Supplementary Information for

Competitive binding of HIF-1α and CITED2 to the TAZ1 domain of CBP from molecular simulations

Irene Ruiz-Ortiz[1] and David De Sancho[1,2]

[1]Donostia International Physics Center (DIPC), 20018 Donostia, Euskadi, Spain

[2]University of the Basque Country, Faculty of Chemistry, Paseo Manuel Lardizabal, 3, 20018 Donostia-San Sebastián, Spain

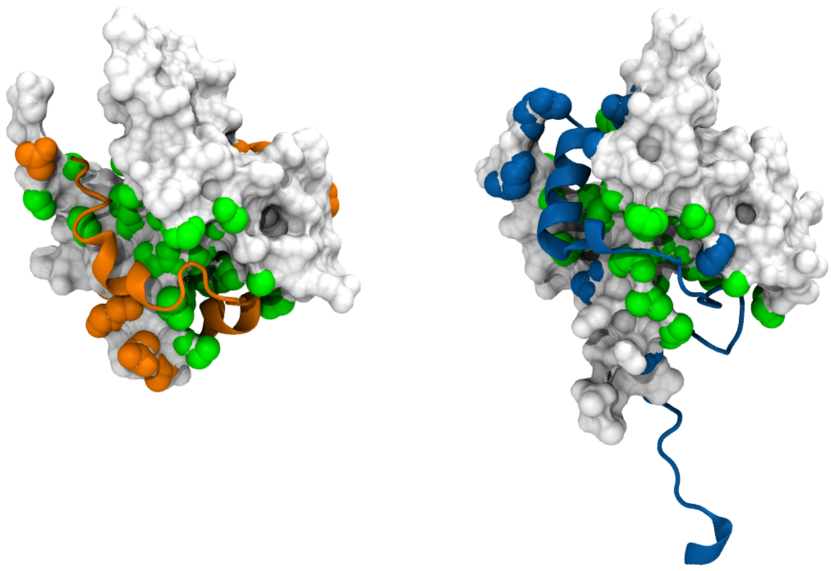

Figure S1: Representation of experimental structures of TAZ1 complexes with HIF-1α (1l8c, left) and CITED2 (1r8u, right). Surface representations in white correspond to TAZ1, and cartoon representations with secondary structure assignment for HIF-1α and CITED2 are shown in orange and blue, as in the main text. Additionally, we show as spheres, side chain atoms of TAZ1 in contact with only HIF-1α (orange), only CITED2 (blue) or both (green).

Structuring of TAZ1 upon binding to HIF-1α and CITED2

In Figure S2 we report statistics for the $RMSD$ and the fraction of intramolecular contacts Q for both the free state and the HIF-1α bound (HB) and CITED2 bound (CB) states. We find that the probability distributions for the bound states are shifted towards low $RMSD$ and high Q values, indicating that TAZ1 gets more structured upon binding. These results are consistent with the increasing rigidification of TAZ1 in the bound state derived from the values of the $RMSF$.

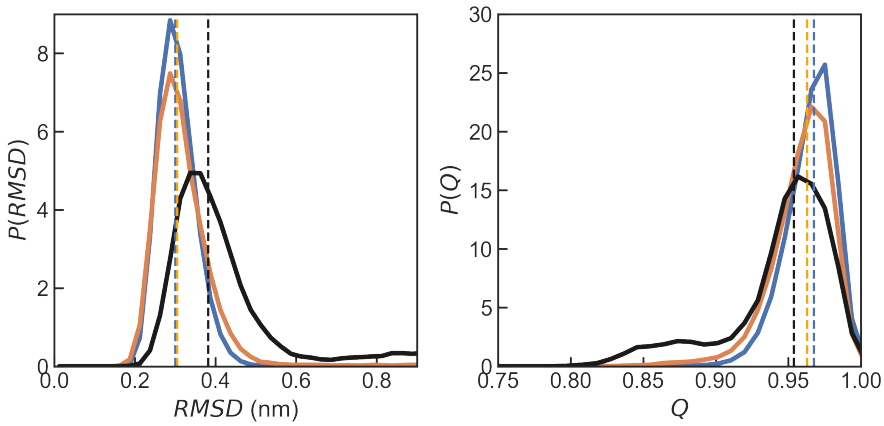

Figure S2: Histograms of the $RMSD$ from native (left) and Q (right) for conformations in the free state (black) or HIF-1α or CITED2 bound states (orange and blue, respectively). Dashed vertical lines mark the median values.

Convergence tests for replica exchange molecular simulations

Convergence of the simulations is essential in order to be able to obtain reliable estimates of equilibrium properties of the system from replica exchange MD simulations. We evaluate convergence in our datasets from cumulative averages of metrics corresponding to the binding of HIF-1α and CITED2 to TAZ1. In Figure S3 we show the d_{RMS} for both complexes at multiple temperatures in the range explored in our replica runs. Despite the difficulty in exhaustively sampling the ternary complex, we find that the average values of the d_{RMS} converge to the mean values (within error) after ~2 μs, while the total duration of our REMD runs is 5 μs at each temperature. Hence, we are confident that our simulations are well converged.

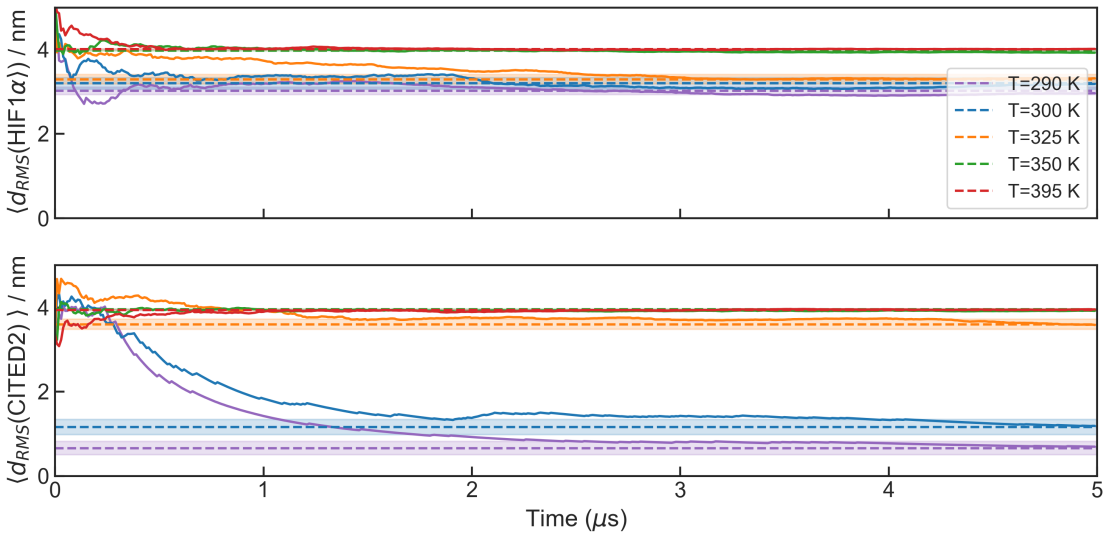

Figure S3: Cumulative averages at multiple temperatures of the d_{RMS} for intermolecular contacts of the HIF-1α-TAZ1 (top) and CITED2-TAZ1 (bottom) complexes. We show the block average with errors indicated as +/- one standard deviation as dashed lines and bands with the same colour code, respectively.

Short trajectories result in jumps between the dominant energy basins

We have run 90 additional trajectories of 100 ns to 1μs duration starting from the different dominant basins for binary complexes (CB, HB and HI, as defined in the main manuscript). In Figure S4 we show the projection of these simulation datasets on the fraction of intermolecular contacts for both HIF-1α and CITED2 for the different initial states. In all cases we find that from the initial states there are transitions into neighbouring free energy basins, resolving the connectivity HB↔HI↔HCI↔CB. We note that, out of 20 100 ns trajectories from the CB state, only one hopped to the HCI state. That is to be compared with 3 out of 30 trajectories run from HI that reached the ternary intermediate. Among the 40 trajectories started from the HB state, 20 were extended to reach the 1 μs mark. Thirteen trajectories hopped between the HB and HI states, four of which reached the ternary intermediate (one of them in the initial 100 ns), and another one (pink in the left panel of Fig. S4) reached the CB state in ~500 ns.

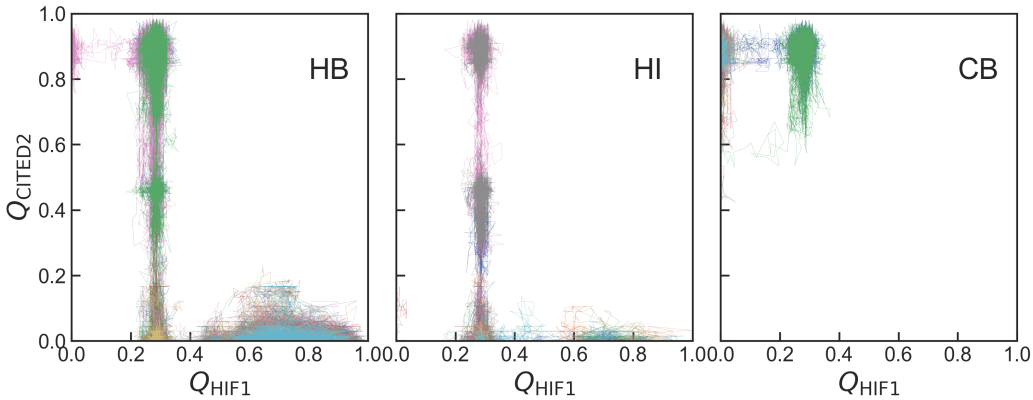

Figure S4: Projection of simulation trajectories on the fraction of intermolecular contacts for HIF-1α (Q_{HIF1}) and CITED2 (Q_{CITED2}). Labels indicate the initial state for each of the sets.