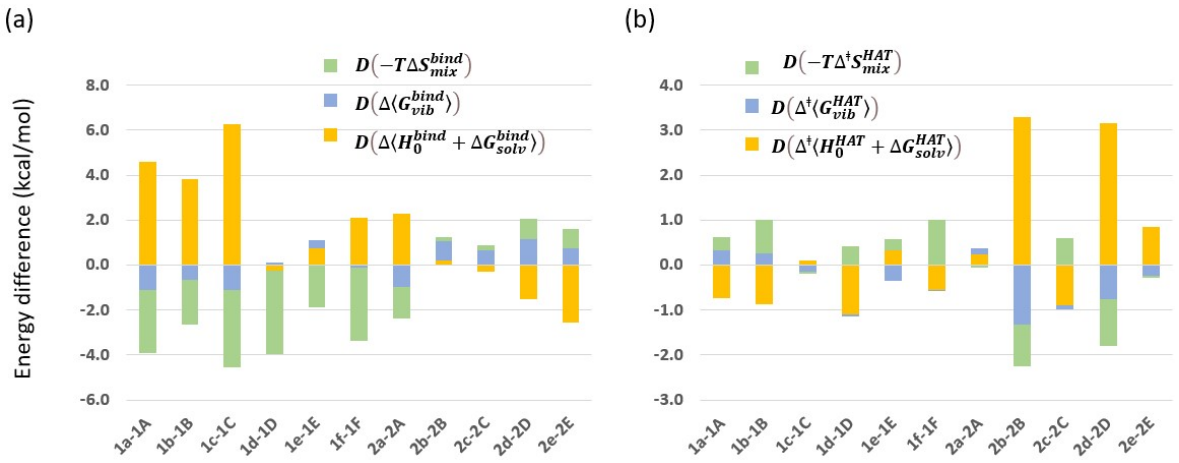

Figure S4. Compensation effects in the reactant binding step (a) and HAT step (b) of the MCMS-TST.

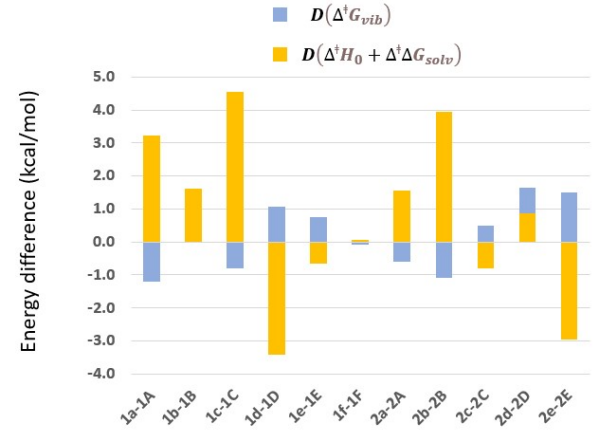

Figure S5. Compensation between $D(\Delta^{\ddagger} H_0 + \Delta^{\ddagger}\Delta G_{solv})$ and $D(\Delta^{\ddagger} G_{vib})$ calculated based on the most stable conformations.

Other Supporting Figures and Table

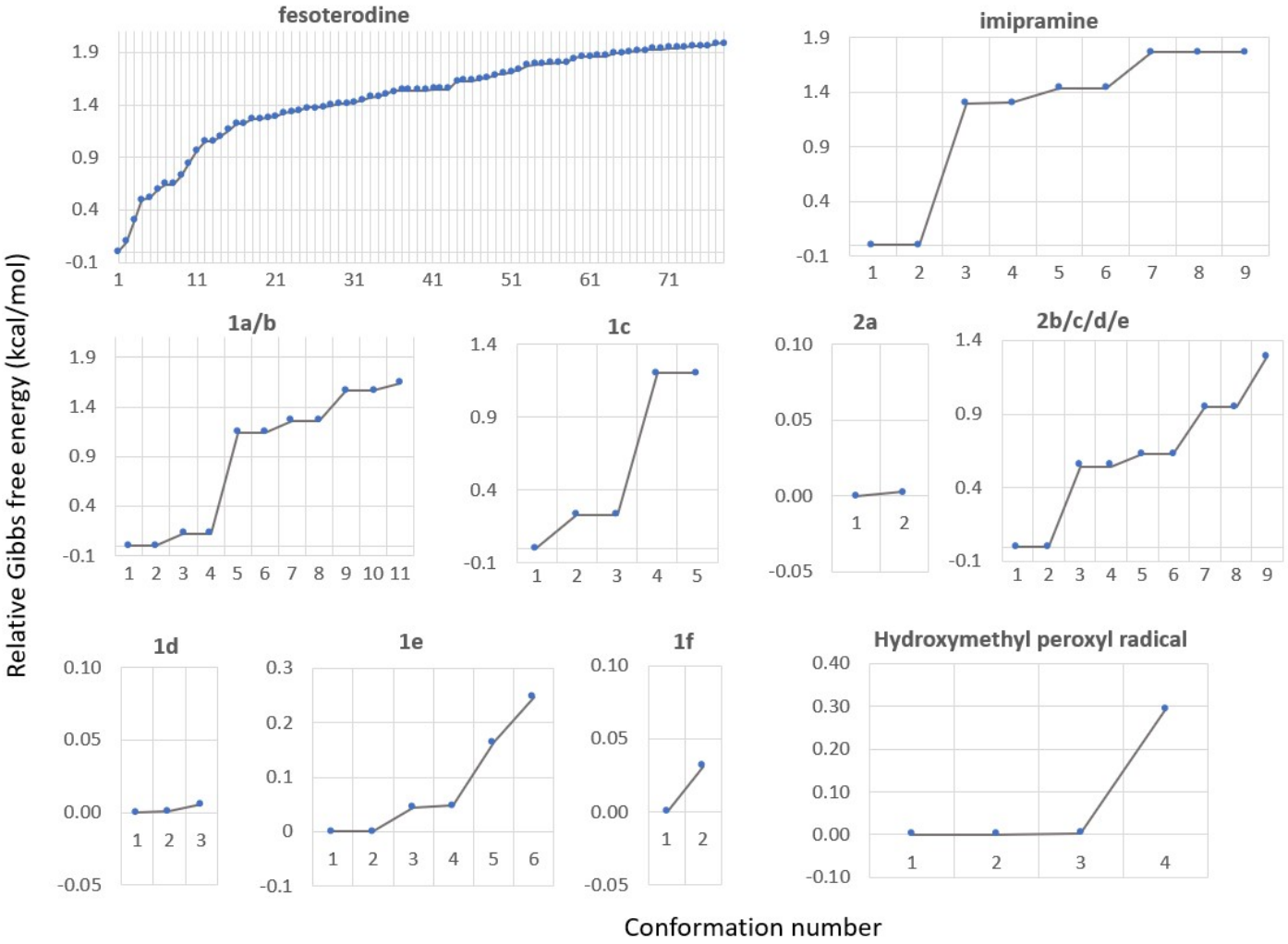

Figure S6. Conformational free energy distribution of reactants. The first conformation is the most stable conformation whose free energy is set to zero. Only conformations within a 2 kcal/mol window are displayed.

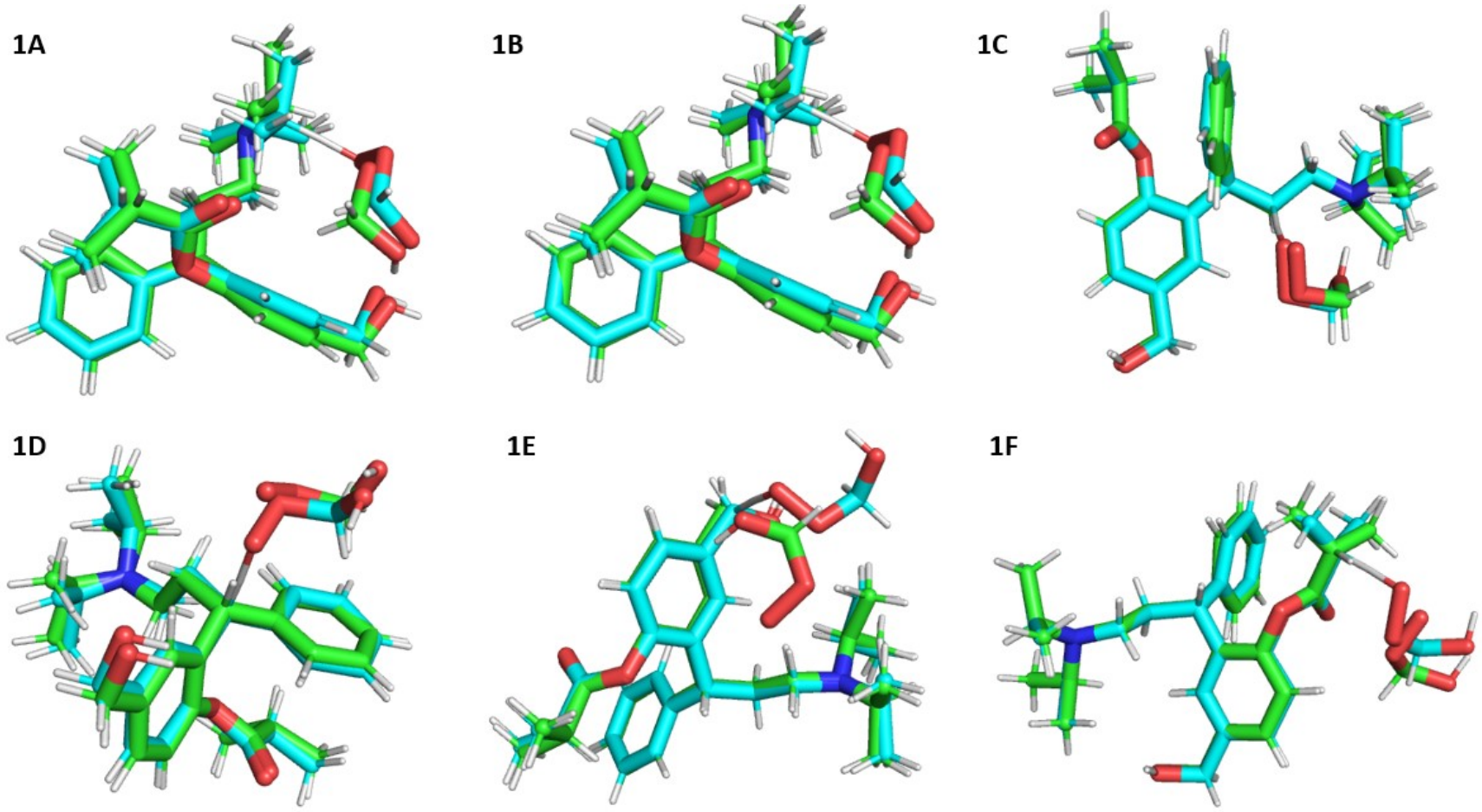

Figure S7. Superposition of reactant complexes (green) and TS's (cyan) of fesoterodine. Only the most stable TS and its corresponding reactant complex are displayed.

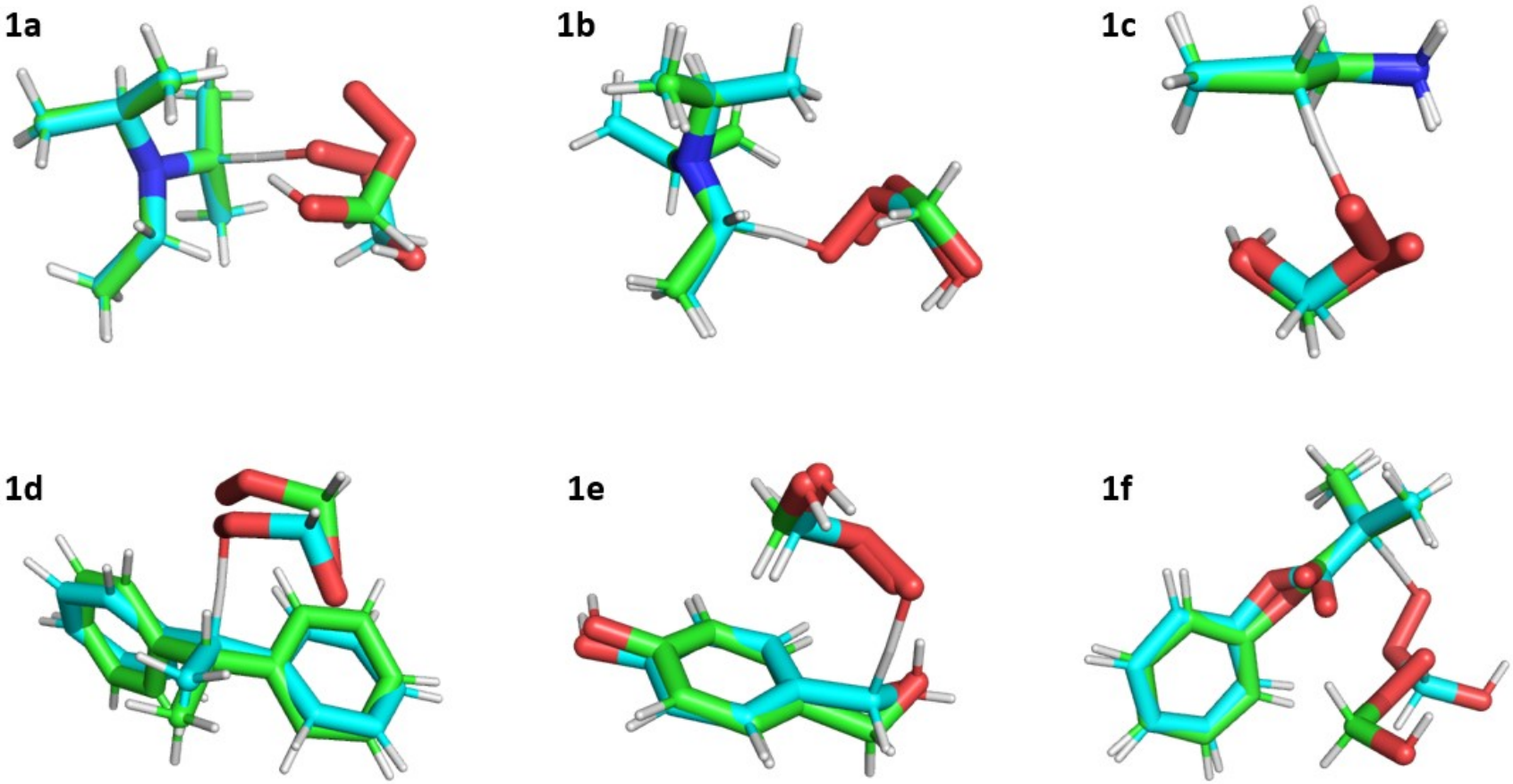

Figure S8. Superposition of reactant complexes (green) and TS's (cyan) of fesoterodine fragments. Only the most stable TS and its corresponding reactant complex are displayed.

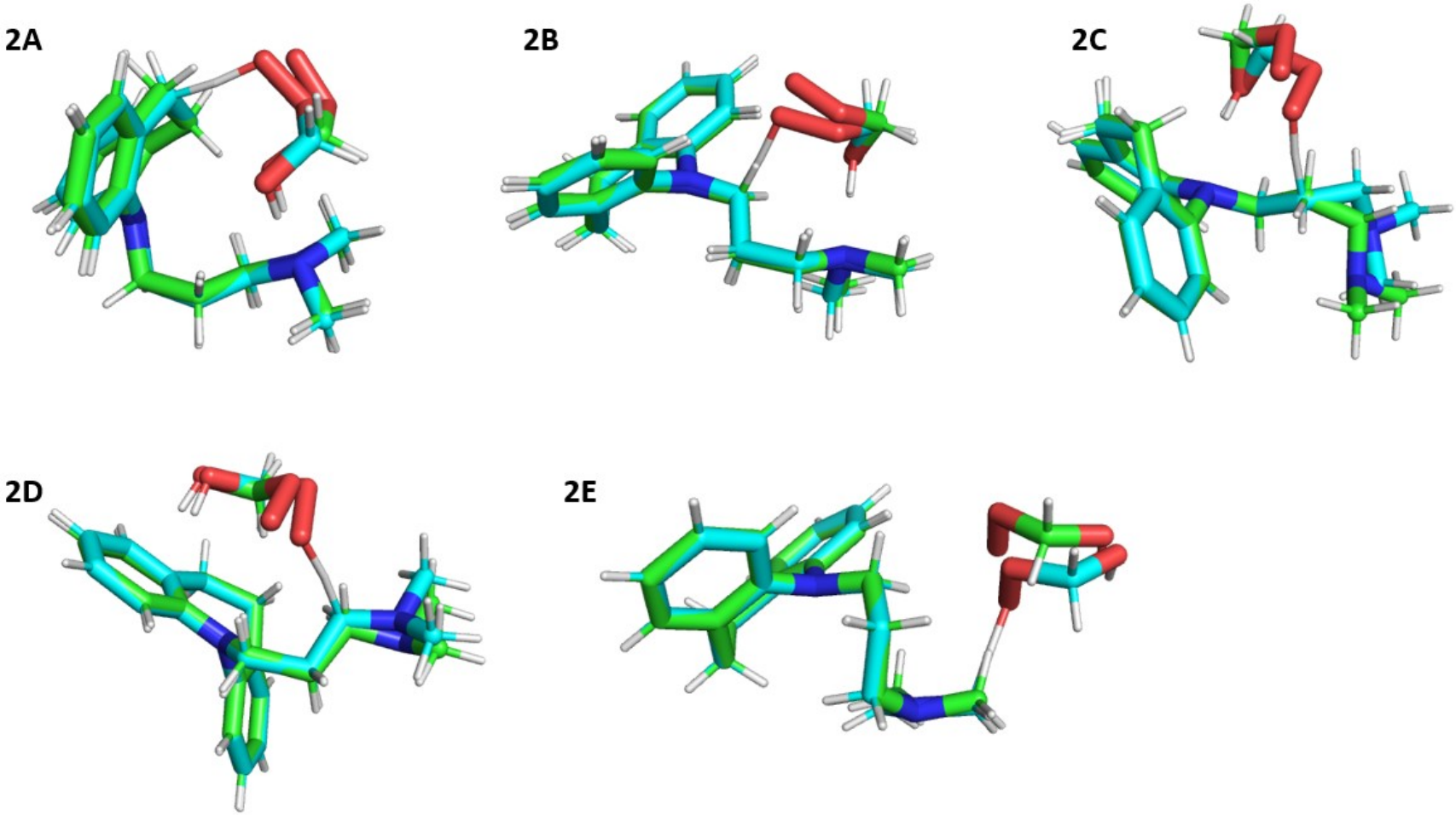

Figure S9. Superposition of reactant complexes (green) and TS's (cyan) of imipramine. Only the most stable TS and its corresponding reactant complex are displayed.

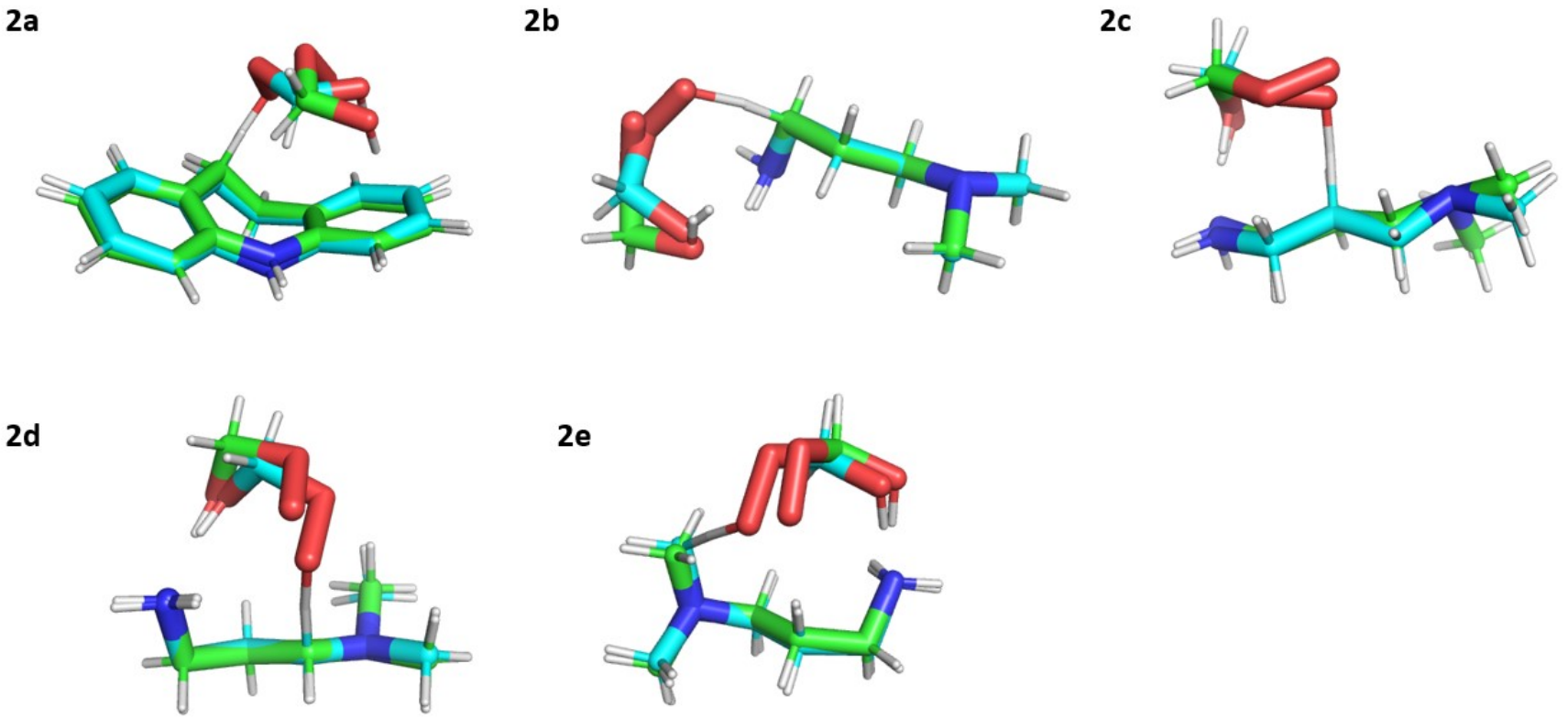

Figure S10. Superposition of reactant complexes (green) and TS's (cyan) of imipramine fragments. Only the most stable TS and its corresponding reactant complex are displayed.

Table S1. Changes of electronic energy U_e, zero-point vibrational energy $ZPVE$, and solvation free energy ΔG_{solv} along the multi-step reaction path.

	substrate (Hartree)			radical (Hartree)			reactant complex (Hartree)			TS (Hartree)			Binding step (kcal/mol)			HAT step (kcal/mol)		
	U_e	$ZPVE$	ΔG_{solv}	U_e	$ZPVE$	ΔG_{solv}	U_e	$ZPVE$	ΔG_{solv}	U_e	$ZPVE$	ΔG_{solv}	ΔU_e	$\Delta ZPVE$	$\Delta\Delta G_{solv}$	$\Delta^{\ddagger}U_e$	$\Delta^{\ddagger}ZPVE$	$\Delta^{\ddagger}\Delta G_{solv}$
1A	-1291.2155	0.5846	-0.0156	-265.0739	0.0484	-0.0118	-1556.3030	0.6349	-0.0275	-1556.2792	0.6308	-0.0334	-8.5	1.1	0.0	14.9	-2.6	-3.7
1B	-1291.2155	0.5846	-0.0156	-265.0739	0.0484	-0.0118	-1556.3033	0.6338	-0.0241	-1556.2788	0.6299	-0.0327	-8.7	0.5	2.1	15.4	-2.4	-5.4
1C	-1291.2155	0.5846	-0.0156	-265.0739	0.0484	-0.0118	-1556.3101	0.6354	-0.0223	-1556.2755	0.6292	-0.0252	-13.0	1.5	3.2	21.8	-3.9	-1.8
1D	-1291.2155	0.5846	-0.0156	-265.0739	0.0484	-0.0118	-1556.3016	0.6336	-0.0230	-1556.2701	0.6283	-0.0251	-7.7	0.4	2.8	19.8	-3.4	-1.3
1E	-1291.2155	0.5846	-0.0156	-265.0739	0.0484	-0.0118	-1556.3019	0.6346	-0.0284	-1556.2689	0.6286	-0.0292	-7.9	1.0	-0.6	20.7	-3.8	-0.5
1F	-1291.2155	0.5846	-0.0156	-265.0739	0.0484	-0.0118	-1556.3060	0.6346	-0.0234	-1556.2675	0.6279	-0.0253	-10.4	1.0	2.5	24.2	-4.2	-1.2
1a	-370.3599	0.2622	0.0011	-265.0739	0.0484	-0.0118	-635.4382	0.3116	-0.0119	-635.4171	0.3081	-0.0169	-2.7	0.6	-0.7	13.2	-2.2	-3.2
1b	-370.3599	0.2622	0.0011	-265.0739	0.0484	-0.0118	-635.4371	0.3111	-0.0117	-635.4182	0.3071	-0.0159	-2.0	0.3	-0.6	11.8	-2.5	-2.6
1c	-174.1781	0.1217	-0.0060	-265.0739	0.0484	-0.0118	-439.2554	0.1710	-0.0186	-439.2212	0.1647	-0.0208	-2.1	0.5	-0.5	21.5	-3.9	-1.4
1d	-540.8972	0.2380	-0.0025	-265.0739	0.0484	-0.0118	-805.9827	0.2872	-0.0112	-805.9520	0.2817	-0.0140	-7.2	0.5	2.0	19.3	-3.5	-1.8
1e	-421.2936	0.1374	-0.0157	-265.0739	0.0484	-0.0118	-686.3869	0.1883	-0.0213	-686.3457	0.1817	-0.0293	-12.2	1.5	3.9	25.9	-4.1	-5.0
1f	-537.8249	0.1987	-0.0061	-265.0739	0.0484	-0.0118	-802.9133	0.2488	-0.0127	-802.8725	0.2419	-0.0175	-9.1	1.0	3.3	25.6	-4.3	-3.0
2A	-846.4018	0.3932	-0.0082	-265.0739	0.0484	-0.0118	-1111.4960	0.4434	-0.0142	-1111.4654	0.4382	-0.0194	-12.7	1.1	3.7	19.2	-3.3	-3.3
2B	-846.4018	0.3932	-0.0082	-265.0739	0.0484	-0.0118	-1111.4969	0.4439	-0.0134	-1111.4713	0.4388	-0.0162	-13.3	1.4	4.2	16.1	-3.2	-1.8
2C	-846.4018	0.3932	-0.0082	-265.0739	0.0484	-0.0118	-1111.4966	0.4430	-0.0123	-1111.4506	0.4365	-0.0190	-13.1	0.9	4.9	28.8	-4.1	-4.2
2D	-846.4018	0.3932	-0.0082	-265.0739	0.0484	-0.0118	-1111.4889	0.4428	-0.0176	-1111.4602	0.4375	-0.0256	-8.3	0.7	1.5	18.0	-3.3	-5.0
2E	-846.4018	0.3932	-0.0082	-265.0739	0.0484	-0.0118	-1111.4883	0.4426	-0.0157	-1111.4559	0.4380	-0.0233	-7.9	0.6	2.8	20.4	-2.9	-4.8
2a	-594.9666	0.2343	-0.0075	-265.0739	0.0484	-0.0118	-860.0524	0.2836	-0.0173	-860.0209	0.2783	-0.0230	-7.5	0.6	1.3	19.8	-3.3	-3.6
2b	-307.9040	0.1953	-0.0106	-265.0739	0.0484	-0.0118	-573.0005	0.2465	-0.0145	-572.9644	0.2404	-0.0216	-14.2	1.7	5.0	22.7	-3.8	-4.4
2c	-307.9040	0.1953	-0.0106	-265.0739	0.0484	-0.0118	-572.9957	0.2461	-0.0193	-572.9538	0.2395	-0.0231	-11.2	1.5	2.0	26.3	-4.1	-2.4
2d	-307.9040	0.1953	-0.0106	-265.0739	0.0484	-0.0118	-572.9984	0.2462	-0.0165	-572.9649	0.2408	-0.0240	-12.8	1.5	3.7	21.0	-3.4	-4.7
2e	-307.9040	0.1953	-0.0106	-265.0739	0.0484	-0.0118	-572.9940	0.2461	-0.0200	-572.9580	0.2410	-0.0293	-10.1	1.4	1.6	22.6	-3.2	-5.9

References:

1. M. S. Searle, M. S. Westwell and D. H. Williams, Application of a Generalized Enthalpy-Entropy Relationship to Binding Cooperativity and Weak Associations in Solution, *J Chem Soc Perk T 2*, 1995, DOI: DOI 10.1039/p29950000141, 141-151.
2. J. M. Fox, M. X. Zhao, M. J. Fink, K. Kang and G. M. Whitesides, The Molecular Origin of Enthalpy/Entropy Compensation in Biomolecular Recognition, *Annu Rev Biophys*, 2018, **47**, 223-250.
3. U. Ryde, A fundamental view of enthalpy-entropy compensation, *Medchemcomm*, 2014, **5**, 1324-1336.