Supplementary Information for: Binding Energy Distribution Analysis Method (BEDAM): Hamiltonian replica exchange with torsional flattening for binding mode prediction and binding free energy estimation

Ahmet Mentes,[†,‡] Nanjie Deng,[†,‡] R. S. K. Vijayan,[†,‡] Junchao Xia,[†,‡] Emilio Gallicchio,[¶] and Ronald M. Levy[*,†,‡]

Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122, United States, Center for Biophysics and Computational Biology and Institute for Computational Molecular Science, Temple University, Philadelphia, Pennsylvania 19122, United States, and Department of Chemistry, Brooklyn College, The City University of New York, Brooklyn, New York 11210, United States

E-mail: ronlevy@temple.edu

[*]To whom correspondence should be addressed

[†]Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122, United States

[‡]Center for Biophysics and Computational Biology and Institute for Computational Molecular Science, Temple University, Philadelphia, Pennsylvania 19122, United States

[¶]Department of Chemistry, Brooklyn College, The City University of New York, Brooklyn, New York 11210, United States

Table S1: **Structural comparison of initial docked structures with the crystal structures**

F. Neg. (25) (Core)	T. Pos. (28) (Core)	F. Neg. (25) (Complete)	T. Pos. (28) (Complete)
7 bad (28 %)	11 bad (39.3 %)	19 bad (76 %)	16 bad (57.1 %)
18 good (72 %)	17 good (60.7 %)	6 good (24 %)	12 good(42.9 %)

Structural analysis for all ligands based on their core scaffolds (the benzoic acid or benzodioxole/dioxine carboxylate core of the ligands) and based on their complete structures are compared to their corresponding initial crystal structures. 7 out of 25 false negatives and 11 out of 28 true positives have initial docked structures with RMSD of larger than 2 Å when only the core of the ligands are considered. This number changes when all ligand atoms considered, so that 19 out of 25 for false negatives and 16 out of 28 true positives have initial docked structures with RMSD of larger than 2 Å. We have compared structurally core scaffold vs. all ligand atoms because of the understanding of how ligand structures differ from their crystal structures when ligand functional groups are considered.

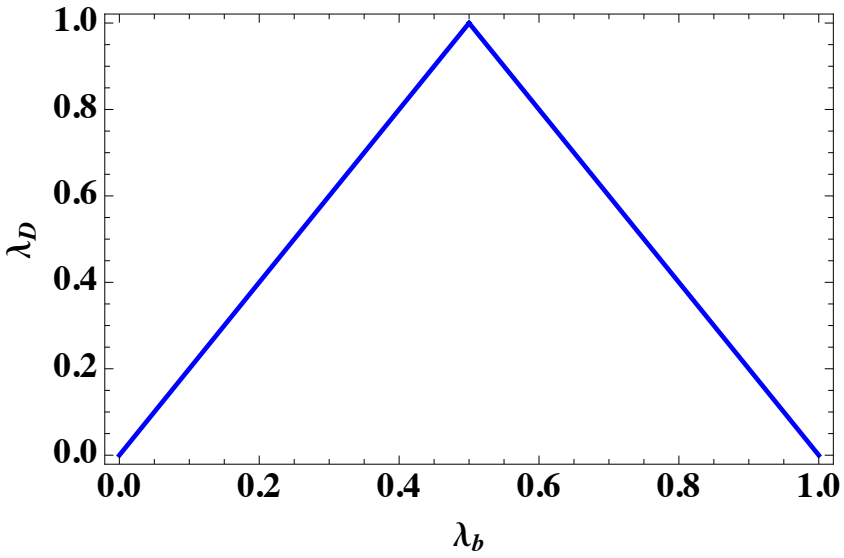

Figure S1: Schematic representation of relation between binding parameter λ_b and torsional flattening parameter λ_D.

Table S2: **Comparision of binding free energies of crystal structure vs docked structure**

False Negatives (25)	True Positives (28)
22 Δ G (crystal) $<$ Δ G (docked)	21 Δ G (crystal) $<$ Δ G (docked)
3 Δ G (crystal) $>$ Δ G (docked)	7 Δ G (crystal) $>$ Δ G (docked)

Binding free energy estimates from simulations starting from docked structures are compared to their corresponding initial crystal structures. For false negatives, 22 out of 25 (88 %) binding free energy estimates starting from crystal structures have more favorable binding free energy (Δ G) value than the corresponding results starting from docked structures. Only 3 Δ G starting from crystal structures have more positive than when starting from the docked structures. For true positives, 21 out of 28 (75 %) Δ G's starting from crystal structures have more favorable binding free energy (Δ G) value than the corresponding results obtained when starting from docked structures. 7 Δ G's starting from crystal structures have more positive value than the results starting from docked structures. However, 4 out of 7 true positives are still more favorable than -4 kcal/mol. e.g. Δ G (crystal) = -6.57 vs Δ G (docked) = -7.36 kcal/mol. In total, 43 out of 53 (81 %) free energy estimates based on choosing crystal structure as the initial pose are more favorable than the free energy estimates starting from docked structures, and also 47 out of 53 (88.7 %) Δ G's of crystal structures have binding free energy estimation better -3.0 kcal/mol (in Table S3).

Table S3: **Classification of binding free energy of crystal structures**

	False Negatives (25)	True Positives (28)
Δ G $<$ -3 kcal/mol	22 (88 %)	25 (89.3 %)
Δ G $>$ -3 kcal/mol	3 (12 %)	3 (10.7 %)

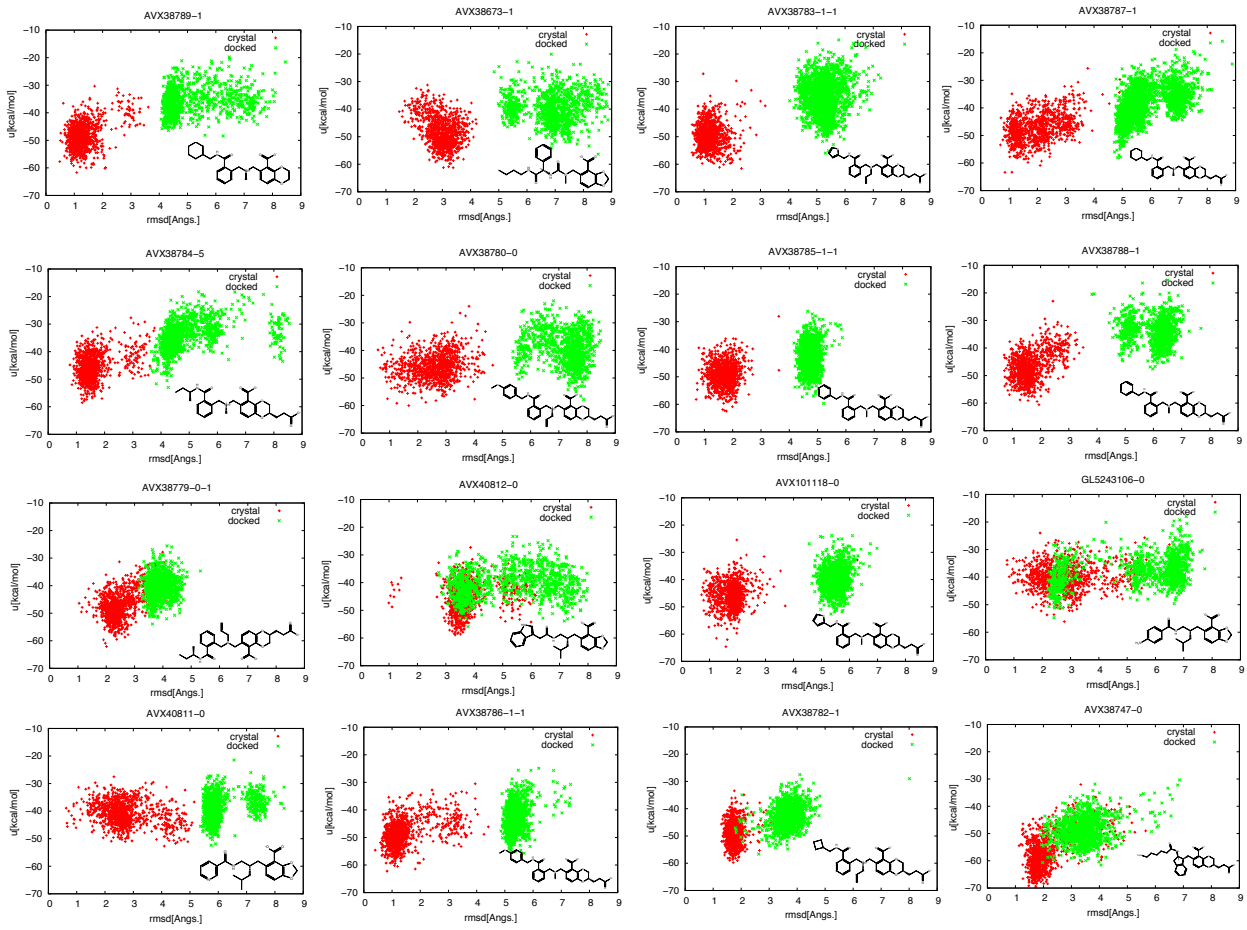

Figure S2: Funnel type binding energy landscapes of 16 false negatives from BEDAM simulations at λ_b=1; green: starting conformation is docked structure without torsional flattening,red: starting conformation is crystal structure without torsional flattening.

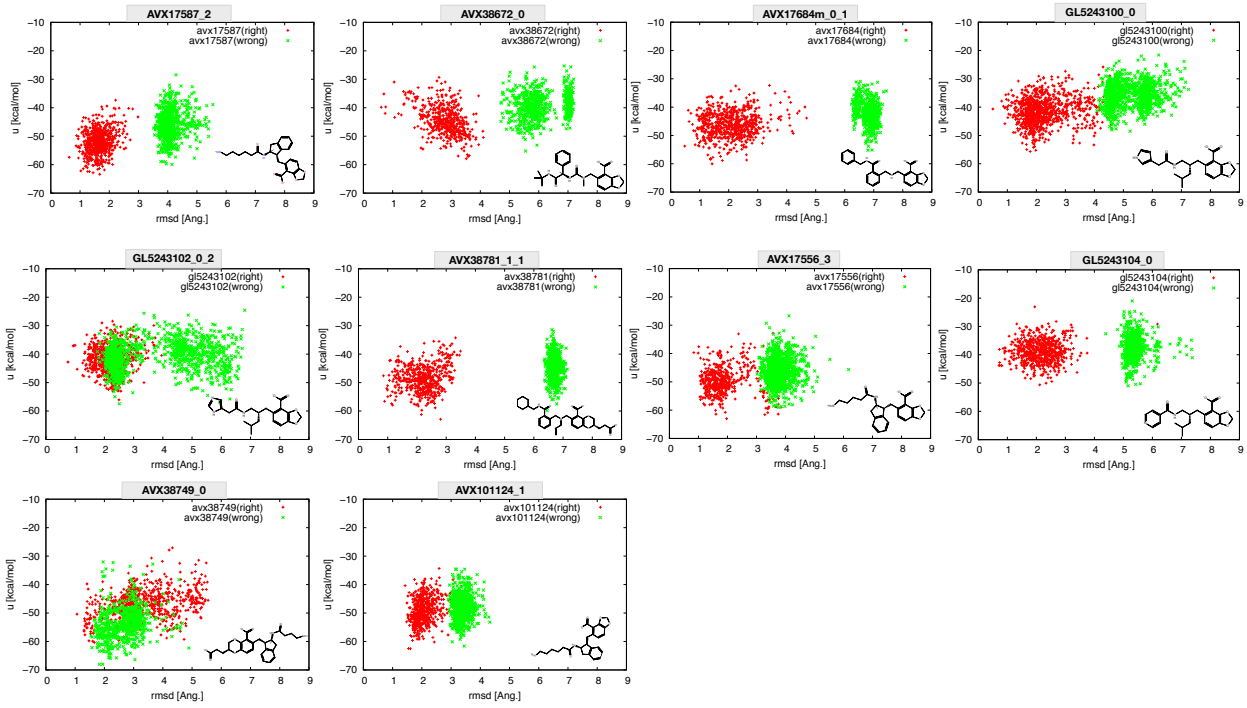

Figure S3: Funnel type binding energy landscapes of 10 true positives from BEDAM simulations at λ_b=1; green: starting conformation is docked structure without torsional flattening, red: starting conformation is crystal structure without torsional flattening.

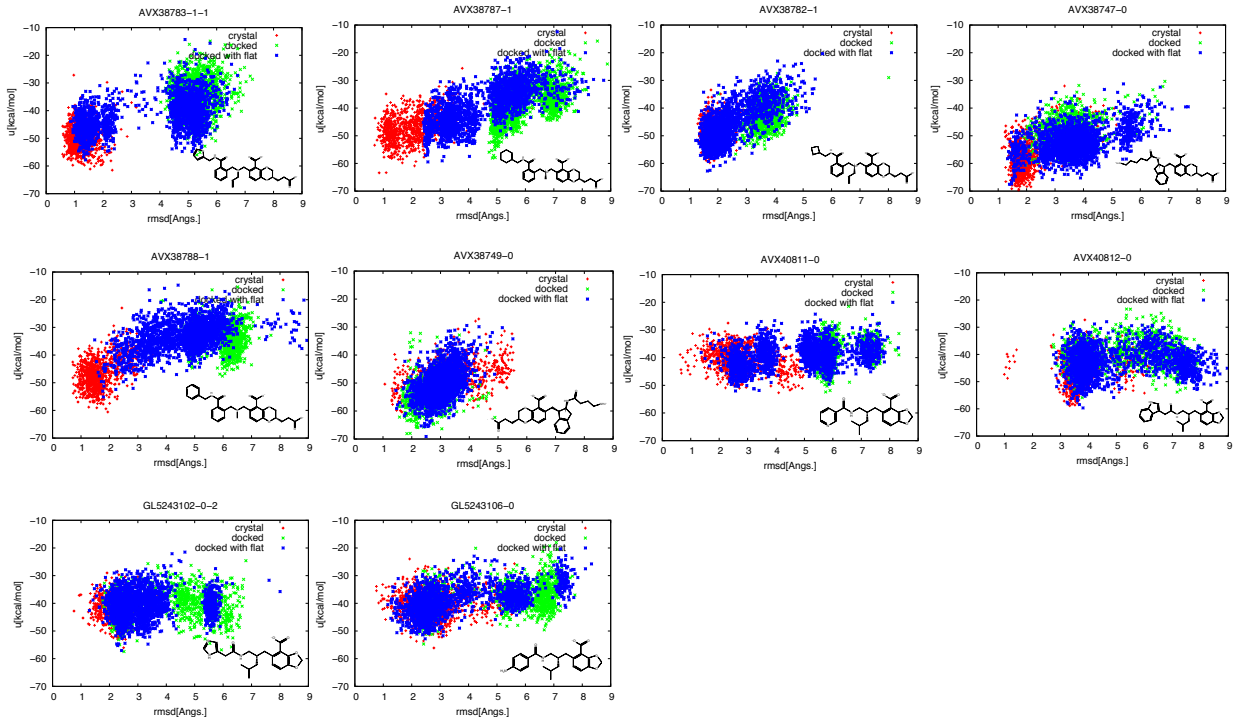

Figure S4: Funnel type binding energy landscapes from BEDAM simulations with and without flattening at λ_b=1; green: starting conformation is docked structure without torsional flattening, red: starting conformation is crystal structure without torsional flattening and blue: starting conformation is docked structure with torsional flattening.

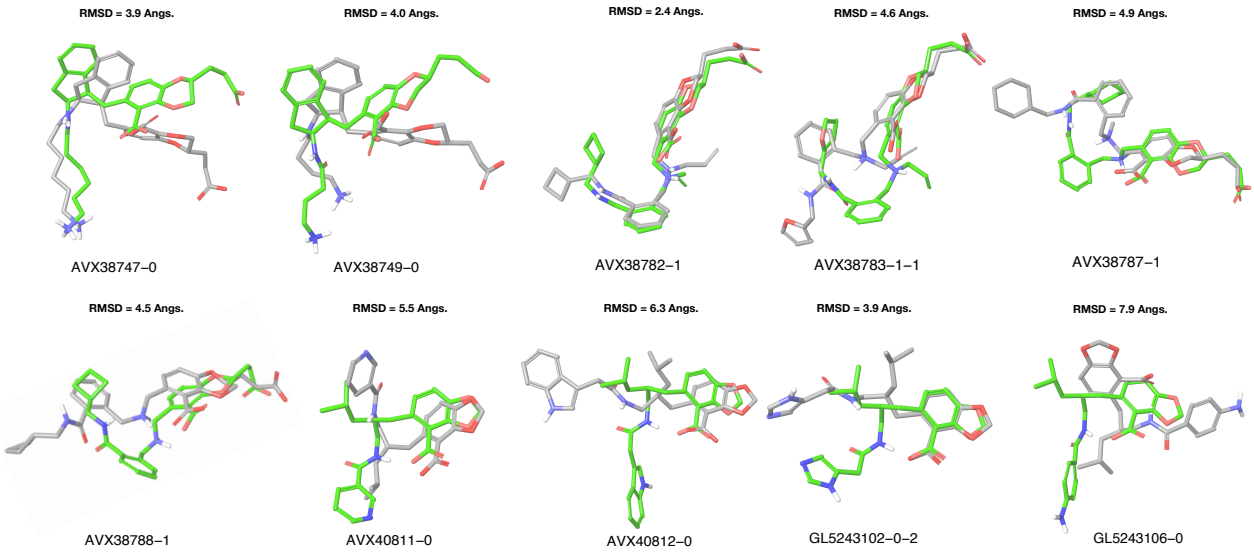

Figure S5: The RMSD between starting docked pose (grey) and the validation (crystallographic) pose (green).

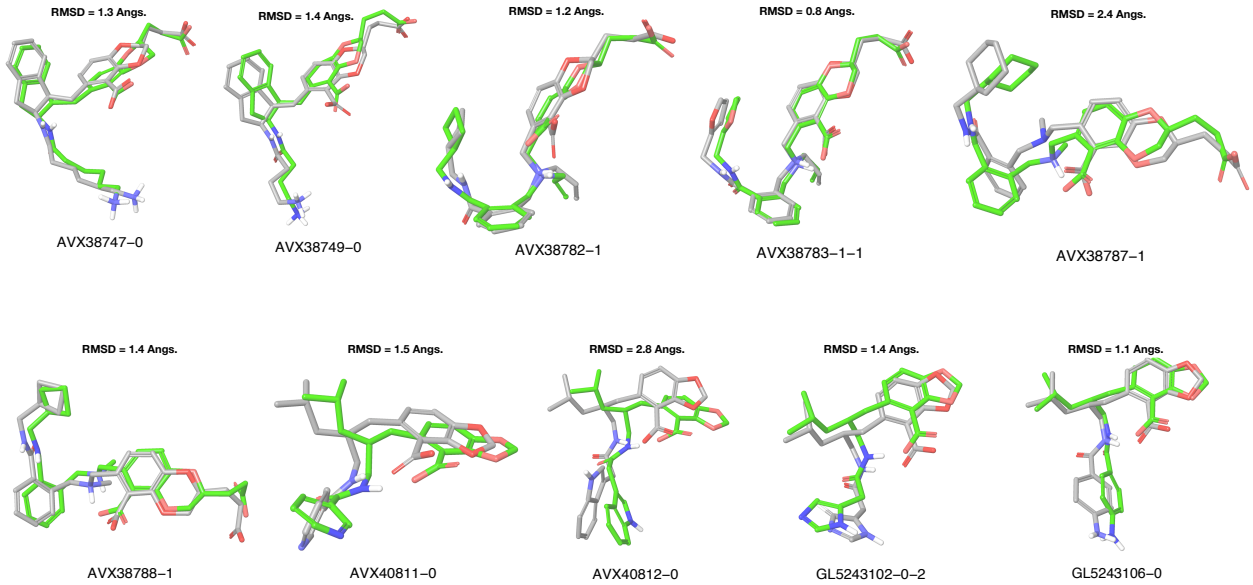

Figure S6: The RMSD between predicted pose (grey) and the validation (crystallographic) pose (green) after flattening applied.

Advantages of BEDAM sampling (translational and rotational) of ligand with respect to receptor

As mentioned in the main text, the BEDAM method samples efficiently the rotational and translational coordinates of the ligand with respect to the receptor in the binding site. In order to represent the rotational and translational sampling of the ligand with respect to the receptor, two angles (the pitch and in-plane rotational angles pairs (Θ_n, Θ_p)) are defined in the original BEDAM paper[1] in which the atoms of aromatic ring of ligand scaffold are used to compute these angles. The pitch angle (Θ_n) is an angle between normal vectors to the planes of the aromatic ring of reference ligand scaffold and of ligand scaffold from the simulations. The in-plane rotational angle (Θ_p) is defined as the angle between vector "OC" and the projection of "O$C^{'}$" segment on the the plane of the ring of the reference ligand pose. "O" is the centroid of the heavy atoms of the ring and "C" and "$C^{'}$" are coordinates of atom in the reference ligand pose and poses from the simulations, respectively (for the detailed description of the pitch and in-plane rotational angles, see the Figure S11 and the reference[1]).

In order to show the rotational and translational sampling of the ligand with respect to the receptor when the protein and ligand is fully coupled, weakly coupled and uncoupled, we have computed the pitch and in-plane rotational angles pairs (Θ_n, Θ_p) for the scaffold of ligand AVX38783-1-1 bound to the binding site of HIV-1 Integrase from different BEDAM simulations. In Figure S7, we showed the sampling of pitch and in-plane rotational angles pairs (Θ_n, Θ_p), **a**) at $\lambda_b = 1$ (receptor and ligand are fully coupled), **b**) at $\lambda_b = 0.5$ (weakly coupled), **c**) at $\lambda_b = 0$ (receptor and ligand are uncoupled) starting from docked structure without torsional flattening (green), starting from crystal structure without torsional flattening (red), and starting from docked structure with torsional flattening (blue). In Figure S7-a, closer view of sampling of pitch and in-plane rotational angles are shown as inset figures. For the sampling of the angles at weakly coupled and uncoupled λ_b values, only data from BEDAM

simulations with torsional flattening has been showed because it is previously reported that the BEDAM is good at sampling intermolecular degrees of freedom when flattening was not considered.[1]

Two conformational macrostates are identified, corresponding to the crystal structure (red) with $\Theta_p = 0^o \pm 30^o$ and Θ_n between 0^o and 45^o, and another one corresponds to the docked structure (green) with $\Theta_p = 0^o \pm 40^o$ and Θ_n between 40^o and 80^o using the standard BEDAM calculations with no torsional flattening. Indeed, theses two macrostates belong to one macrostate (blue) with $\Theta_p = 0^o \pm 50^o$ and Θ_n between 0^o and 85^o as shown from the calculations with torsional flattening and in the BEDAM simulations with torsional flattening, it was possible to sample these two conformational states because slow conformational degrees of freedom relevant to the binding were accelerated by flattening torsional barriers.

In Figure S8, we have also showed the sampling of translational degrees of freedom of ligand AVX38783-1-1 bound to the binding site of HIV-1 Integrase from BEDAM simulation at $\lambda_b = 1$ (receptor-ligand are fully coupled) and at $\lambda_b = 0$ (receptor-ligand are not coupled) using distance between center of mass of ligand and center of mass of receptor. It indicates that BEDAM method is able to sample translational degrees of freedom of ligand with respect to the receptor in the binding volume V_{site} defined by flat-harmonic potential using center of masses of two species (receptor and ligand).

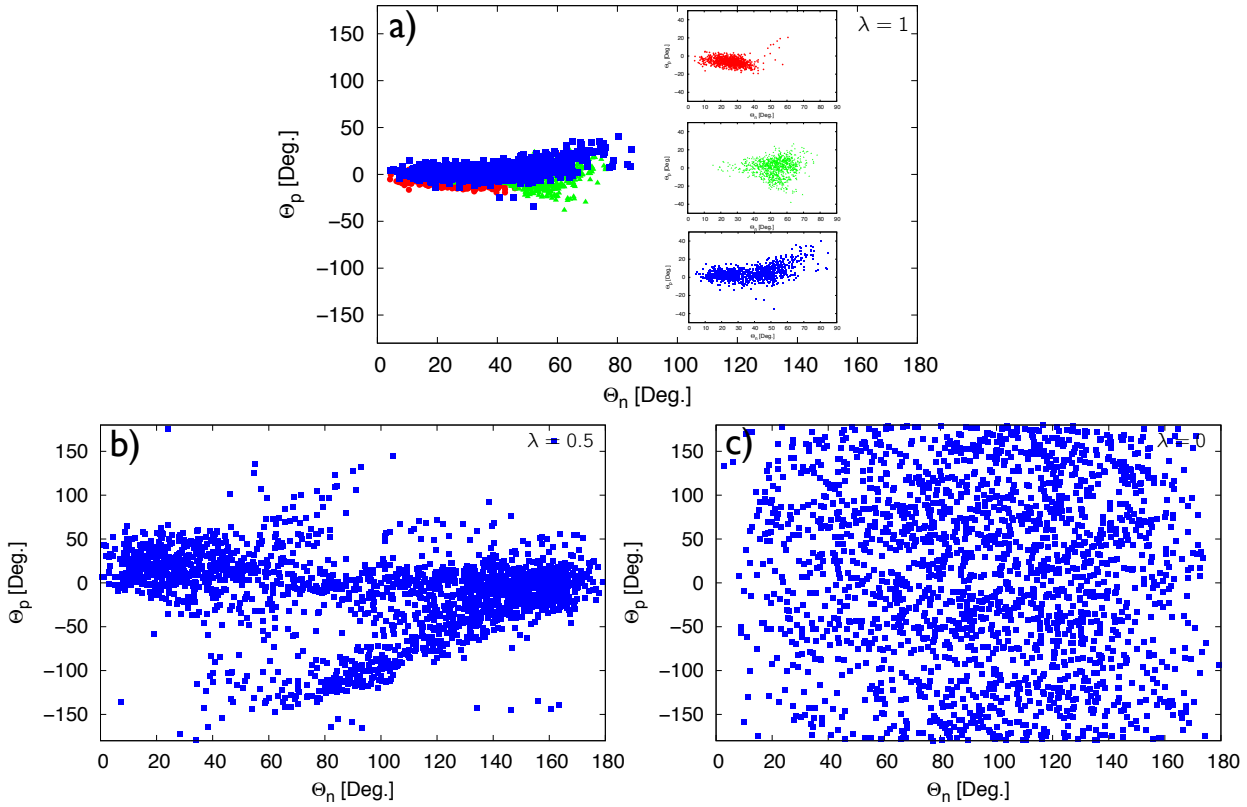

Figure S7: Sampling of pitch and in-plane rotational angles pairs (Θ_n, Θ_p) for the scaffold of ligand AVX38783-1-1 bound to the binding site of HIV-1 Integrase from BEDAM simulations. **a**) at $\lambda_b = 1$ (receptor-ligand are fully coupled), **b**) at $\lambda_b = 0.5$, **c**) at $\lambda_b = 0$ (receptor-ligand are uncoupled); green: starting conformation is docked structure without torsional flattening, red: starting conformation is crystal structure without torsional flattening, blue: starting conformation is docked structure with torsional flattening. In a, inset figures are closer view of sampling of pitch and in-plane rotational angles.

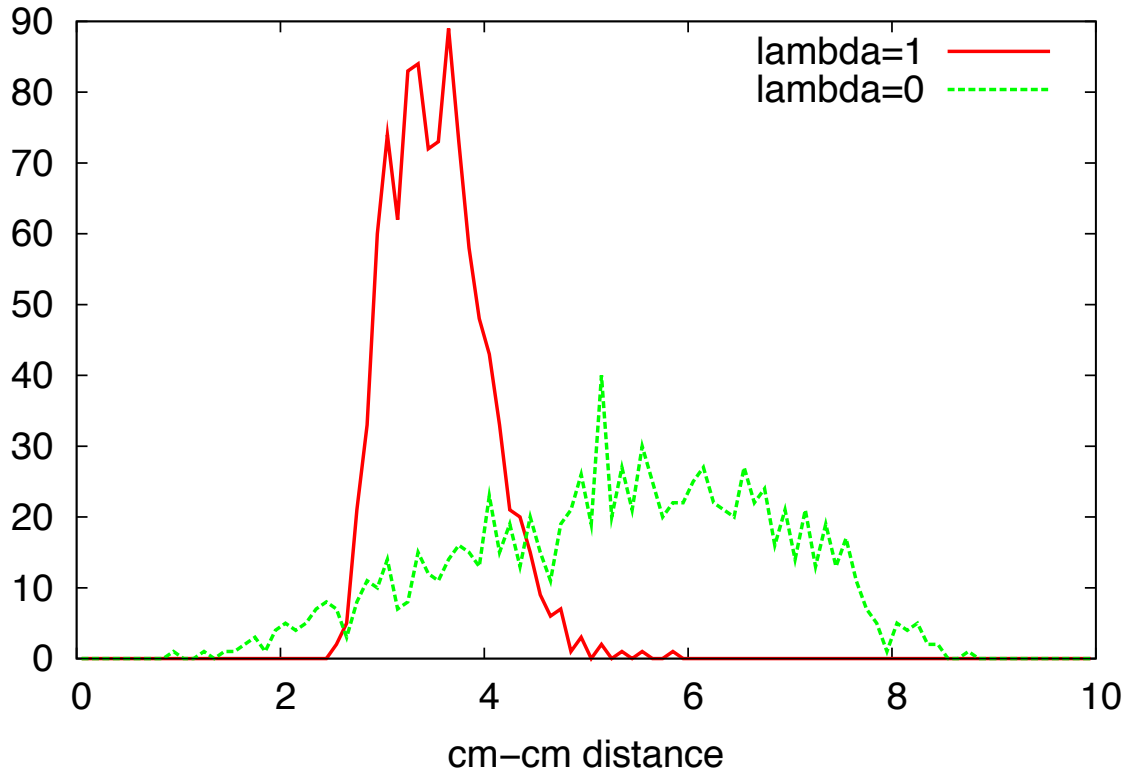

Figure S8: Sampling of translational degrees of freedom of ligand AVX38783-1-1 bound to the binding site of HIV-1 Integrase from BEDAM simulation at $\lambda_b = 1$ (receptor-ligand are fully coupled) and at $\lambda_b = 0$ (receptor-ligand are not coupled) using distance between center of mass of ligand and center of mass of receptor

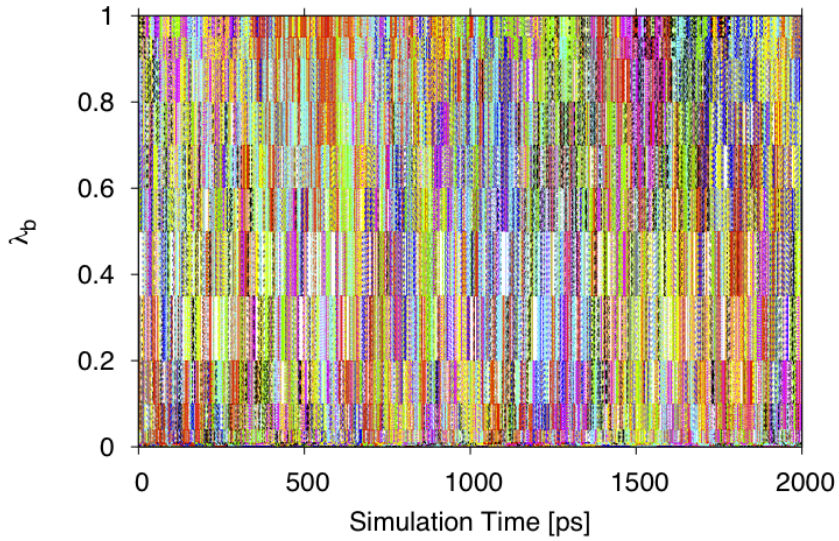

Figure S9: Time evolution of λ_b for each of the HREM replicas with 18 replicas. Each color corresponds to a different replica.

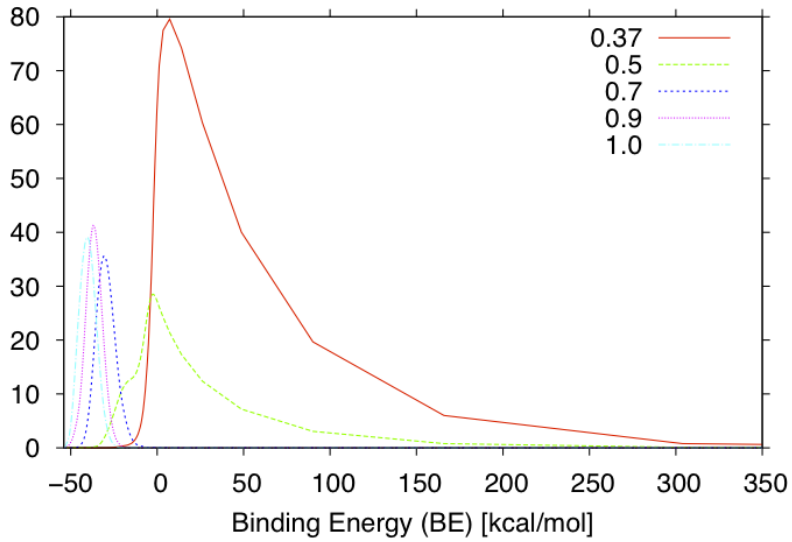

Figure S10: Binding energy distribution at $\lambda_b = 0.37$, 0.5, 0.7, 0.9, and 1. The amount of overlap between these distributions, which is important for accurate binding free energy estimation, is reasonably good with the distribution at the critical value $\lambda_b = 0.5$ acting as a bridge between the other distributions.

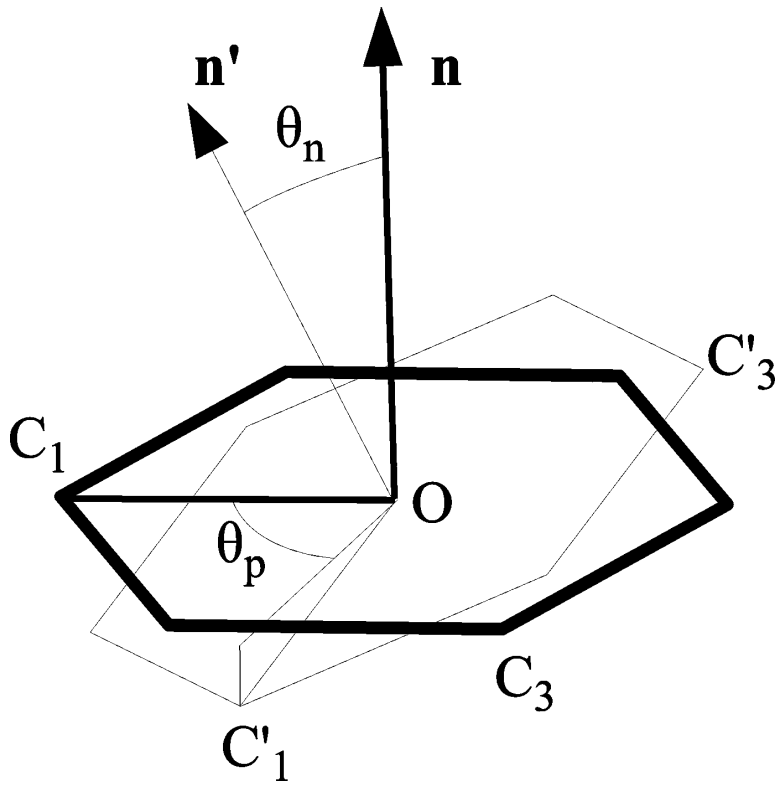

Figure S11: Diagram depicting the definition of the pitch angle Θ_n and in-plane rotation angle Θ_p used in the conformational decomposition analysis (adapted from ref.[1] with permission).

References

[1] E. Gallicchio, M. Lapelosa, R. M. Levy, 2010. Binding Energy Distribution Analysis Method (BEDAM) for Estimation of Protein-Ligand Binding Affinities. *J. Chem. Theory Comput.* 6:2961-2977.