Supplemental Information. Stay Wet, Stay Stable? How Internal Water Helps Stability of Thermophilic Proteins.

Debashree Chakraborty, Antoine Taly, and Fabio Sterpone*

Laboratoire de Biochimie Théorique, IBPC, CNRS, UPR9080, Univ. Paris Diderot, Sorbonne Paris Cité, 13 rue Pierre et Marie Curie, 75005, Paris, France,

E-mail: fabio.sterpone@ibpc.fr

*To whom correspondence should be addressed

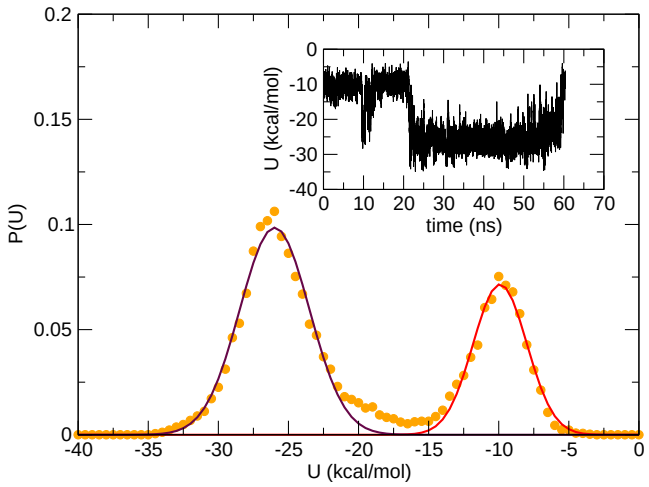

Figure 1: Probability distribution of the potential energy U of a target water molecule residing in the interior of the hyperthermophilic G domain. The two steady states visible in the time evolution of $U(t)$ (see inset) reflect the localisation of the molecule in two different sites in the internal cavity. For each one of these states we applied the Gaussian approximation in order to evaluate the excess chemical potential μ_{ex}^{pw}.

Table 1: Ensemble of long residence water molecules used for the free energy calculations. (*) In a earlier work the number of water used to estimate the internal hydration free energy was of 8 molecules for each homologue[1], each molecule was sampled in three different internal locations.

Mesophile		Thermophile	
Protein	n water	Protein	n water
1B8P	16	1BDM	22
1GCI	20	1THM	25
3H1G	7	1DZ3	9
1IOH	16	3MDS	19
3TL2	22	1A5Z	19
1P3J	18	1ZIN	18
2X8S	22	3CU9	19
1GV1	21	4CL3	26
1EFC*	15	1SKQ*	18

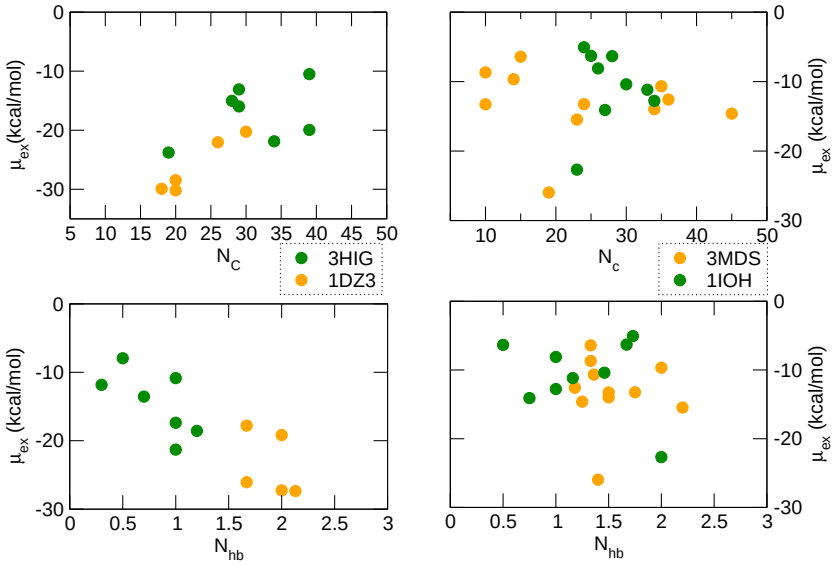

Figure 2: Local chemistry of the internal cavities sampled by long residence water. For each long residence water molecule we counted how many HBs are formed with the wall of the cavity (bottom panels) and how many hydrophobic atoms are found within a radius of 5 Å(top panels). On the left we report data for a pair showing a large difference of the per molecule excess chemical potential between the thermophilic and the mesophilic variants (1DZ3/3H1G). On the right we report data for a pair showing no difference of the per molecule excess chemical potential between the two homologues.

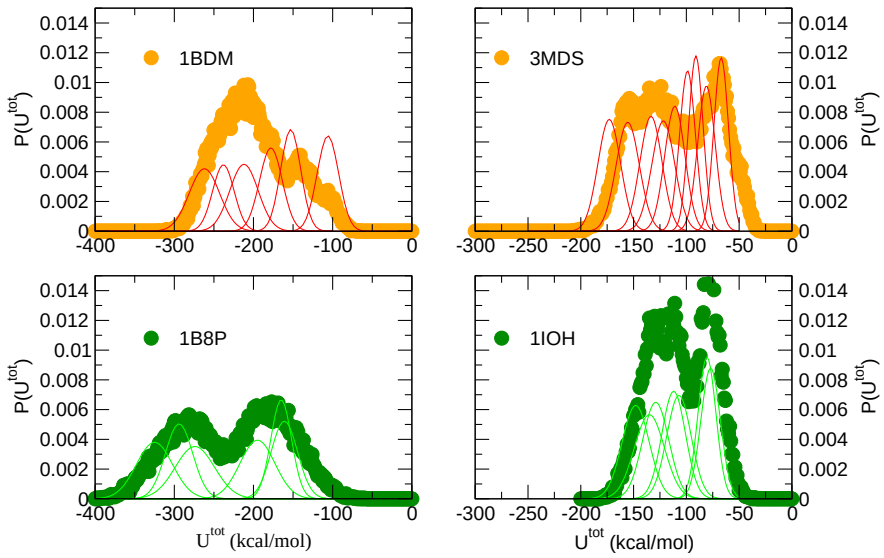

Figure 3: Probability distribution of the total potential energy U^{tot} resulting from the interaction of the internal water with the proteins. On the left we report data for the pair 1BDM/1B8P and on the right the data for the pair 3MDS/1I0H. Data for the thermophiles are plotted in the top panels (oranges curves) and data for the mesophiles are plotted in the bottom panels (green curves). In each panel we report the underlying Gaussian state that are used to sample different steady hydration states of the internal cavities, for each of this state the Gaussian approximation is used to estimate the gain of internal wetting.

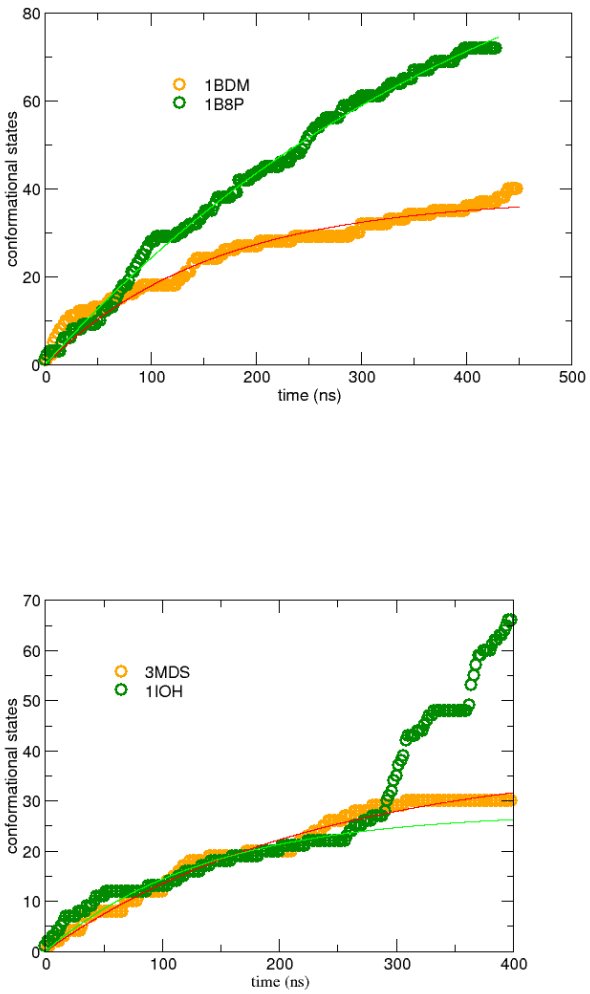

Figure 4: Number of conformational states visited by the mesophilic and thermophilic malate dehydrogenase single domains, 1B8P and 1BDM (top), and by the two manganese superoxide dismutases homologues, 3MDS and 1IOH (bottom). The conformational states are numerated on the basis of relative Cα-RMSD and using clustering leader-follower algorithm with cut-off 1.3 Å. Continuous lines refer to a simple exponential growth model used to fit the curves.

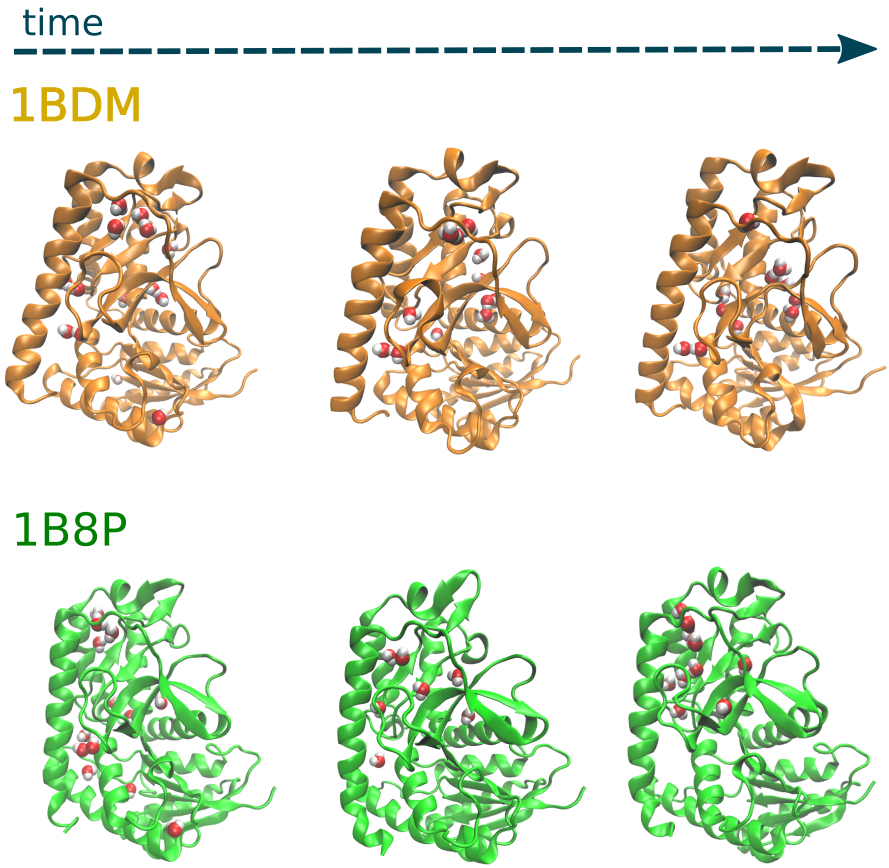

Figure 5: Molecular representation of the internal hydration of the proteins 1BDM/1B8P for selected states extracted along the simulation.

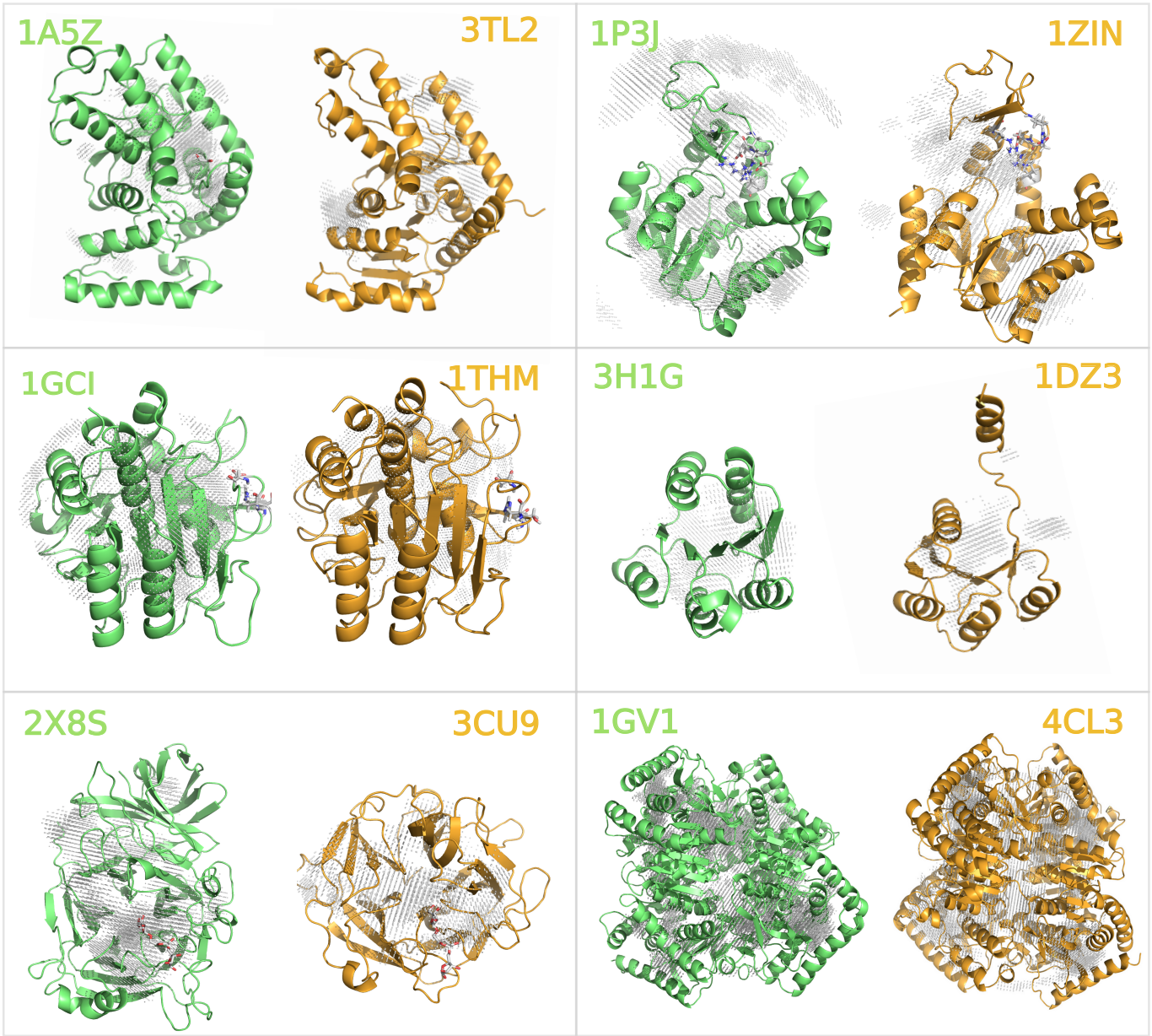

Figure 6: Density map of long residence water with respect to the X-ray protein structures.

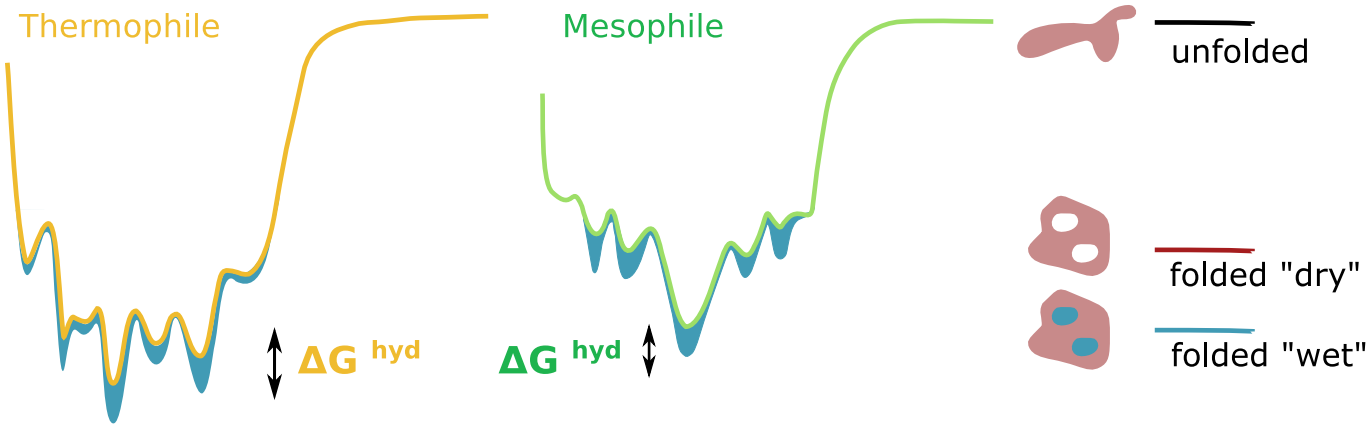

Figure 7: Schematic representation of the contribution of internal hydration to protein stability as computed in our framework.

References

(1) Rahaman, O.; Kalimeri, M.; Melchionna, S.; Hénin, J.; Sterpone, F. *J. Phys. Chem. B* **2015**, *119*, 8939–8949.