Electronic Supplementary Material (ESI) for Dalton Transactions.

Supporting information for: Hunting the human DPP III active conformation: combined thermodynamic and QM/MM calculations

Antonija Tomić[a] and Sanja Tomić[b]

[a] Division of Physical Chemistry, Ruđer Bošković Institute, Bijenička 54, 10 000 Zagreb, Croatia. E-mail: atomic@irb.hr.

[b] Division of Physical Chemistry, Ruđer Bošković Institute, Bijenička 54, 10 000 Zagreb, Croatia. E-mail: sanja.tomic@irb.hr.

Supplementary Figures

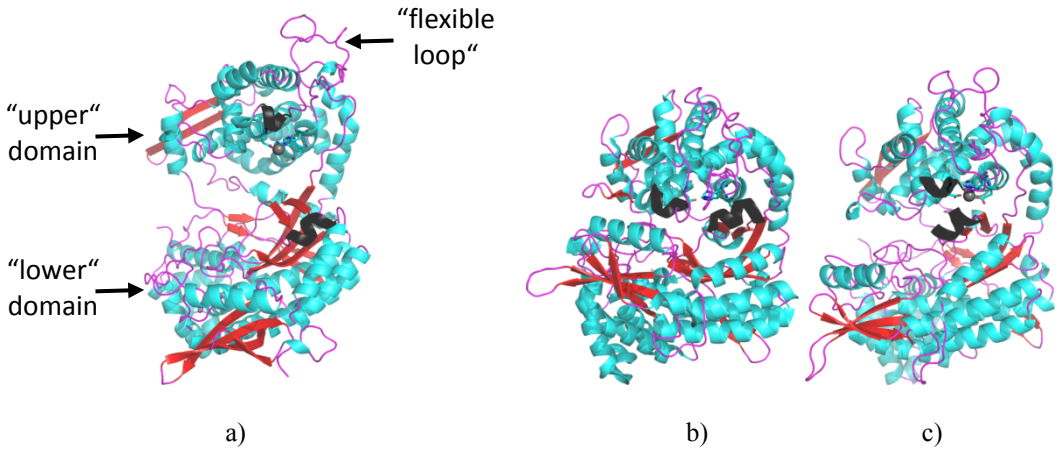
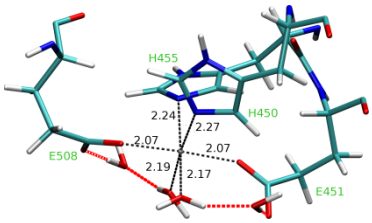
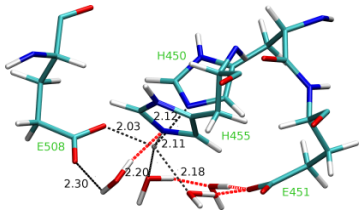


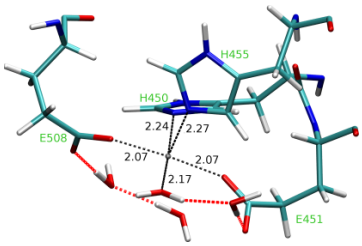
Figure S1. Three different DPP III conformations: a) „open X-ray" (pdb code: 3FVY), b) „closed X-ray" (pdb code: 3T6B) and c) „closed MD" (structure obtained after 100 ns of MD simulations starting from the "open X-ray" structure). Two helices are colored black in order to highlight the difference between these structures. Positions of protein "upper" and "lower" domain are indicated on the "open" DPP III structure.



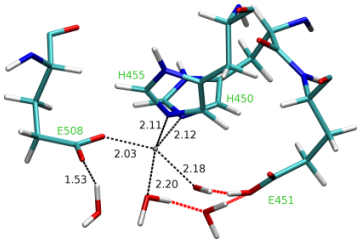
*o*WT, **6[1:2:0]**



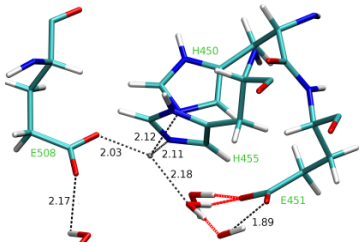
*o*WT, **5[0:2:0]**



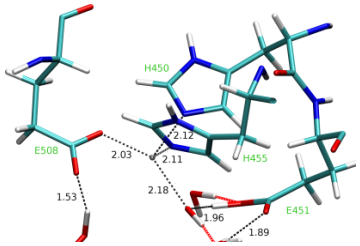
*o*WT, **5[1:1:0]**



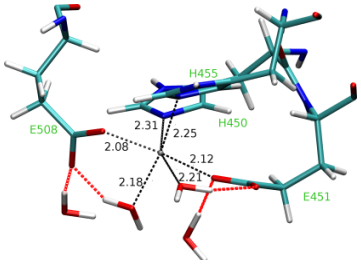
*o*WT, **5[0:1:1]**



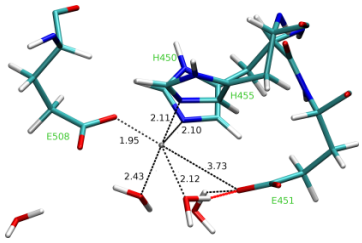
*o*WT, **4[0:1:0]**



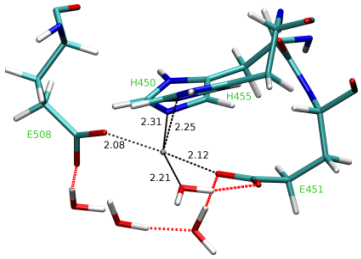
*o*WT, **4[0:0:1]**



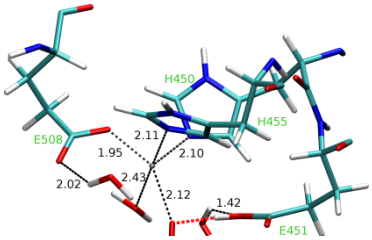
*c*WT, **6[1:2:0]**



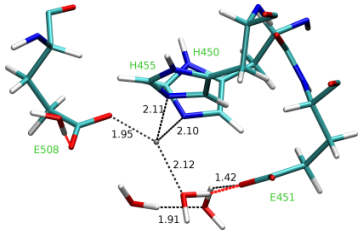
*c*WT, **5[0:2:0]**



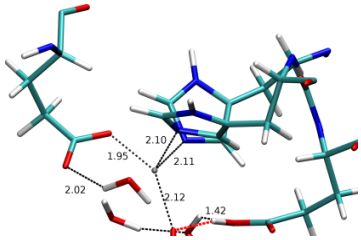
*c*WT, **5[1:1:0]**



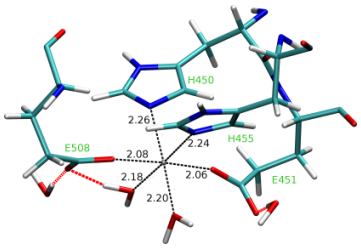
*c*WT, **5[0:1:1]**



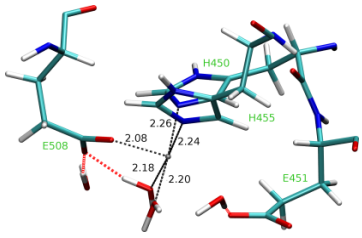
*c*WT, **4[0:1:0]**



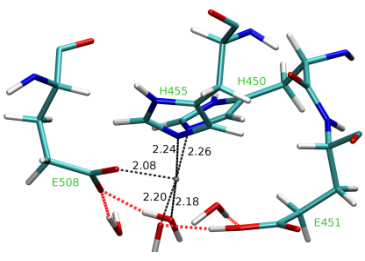
*c*WT, **4[0:0:1]**



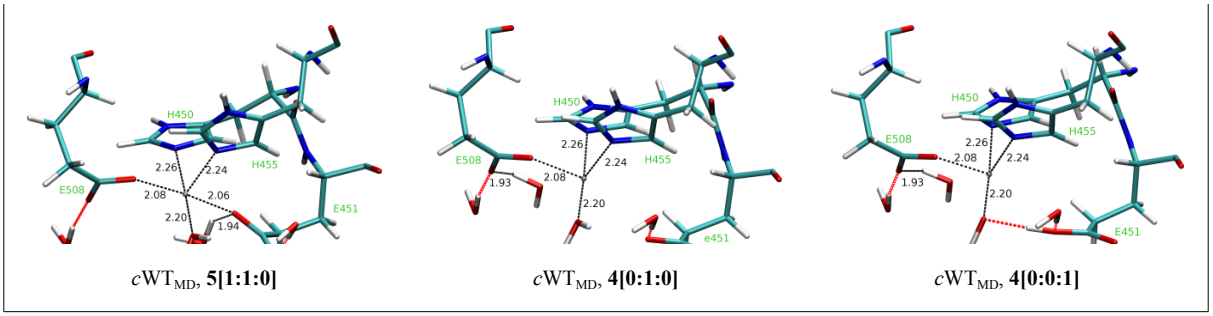
*c*WT$_{MD}$, **6[1:2:0]**



*c*WT$_{MD}$, **5[0:2:0]**



*c*WT$_{MD}$, **5[0:1:1]**



cWT$_{MD}$, **5[1:1:0]** cWT$_{MD}$, **4[0:1:0]** cWT$_{MD}$, **4[0:0:1]**

Figure S2. The initial Zn^{2+} coordination geometries in the systems (enzyme + 52 water molecules) used in ONIOM QM/MM study. The system identifying code consist of the complex name, the metal ion coordination number (the first bolded number) and the numbers in square brackets indicating whether E451, water molecule(s) and OH^- ion, respectively, coordinates the zinc ion (0 means 'NO', while numbers ≥1 indicate the number of corresponding oxygen atoms types coordinating the zinc ion). Distances are in Å. Hydrogen bonds, represented by dashed red lines, were determined using 180 + 30° angle and 3.0 Å distance cutoff values.

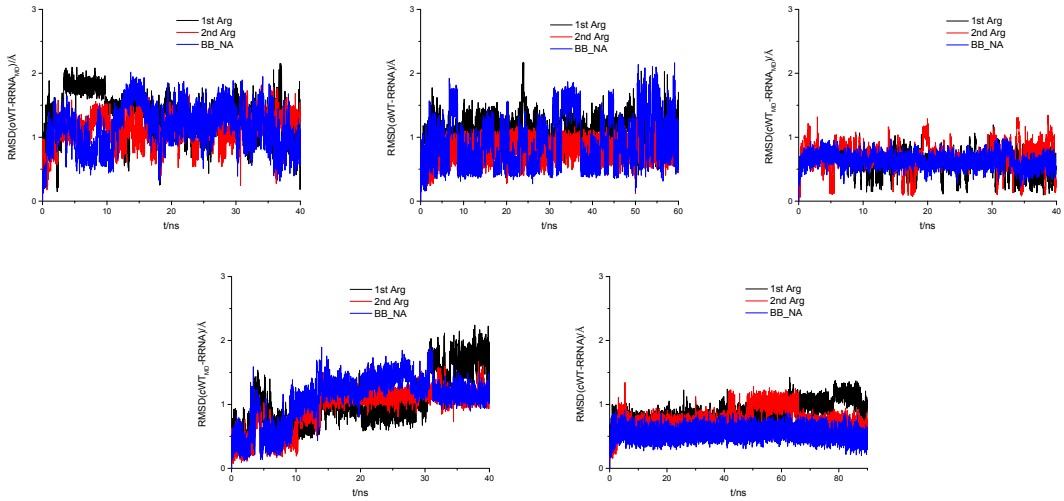
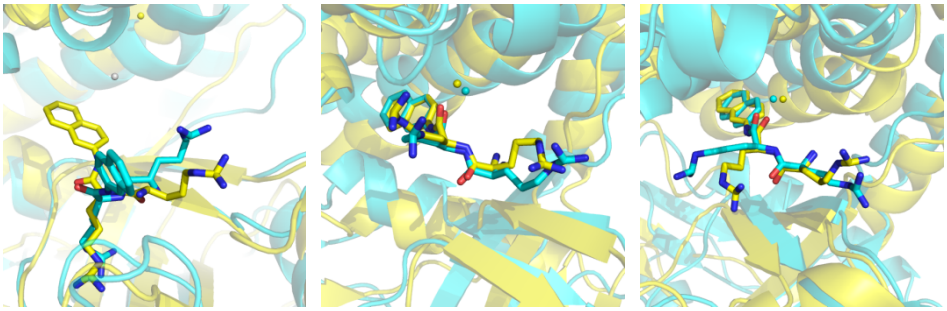
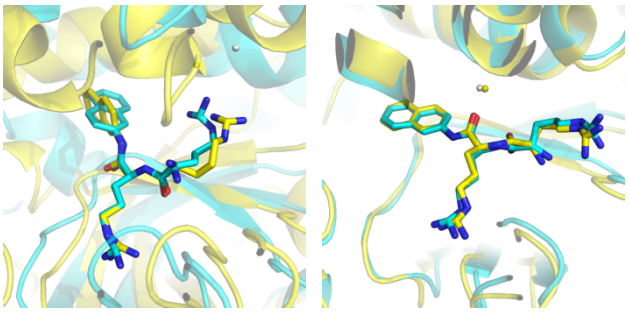


Figure S3. RMSD profiles of the substrate backbone and naphthyl group (bluen), 1st (black) and 2nd (red) arginine residue side chains obtained from MD simulations of all five DPP III-RRNA complexes.



a) b) c)



d) e)

Figure S4. Substrate (RRNA) binding modes obtained after 30ns (or 50 and 80 ns) of MD simulation performed at 300 K (the zinc ion, sphere, and substrate carbon atoms, sticks, colored yellow) and after 10 ns at 350 K (the zinc ion, sphere, and substrate carbon atoms, sticks, are colored cyan); a) *o*WT-RRNA, b) *o*WT-RRNA$_{MD}$, c) *c*WT$_{MD}$-RRNA$_{MD}$, d) *c*WT$_{MD}$-RRNA and e) *c*WT-RRNA complexes.

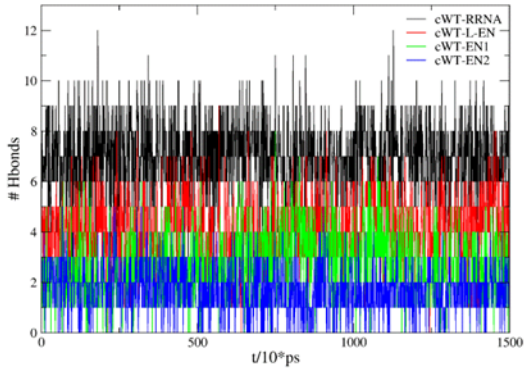


Figure S5. Number of hydrogen bonds (# Hbonds) that substrate, Arg-Arg-2-naphthylamide (RRNA), Leu-enkephalin (L-EN), endomorphin-1 (EN1) and endomorphin-2 (EN2), molecule makes with protein in its "closed X-ray" form (*c*WT). Period from 15th till 30th ns of MD simulations was considered.

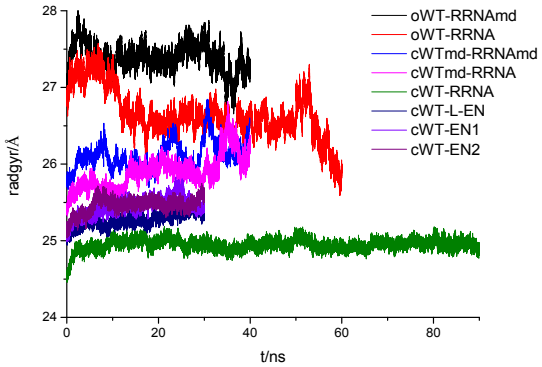


Figure S6. Radius of gyration profile obtained from DPP III-substrate complexes MD simulations. RRNA, L-EN, EN1 and EN2 are substrate molecules Arg-Arg-2-naphthylamide, Leu-enkephalin, endomorphin-1 and endomorphin-2, respectively. Residues 462 to 490, belonging to flexible loop from "upper" protein domain (see Fig.S1) were omitted from calculations.

Supplementary Tables

Table S1. Description of the structures and corresponding codes used in the article text.

structure code	description
unbound DPP III	
"open X-ray" or *o*WT	"open" human DPP III crystal structure, pdb code: 3FVY
"closed X-ray" or *c*WT	"closed" human DPP III crystal structure, pdb code: 3T6B
"closed MD" or cWT_{MD}	"closed" human DPP III structure obtained after 100 ns of MD simulations starting from the "open" human DPP III crystal structure
substrates	
RRNA	Arg-Arg-2-naphthylamide
L-EN	Leu-enkephalin
EN1	endomorphin-1
EN2	endomorphin-2
ligated DPP III	
*o*WT-RRNA	RRNA molecule bound into the "open X-ray" DPP III form with its first two arginine residues forming a β-strand bounded to the five-stranded β-core of the enzyme in the antiparallel fashion
$oWT\text{-}RRNA_{MD}$	RRNA molecule bound into the "open X-ray" DPP III form in a way determined by docking and steered MD simulations
$cWT_{MD}\text{-}RRNA_{MD}$	RRNA molecule bound into the "closed" DPP III form obtained after 72 of MD simulations starting from the "open X-ray" structure in a way determined by docking and steered MD simulations
cWT_{MD}-RRNA	RRNA molecule bound into the "closed" DPP III form obtained after 72 of MD simulations starting from the "open X-ray" structure with its first two arginine residues forming a β-strand bounded to the five-stranded β-core of the enzyme in the antiparallel fashion
*c*WT-RRNA	Substrate (RRNA, L-EN, EN1 and EN2) molecule bound into the "closed X-ray" DPP III form with its first two/three amino acid residues forming a β-strand bounded to the five-stranded β-core of the DPPIII in the antiparallel fashion
*c*WT-L-EN	
*c*WT-EN1	
*c*WT-EN2	

Table S2. Residue based decomposition of the MM-GBSA enthalpy contribution to the binding free energies calculated for Arg-Arg-2-naphthylamide (RRNA), Leu-enkephalin (L-EN), endomorphin-1 (EN1) and endomorphin-2 (EN2). Energies were calculated for the period from the 15th till 30th ns of MD simulations, considering every 100th ps. Shown are energy values that contribute more than ± 1 kcal mol^{-1}. Residues that stabilized all substrates bound in "closed X-ray" protein structure (*c*WT) are given in bold.

*o*WT-RRNA$_{MD}$		*c*WT$_{MD}$-RRNA$_{MD}$		*o*WT-RRNA		*c*WT$_{MD}$-RRNA		*c*WT-RRNA	
residue	Δ*H*/ kcal mol^{-1}	residue	Δ*H*/ kcal mol^{-1}	residue	Δ*H*/ kcal mol^{-1}	residue	Δ*H*/ kcal mol^{-1}	residue	Δ*H*/ kcal mol^{-1}
E451	−3.37	E451	−5.51	E329	−4.75	E316	−5.41	**E316**	−8.57
H568	−2.32	D372	−5.06	E316	−4.62	E329	−4.18	D396	−6.10
D372	−1.68	F404	−3.72	E327	−3.13	E508	−2.10	**N394**	−4.26
V447	−1.44	H568	−1.60	N391	−2.98	A388	−1.94	E329	−4.21
N406	−1.18	E316	−1.49	P387	−1.89	G389	−1.89	E327	−4.16
E508	−1.07	N406	−1.43	I390	−1.55	P387	−1.86	D496	−4.11
Zn	3.51	V447	−1.31	G389	−1.45	H568	−1.73	**Zn**	−3.66
		H455	−1.30	Y318	−1.39	Q566	−1.10	**Y318**	−3.64
		N391	−1.20	G385	−1.28	F443	−1.02	**N391**	−2.62
		H450	−1.20			Zn	1.85	**H568**	−2.17
		Zn	4.26					**H455**	−2.05
								E507	−1.98
								P387	−1.90
								I390	−1.35
								A388	−1.18
								S317	−1.05
								R399	2.36
								K629	1.83

*c*WT-L-EN		*c*WT-EN1		*c*WT-EN2	
residue	Δ*H*/ kcal mol^{-1}	residue	Δ*H*/ kcal mol^{-1}	residue	Δ*H*/ kcal mol^{-1}
R669	−4.01	**E316**	−4.39	**E316**	−3.09
E316	−3.64	**H568**	−3.50	**N391**	−2.76
N394	−2.74	**N391**	−2.65	**H455**	−2.43
F443	−2.48	**P387**	−2.56	**P387**	−2.38
H455	−2.38	**Y318**	−2.31	**Y318**	−2.17
N391	−2.30	**Zn**	−2.18	**Zn**	−1.90
H568	−2.14	**H455**	−2.16	**N394**	−1.87
P387	−2.01	**N394**	−2.15	**H568**	−1.79
R572	−1.63	R572	−1.68	V447	−1.71
Zn	−1.37	V447	−1.67	G389	−1.60
A388	−1.27	**I390**	−1.50	**I390**	−1.59
E507	−1.24	**A388**	−1.49	**A388**	−1.53
Y318	−1.22	E512	−1.48	F443	−1.51
I390	−1.17	G389	−1.34	H450	−1.47
G389	−1.15	F109	−1.28	F109	−1.17
F109	−0.99	H450	−1.17	**E507**	−1.10
		E507	−1.11		
		F443	−0.97		

Table S3. Energies[a] (in kcal mol^{-1}) of the partly solvated DPP III (52 water molecules) obtained by *2-layer* ONIOM optimization (M06 or B3LYP applying mechanical and electrostatic embedding protocols, ME and EE, respectively). Only the energies obtained for the same protein conformation are compared. Two different basis sets; 6-31+g(d,p) for the H, N, C and O atoms, and LANL2DZ-ECP for the Zn^{2+} atom, were used. The MM part was treated by the AMBER force field (param96). Numbers show in square brackets indicate whether E451, water molecule(s) and OH^- ion, respectively, coordinates the zinc ion and how (0 means 'NO', while numbers ≥1 indicate the number of corresponding oxygen atoms coordinating the zinc ion). Calculations with MM water molecules fixed during minimization are shown in italic. MM waters were fixed as there occasional movement in QM region in the "open X-ray" DPP III structure (where there are lots of water molecules found in the proximity of the Zn^{2+} ion) unable the optimization to converged.

initial structures		M06 with ME		M06 with EE		B3LYP with EE	
name	CN[E451:WAT:OH]	CN[E451:WAT:OH]	ΔE^{ONIOM}	CN[E451:WAT:OH]	ΔE^{ONIOM}	CN[E451:WAT:OH]	ΔE^{ONIOM}
*c*WT	6[1:2:0]	6[1:2:0]	0.00	6[1:2:0]	0.00	6[1:2:0]	2.38
	5[0:2:0]	5[0:2:0]	15.31	6[1:2:0]	7.10	5[0:2:0]	3.71
	5[1:1:0]	5[1:1:0]	9.43	5[1:1:0]	4.60	5[1:1:0]	0.00
	5[0:1:1]	6[0:3:0]	35.96	6[0:3:0]	31.39	5[0:2:0]	23.86
	4[0:1:0]	4[0:0:1]	8.82	4[0:0:1]	12.29	4[0:0:1]	8.85
	4[0:0:1]	4[0:0:1]	5.07	4[0:0:1]	8.84	4[0:0:1]	5.57
*c*WT$_{MD}$	6[1:2:0]	6[1:2:0]	0.00	6[1:2:0]	0.00		
	5[0:2:0]	6[1:2:0]	1.66	6[1:2:0]	0.00		
	5[0:1:1]	5[0:1:1]	17.20	5[0:2:0]	17.31		
	5[1:1:0]	5[1:1:0]	14.92	5[1:1:0]	13.96		
	4[0:1:0]	-	-	6[1:2:0]	0.00		
	4[0:0:1]	5[0:1:1]	16.99	5[1:1:0]	11.85		
*o*WT	6[1:2:0]	6[1:2:0]	0.00	6[1:2:0]	3.78		
	5[0:2:0]	6[0:3:0]	28.26	*6[0:3:0]*	*36.69*		
	5[1:1:0]	*5[1:1:0]*	*27.03*	*5[1:1:0]*	*28.12*		
	5[0:1:1]	5[0:2:0]	31.01	*5[0:2:0]*	*42.81*		
	4[0:1:0]	6[0:3:0]	19.32	6[0:3:0]	24.42		
	4[0:0:1]	6[1:2:0]	7.64	6[1:2:0]	0.00		

[a] $E^{ONIOM}=E_{MM}(S)+E_{QM}(SM)-E_{MM}(SM)$

Details of correlation coefficient calculations:

The correlation coefficient was calculated between theoretically determined fraction of $K_d(1)/K_d(2)$ values and experimentally determined $K_m(1)/K_m(2)$ fraction. $K_d(1)/K_d(2)$ values were calculated using formula $K_d(1)/K_d(2)=\exp[-\Delta\Delta G_{bind}/(RT)]$, where $\Delta\Delta G_{bind}$ values are calculated relative to complex with lowest binding free energy, $\Delta G_{bind,2}$, according to $\Delta\Delta G_{bind}=\Delta G_{bind,1}-\Delta G_{bind,2}$.