

Decoding the patterns of ubiquitin recognition from free energy simulations - Supporting information

Benjamin Bouvier

Bioinformatics: Structures and Interactions, Bases Moléculaire et Structurales des Systèmes
Infectieux, Univ. Lyon I/CNRS UMR 5086, IBCP, 7 passage du Vercors, Lyon 69367, France.

WHAM consistency tests

Theory

The free energy profile obtained from the WHAM procedure is a function of the chosen reaction coordinate, with all other degrees of freedom integrated out. This implies that these orthogonal degrees of freedom are adequately sampled in each window, which might require much longer timescales than typical window lengths if they feature local minima separated by high free energy barriers. In addition, an individual window trajectory where the system remains in a minimum without visiting any other basins will not show symptoms of insufficient sampling, making the identification of the problem difficult. Zhu and Hummer¹ have proposed that insufficient sampling can be detected by checking the consistency of histograms in neighboring simulation windows: if different states of the orthogonal coordinates are visited in adjacent windows, inconsistent probability distributions of the reaction coordinate will ensue.

Consider a virtual simulation window halfway between two adjacent windows labeled 1 and 2, centered at $d^*=(d_1+d_2)/2$ and with biasing potential $E^*=k(d-d^*)^2$. The corresponding probability distribution $p^*(d)$ can be computed from both $p_1(d)$ and $p_2(d)$:

$$p_i^*(d) = \frac{p_i(d) \exp\left\{\left[E_i(d) - E^*(d)\right]/kT\right\}}{\int_{-\infty}^{+\infty} p_i(x) \exp\left\{\left[E_i(x) - E^*(x)\right]/kT\right\} dx}, \quad i \in \{1,2\}$$

The inconsistency between the two distributions, based on a Kolmogorov-Smirnov test, can be written as:

$$\theta_{1,2} = \sqrt{\frac{N_1 N_2}{N_1 + N_2}} \max_d \left| \int_{-\infty}^d [p_1^*(x) - p_2^*(x)] dx \right|$$

Where the number of independent samples N_1 and N_2 can be evaluated from the variance of the distribution of d and its average over subsets of the corresponding window trajectories (see reference¹ for details).

Results

Figure S1 presents graphs of θ as a function of interpartner distance for all monoUbq/UBA complexes under study, where the problem is most susceptible to appear (in the Ubq₂ complexes, the presence of a linker between the two Ubq subunits strongly limits the accessible conformational space). There is a gradual rise in θ values with the monomer separation distance, hinting at the progressive opening of the volume of conformational space available to the system and a gradual degradation of the quality of the sampling achieved with the nanosecond-order of magnitude simulation windows employed here. However, large values of θ are reached for separation distances larger than 4.5 Å, at which point the free energy profile has in all cases already reached its plateau value. The absence of sharp peaks in the θ plots is an indication that sampling is globally consistent from one window to the next: hysteresis issues that could appear in case of a subdivision of the main complex dissociation pathway into several branches separated by high free energy barriers are not encountered, and the quality of the sampling degrades gracefully rather than abruptly.

Hence, major contamination of the free energy profiles presented here from insufficient sampling are not expected, although the issue clearly exists and would prevent the exploration of larger interpartner separation distances using the same method.

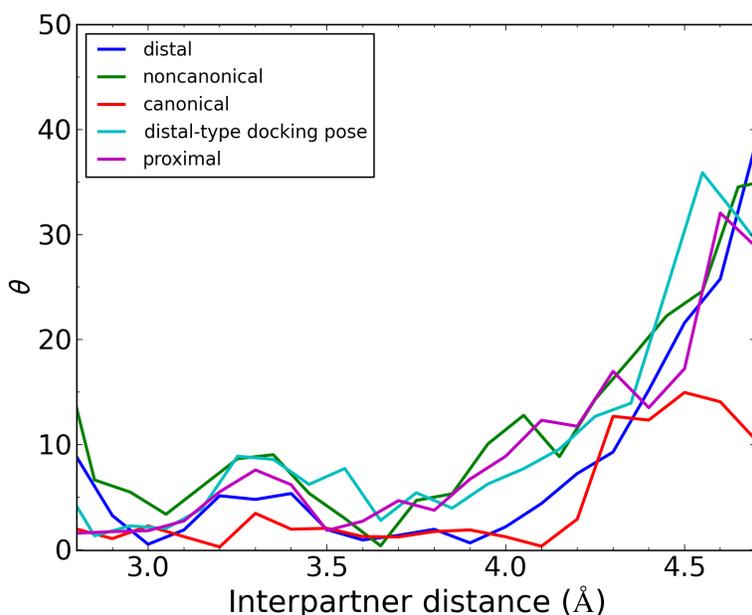


Figure S1: Inconsistency between successive umbrella windows as a function of interpartner distance for all monoUbq/UBA complexes under study (see text for details).

Evaluation of translation-rotation entropy variations

The translational and rotational entropy loss due to the association of two macromolecules can be written as:

$$\Delta G_{\text{rot+trans}} = T\Delta S_{\text{bound-free}} = RT \ln \frac{V_b}{V_f}$$

where V_b is the volume sampled by molecule B in its bound state, in the reference frame of molecule A, and V_f is the volume sampled by the free, isolated molecule B.

Because of the limitations mentioned above, the separation simulations performed in this work do not reach the fully dissociated state where the partners can be considered completely isolated. Thus, it is possible that the translational and rotational entropy lost during the formation of the complex has not been fully recovered. To assess the significance of this, I evaluated the residual translation-rotation entropy loss at maximal interpartner separation. To this effect, V_b was computed from the umbrella sampling window trajectories at this maximal separation, while V_f was computed analytically, as explained below.

Translational contribution

The trajectory frames were aligned by performing a LSQ fit on the C_α atoms of ubiquitin. The position of the center of mass of UBA was recorded at each timestep. The diagonalization of the covariance matrix of center of mass positions yielded three eigenvalues λ_1 , λ_2 and λ_3 . The volume sampled by the center of mass was then written as:

$$V_b^{\text{trans}} = \sqrt{(2\pi)^3 \lambda_1 \lambda_2 \lambda_3}$$

As an alternative to this quasiharmonic analysis, V_b^{trans} can be estimated directly from the minimal and maximal values of the projection of the center of mass position C along each referential axis:

$$V_b^{\text{trans}} = (C_x^{\text{max}} - C_x^{\text{min}})(C_y^{\text{max}} - C_y^{\text{min}})(C_z^{\text{max}} - C_z^{\text{min}})$$

The volume accessible to the translational degrees of freedom of the free UBA was computed from the standard volume of one solute molecule in a 1M solution:

$$V_f^{\text{trans}} = 1660 \text{ \AA}^3$$

Rotational contribution

The principal axes of UBA were recorded at each timestep of the umbrella sampling windows under scrutiny. The rotation matrix between the set of principal axes of UBA at timestep 0 and timestep $n>0$ was determined and converted to three rotation angles θ_x , θ_y and θ_z around the referential axes. Principal component analysis was performed on the set of $(\theta_x, \theta_y, \theta_z)$ vectors, and the rotational volume V_{rot} was computed from the eigenvalues λ_1 , λ_2 and λ_3 of the covariance matrix as for the translational case above:

$$V_{\text{b}}^{\text{rot}} = \sqrt{(2\pi)^3 \lambda_1 \lambda_2 \lambda_3}$$

As an alternative to this quasiharmonic analysis, $V_{\text{b}}^{\text{rot}}$ can be obtained by integrating the Euler angles corresponding to the rotation matrices obtained above over the range of values they span during a simulation:

$$V_{\text{b}}^{\text{rot}} = \int_{\phi_{\text{min}}}^{\phi_{\text{max}}} d\phi \int_{\psi_{\text{min}}}^{\psi_{\text{max}}} d\psi \int_{\theta_{\text{min}}}^{\theta_{\text{max}}} d\cos(\theta)$$

The integration over the entire allowable angle range gives a value of $8\pi^2$ for $V_{\text{f}}^{\text{rot}}$.

Results

The final expressions of translation and rotation entropy loss are:

$$\begin{cases} \Delta G_{\text{trans}} = -RT \ln(C_0 V_{\text{trans}}) \\ \Delta G_{\text{rot}} = -RT \ln\left(\frac{V_{\text{rot}}}{8\pi^2}\right) \end{cases}$$

V_{trans} and V_{rot} were computed using the two approaches described above (quasiharmonic or simple maximum-minimum). Table S1 reports the range of values obtained for each system.

	ΔG_{trans}	ΔG_{rot}	$\Delta G_{\text{rot+trans}}$
Canonical	[0.29, 1.84]	[0.01, 1.69]	[0.30, 3.53]
Proximal	[-0.47, 1.58]	[3.59, 3.67]	[3.12, 5.25]
Distal	[-0.32, 0.81]	[0.36, 2.44]	[0.04, 3.25]
Distal-type docking pose	[0.04, 1.64]	[0.67, 2.29]	[0.71, 3.93]
Noncanonical	[-0.39, 1.44]	[0.29, 1.36]	[-0.10, 2.80]

Table S1: minimum and maximum estimates of the residual translation and rotation entropies at the largest simulated interpartner distance, using the models described above (kcal mol^{-1}).

These results suggest that the absolute values of the variation of the free energy between the endpoints of the PMFs presented in this work should typically be diminished by an estimated 3 kcal mol^{-1} when compared to experimental binding free energies. Although non-negligible, this will not change the stability ranking of the different complexes presented here. It should also be stressed that due to the difficulty of precisely evaluating the conformational volume spanned by a system, these figures are indicative.

Additional figures

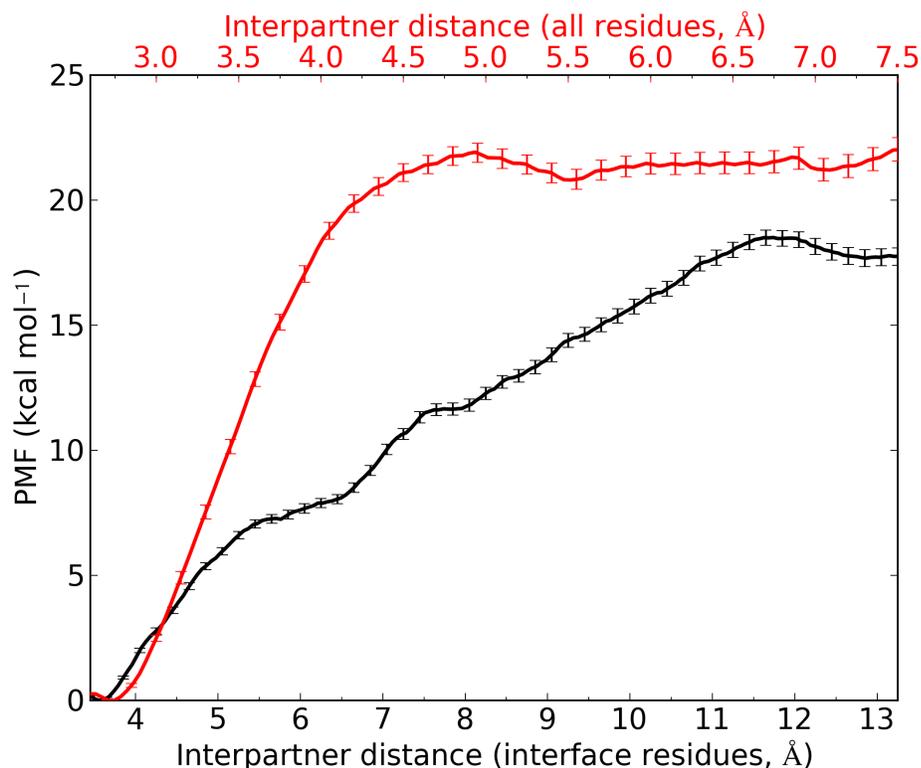


Figure S2: PMF as a function of interpartner separation for the opening of the closed Ubq₂ complex, with the minimum distance restraint acting on the heavy atoms of interface residues only (L8, I44, A46, H68, V70, L71 – black plot, lower abscissa) or on all heavy atoms (red plot, upper abscissa).

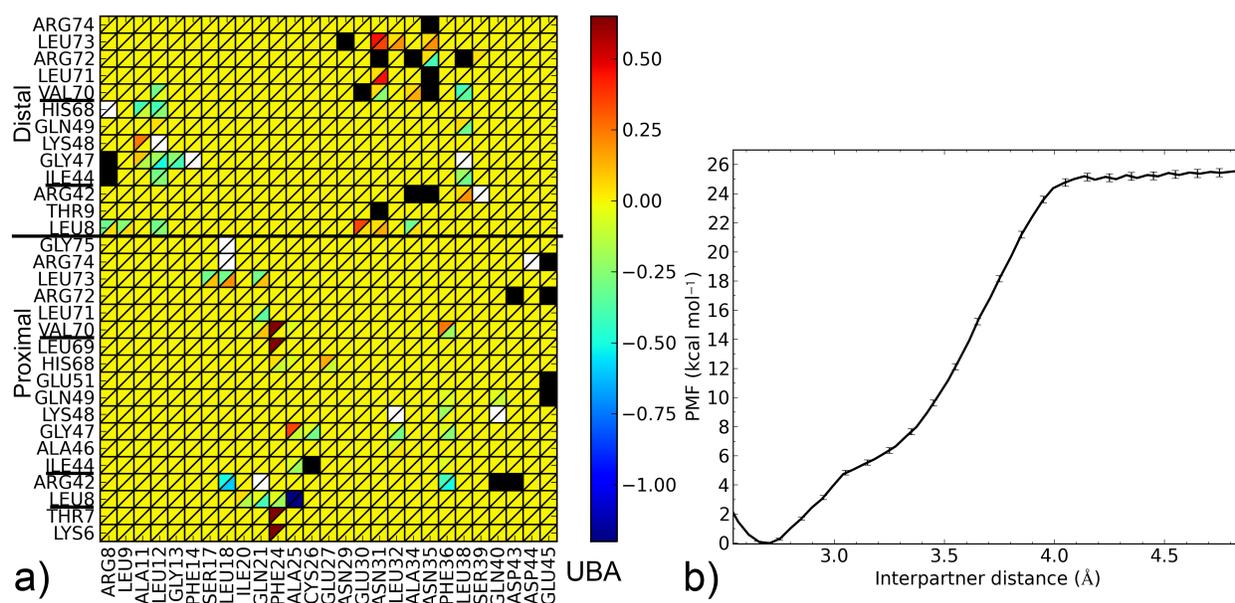


Figure S3: Dissociation of the HHR23a UBA domain from Ubq₂. a) Contact map showing the deviation of the mechanism from the dissociation pathway of UBA bound to the corresponding isolated monomers; negative (resp. positive) values correspond to contacts forming or breaking at smaller (resp. larger) distances in Ubq₂/UBA than in Ubq/UBA. b) PMF as a function of interpartner separation for Ubq₂/UBA.

Supporting files

Multi-frame PDB files containing representative conformations along the dissociation pathways for all complexes under study, available for download from the journal website.

File name	Description
canonical.pdb	Canonical Ubq/UBA complex (Ubq/EDD-UBA, PDB Id. 2QHO).
noncanonical.pdb	Noncanonical Ubq/UBA complex (Ubq/CBL-B UBA, PDB Id. 2OOB).
ubq2Proximal.pdb	UBA complexed to proximal Ubq monomer from Ubq ₂ /HHR23A-UBA complex (PDB Id. 1ZO6).
dockingProximal.pdb	Best docking pose of the monomers of Ubq/CBL-B UBA complex (PDB Id. 2OOB).
ubq2Distal.pdb	UBA complexed to distal Ubq monomer from Ubq ₂ /HHR23A-UBA complex (PDB Id. 1ZO6).
ubq2InterfaceResidues.pdb	Ubq ₂ complex (chains A and B of PDB Id. 3M3J), separation bias on interface residues only (L8, I44, A46, H68, V70, L71).
ubq2AllResidues.pdb	Ubq ₂ complex (chains A and B of PDB Id. 3M3J), separation bias on all residues.
ubq2UBA.pdb	Ubq ₂ /HHR23A-UBA complex (PDB Id. 1ZO6).

References

- 1 F. Zhu and G. Hummer, *J. Comput. Chem.*, 2012, **33**, 453.