

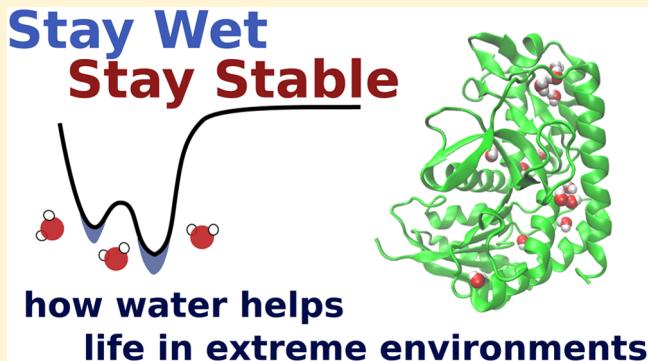
Stay Wet, Stay Stable? How Internal Water Helps the Stability of Thermophilic Proteins

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

S Supporting Information

ABSTRACT: We present a systematic computational investigation of the internal hydration of a set of homologous proteins of different stability content and molecular complexities. The goal of the study is to verify whether structural water can be part of the molecular mechanisms ensuring enhanced stability in thermophilic enzymes. Our free-energy calculations show that internal hydration in the thermophilic variants is generally more favorable, and that the cumulated effect of wetting multiple sites results in a meaningful contribution to stability. Moreover, thanks to a more effective capability to retain internal water, some thermophilic proteins benefit by a systematic gain from internal wetting up to their optimal working temperature. Our work supports the idea that internal wetting can be viewed as an alternative molecular variable to be tuned for increasing protein stability.



INTRODUCTION

Life is found at both severe cold and hot climates. Extremophilic organisms equipped with suitable molecular machinery challenge these extreme conditions. For instance, proteins from thermophilic bacteria or archaea are stable and functional at very high temperatures,¹ in some cases up to the boiling point of water (100 °C).

Thermophilic enzymes are considered a natural template for understanding the elementary molecular factors concurring with the stability and the stability-and-function trade-off of proteins.^{1–3} Structural analysis and biochemical engineering have successfully singled out several mechanisms of thermal stabilization. These include the location of strategic ion pairs,^{4–6} the optimal distribution of charged amino acids at the surface,^{7,8} and optimal local packing.⁹ However, it is widely accepted that different molecular mechanisms can be combined, resulting in a variety of thermodynamic routes to thermal adaptation.^{10–12}

The design of enzymes resistant to high temperatures or other harsh conditions, e.g., pH and denaturants, is appealing in biotechnology, chemical processing, and other technologies that aim to exploit the power of enzymes.^{13–16} The possibilities for enlarging the molecular variables so as to tune them to achieve enhanced stability of the enzymes is therefore of practical interest.

Inspired by seminal work on globular proteins that attempted to quantify the contribution of cavity hydration on protein stability,^{17–19} we started examining whether internal wetting is a key stabilizing factor of thermophilic proteins when compared to results from their mesophilic variants.²⁰ The goal is to reveal an alternative strategy in protein engineering. In fact, several

investigation have suggested that specific solvation could contribute to the enhanced stability of the thermophilic enzymes.^{21–27} Moreover, it is important to recall that thermophilic proteins generally sustain high pressure too.¹ The presence of strategic internal water helps in explaining this double resistance as the effect of optimized internal packing²⁸ and extended molecular hydrogen-bond patterns, which are less sensitive to temperature and pressure.^{22,29}

Experimentally, by applying single-point mutation (which modifies the local environment of internal cavities too), it was found that buried water acts as a stabilizing agent for several proteins such as BPTI,¹⁹ subtilisin,¹⁸ lysozyme,³⁰ and lipases.²² However, for some other cases, the opposite is true (e.g., iso-1-cytochrome,³¹ protease,³² and lysozyme T4 mutants).¹⁷ Computational studies^{33–39} have provided complementary atomistic details on water penetration and correlated the internal wetting to the chemical nature of protein pockets,^{36–38,40} protein stability,^{36,41} and their function.^{42–51}

How the internal wetting acts differently on the stability and function of mesophilic and thermophilic homologues, however, has never been addressed in detail. In a seminal work, Yin, Hummer, and Rasaiah²⁶ investigated the hydration of the internal cavities of tetrabrachion protein from hyperthermophilic *Staphylothermus marinus*, whose optimal growth temperature is ~365 K. They pointed out that at high temperatures, dewetting of the internal cavities leads to denaturation. Recently, some of

Received: June 17, 2015

Revised: September 3, 2015

Table 1. Pairs of Mesophilic and Thermophilic Homologues Studied in This Work^a

PDB code	protein	organism	T_{opt} (°C)	N_{aa}	N_w	simulation length (ns)
1EFC	G-domain EF-Tu	<i>Escherichia coli</i>	37	196	7440	600
1SKQ (A)	G-domain EF- α	<i>Sulfulobos solfataricus</i>	75	226	10673	600
1B8P (A)	malate dehydrogenase	<i>Aquaspirillum articum</i>	4–10	327	9944	430
1BDM (A)	malate dehydrogenase	<i>Thermus thermophilus</i>	65	317	8837	450
1GCI	subtilisin	<i>Bacillus lentus</i>	10–35	269	7324	410
1THM	thermitase (subtilisin-like)	<i>Thermoactinomyces vulgaris</i>	55	279	7142	428
3H1G	chemotaxis (CheY-like)	<i>Helicobacter pylori</i>	30–37	123	4720	450
1DZ3	SpoOA (CheY-like)	<i>Geobacillus stearothermophilus</i>	55–65	123	6727	400
1I0H	manganese superoxide dismutase	<i>E. coli</i>	37–49	205	7226	430
3MDS	manganese superoxide dismutase	<i>T. thermophilus</i>	65	203	6680	420
3TL2 (A)	malate dehydrogenase	<i>Bacillus anthracis</i>	37	315	10014	400
1ASZ (A)	L-lactate dehydrogenase	<i>Thermotoga maritima</i>	80	311	11609	400
1P3J	adenylate kinase	<i>Bacillus subtilis</i>	25–35	212	6395	400
1ZIN	adenylate kinase	<i>G. stearothermophilus</i>	55–65	217	8461	400
2X8S	L-arabinanase	<i>B. subtilis</i>	28–30	470	11457	210
3CU9	L-arabinanase	<i>G. stearothermophilus</i>	55–65	314	9739	200
1GV1	malate dehydrogenase	<i>Chlorobium vibrioforme</i>	15–45	1240	35485	600
4CL3	malate dehydrogenase	<i>Chloroflexus auranticus</i>	>50	1236	35422	600

^aThe optimal temperature of the host organism is reported in column 4. The number of amino acids and of water molecules solvating the protein in the simulation box and the simulation length are reported in the last three columns. For the proteins 1B8P/1BDM and 3TL2/1ASZ, the simulated monomer corresponds to the chain A in the PDB.

our group members performed a comparative study between mesophilic and hyperthermophilic G domains.^{20,23,52,53} The internal hydration was found to contribute to the stability gap between the two homologues. A gain of about 1.3–2.5 kcal/mol was estimated in favor of the hyperthermophilic domain, which corresponds to the shift of the melting temperature of about 6 K, almost $\frac{1}{6}$ of the experimental shift between the two proteins.²⁰ It was also clearly shown that internal hydration correlates to the flexibility and rigidity of the protein matrix.^{20,53} This finding potentially opens a new perspective with which to elicit the relationship among protein mechanical and thermal stabilities, and hydration.³

In this work, we extend our investigation and present the results of a systematic comparative study on a set of homologues of different stability contents. We used molecular dynamic (MD) simulations stretching to hundreds of nanoseconds to explore the exchange dynamics of water penetrating inside the protein structures. On the basis of this analysis and using a convenient computational approach, we estimated the hydration free energy for buried water. Our results show that thermophilic proteins can actually benefit from internal hydration to ensure their stability in a broad range of temperatures. This is caused by a more favorable interaction of water with the internal cavities and by a stronger capability to retain internal wetting in a broader window of temperatures.

METHOD

Systems and MD Simulations. We have investigated eight pairs of homologous mesophilic and thermophilic proteins belonging to different families. As a starting point, we have considered the list of proteins analyzed in a previous work by Pechkova et al.,⁵⁴ in which the authors looked for a direct correlation between thermophilicity and water content in X-ray structures. We have extended the investigation by searching in the Protein Data Bank for homologues of high structural similarity and focusing on the content of structural water. Our analysis was not conclusive, and the overall results appear to be

extremely dependent on the different resolutions at which the structures were resolved. To have a better understanding without limiting ourselves to crystallographic data, we have selected a pool of seven pairs for our atomistic simulations, and the proteins in the selected pairs have comparable X-ray resolutions. For computational reasons, the selected monomeric proteins have relatively small sizes (200 to 400 amino acids). To these seven pairs, we added another one already investigated by some of us in a different context, a pair of tetrameric malate dehydrogenases.⁵⁵ A second extra pair was used for the sake of comparison of the free-energy calculations (see below in the text); this is a pair of G domains from elongation-factor proteins.²⁰ The molecular details of these systems are summarized in Table 1, and a molecular view of the structural superimposition of the eighth pairs studied specifically here is given in Figure 1. For the G domains we refer to the previous work.²⁰ The size of the systems ranges from 123 up to 1240 amino acids, and each protein fold contains both α helices and β structures. A set of three pairs belong to the malate and lactate dehydrogenase family. Of these three, for two we considered only one monomer of the multimeric protein, while for the other pair (1GV1/4CL3), all of the tetrameric protein is considered. This selection helps to filter the effect of interdomain interface on water confinement.

For each protein, we performed MD simulations in aqueous solution and at ambient conditions. The proteins were modeled by the charmm22 force field combined with the charmm-TIP3P model for water.⁵⁶ The proteins were placed in a simulation box, ensuring a good solvation conditions (see Table 1). Crystallographic water was maintained during the system setup. The simulations were carried out using the NAMD package.⁵⁷ After an equilibration phase in the NPT ensemble, the trajectories were integrated with a time step of 2 fs in combination with the Langevin thermostat to keep the temperature constant (characteristic time of $\tau_T = 0.2 \text{ ps}^{-1}$) and sampling the canonical ensemble (NVT). Electrostatics was handled by the PME scheme with a resolution of 1 Å for the reciprocal part. The short-range part of electrostatic interactions and vdW interactions were truncated at 10 Å. The configurations along the trajectories were

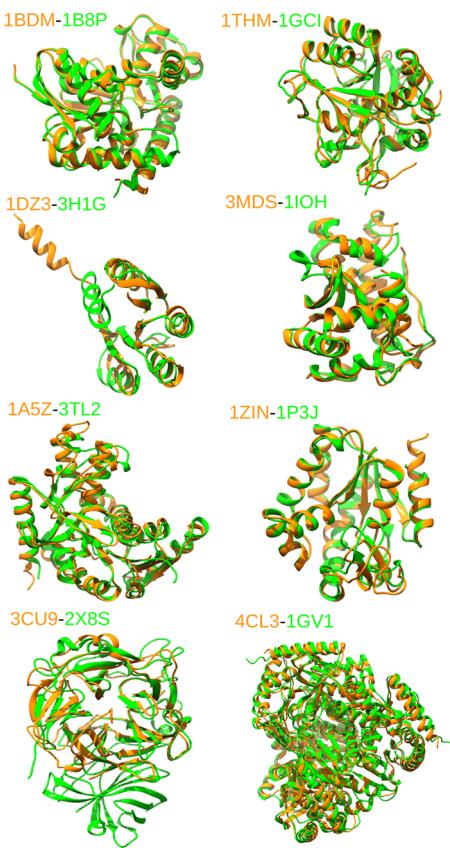


Figure 1. Molecular representation of the eight homologous pairs studied in this work. The mesophilic variant is colored in green, and the thermophilic variant is colored in orange.

stored with a frequency of 2 ps. The simulations were extended for about 400 or 600 ns. For selected pairs, complementary simulations of various lengths (from 60 to 100 ns) were performed at high temperatures, namely 320, 340, and 360 K.

Internal Water. To estimate the free-energy contribution of internal hydration to protein stability, it is necessary, as first step, to individuate the molecules hydrating the internal sites of the protein. For this purpose, we used a kinetic criterion and selected the water molecules that reside continuously in the protein hydration shell for more than 4.5 ns. In fact, it is well-established that water molecules buried in protein internal cavities or pockets exchange with the external solution with a characteristic time exceeding the nanosecond.^{35,58–60} For each of the individuated water molecules, we estimated the hydration free energy that represents the work for transferring the molecule from the external bulk solution to the protein internal site. The framework used for these calculations is discussed below. The hydration shell is defined by the water molecules that are within the spherical cut-off ($r_c = 4.5 \text{ \AA}$) from the heavy atoms of the protein. The continuous location of water in the hydration shell was monitored by the survival probability function, $N_w(t)$, as described in refs 20 and 35. The analysis of internal hydration was restricted to a representative stretch of the trajectory (~100 or 200 ns, depending on the system and the exchange dynamics). In the specific case of the tetrameric malate dehydrogenases (1GV1/4CL3), because of the size of the system and the presence of an internal water pool at the interface with the four domains, the kinetic cutoff was made more strict (20 ns) to select the buried water molecules and filter out the molecules in the internal pool. From the ensemble of long-residence water

molecules, a smallest set constituted of the more stable molecules located in the interior of the proteins was used for the free-energy calculations. The number of internal water molecules used for calculating hydration free energies is reported in Table S1 for all pairs but one, and ~20 molecules per protein were used as a sample.

Free-Energy Calculations. Hydration free energy can be rigorously computed by combining the free-energy perturbation (FEP) method and a localization constraint.^{34,36,61} In short, a non-interacting water molecule is localized in the protein reference site, where it is slowly energetically coupled to the protein environment. The same procedure (without localization) is performed for a molecule in the pure water solution. The free-energy differences between the two processes corrected by the localization contribution represents the free-energy gain or penalty to the hydration of the protein internal site. For a complete description, we refer to the work of Rick and co-workers.³⁶ However, this approach is computationally expensive, and when the goal is a systematic comparison between different systems, a more efficient method, although approximated, is preferred. Several strategies have been recently reported in the literature.^{62,63}

In this work, we rely on the so-called Gaussian approximation⁶⁴ that, in combination with the particle-insertion method, provides a reasonable compromise for estimating hydration free energy. In the past, this approach was successfully tested for different molecular environments, such as proteins and micelles.^{65–67}

The excess chemical potential μ_{ex} associated with the hydration of a protein internal site can be written as the sum of two contributions

$$\mu_{\text{ex}} = \mu_{\text{ex}}^0 + \mu_{\text{ex}}^{\text{pw}} \quad (1)$$

The first term on the right side represents the work to create an empty volume in the disordered protein medium capable to host a water molecule, and the second term measures the free-energy contribution due to the specific interactions of the molecule with the protein, i.e., electrostatic and vdW interactions. This term is formally given by the ensemble average $\frac{1}{\beta} \ln \langle e^{\beta \Delta U} \rangle_1$, where the contribution $\Delta U = U_1 - U_0$ is the difference between the potential energy calculated for a given configuration, assuming a fully interacting particle (final state, U_1) and a noninteracting particle (initial state, U_0). Because the initial state of the hydration process is represented by the noninteracting water molecule located in the protein site, the term is $U_0 = 0$. The ensemble average $\langle \dots \rangle_1$ is evaluated in the final state representing the interacting water molecule inside the protein.

The numerical estimate of the term $\langle e^{\beta \Delta U} \rangle_1$ is difficult to converge in a standard simulation because of the oscillatory nature of the exponential function, and an approximation is often used. In this work, we apply the so-called Gaussian approximation that reduces the expression to

$$\mu_{\text{ex}}^{\text{pw}} = \frac{1}{\beta} \ln \langle e^{\beta \Delta U} \rangle_1 \approx \langle \Delta U \rangle_1 + \frac{1}{2} \beta \sigma_1^2 \quad (2)$$

where the average of ΔU and its fluctuations $\sigma^2 = \langle (\Delta U - \langle \Delta U \rangle)^2 \rangle$ are extracted from the simulations in state 1 (internal site hydrated). Similarly, the excess chemical potential of a water molecule in bulk can be computed. For the TIP3P model at ambient conditions, it is estimated to be ~9 kcal/mol.^{66,67} The advantage of the Gaussian approximation is that by using a single long simulation where the hydration of internal

sites is sampled at once, it is possible to extract the hydration free energy of the sites of interest.

The cavity term μ_{ex}^0 can be calculated according to the Widom's particle-insertion theory.⁶⁸ In previous work, Garcia and co-worker showed that this term is approximately the same when estimated for an internal protein site or for the bulk water, $\approx 4 \text{ kcal/mol}$,^{65,66} and differences are in the order of 0.2 kcal/mol. Therefore, this contribution cancels out when the total free-energy difference, $\Delta\mu_{\text{ex}} = \mu_{\text{ex}}^{\text{pw}} - \mu_{\text{ex}}^{\text{bulk}}$, is calculated.

After a preliminary check, we have verified that generally, the dominant contribution to the hydration free energy stems from the direct interaction of the tagged internal water with the protein matrix, so for our systematic calculations, we have just considered this contribution to the interaction potential energy entering in the estimate of $\mu_{\text{ex}}^{\text{pw}}$.

RESULTS AND DISCUSSION

Free Energy of Internal Hydration. The quality of the Gaussian approximation is first verified. In a previous study,²⁰ we have constructed the distribution of the per-molecule hydration free energies ($\Delta\mu_{\text{ex}}$) at ambient temperature for water buried in the internal cavities of two homologous G domains by applying rigorous FEP calculations. We found that although the two distributions look very similar, the internal hydration of the thermophilic proteins is, however, slightly more favorable than that of the mesophilic variant. For the sake of comparison, they are reported in the top panel of Figure 2. We have constructed

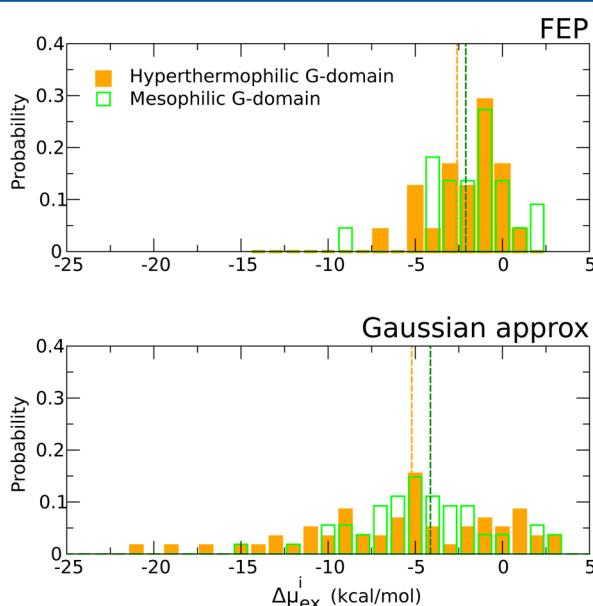


Figure 2. Probability distribution of the free energy to transfer a single water molecule from the external solution to the interior of the G-domain proteins ($\Delta\mu_{\text{ex}}$). The top panel refers to FEP calculations (see ref 20). The bottom panel refers to calculations performed applying the Gaussian approximation. The values obtained for the mesophilic protein are reported in green, and those for the thermophilic variant are reported in orange.

the equivalent probability distribution by applying the Gaussian approximation. Because of the less-demanding computational cost, the calculations were performed on a larger set of internal water than in the case of FEP calculations. The final result is reported in the bottom panel of Figure 2. As observed using the FEP calculations, the distributions for the mesophilic and

thermophilic species are very similar, and for the majority of the considered water molecules, the location in the protein internal sites is energetically favorable. When the results are compared to the FEP distributions, we notice that when applying the Gaussian approximation, the values of the hydration free energy have a broader spread, ranging from +2 to -20 kcal/mol . As an effect of this broadening, the average hydration free energy for the thermophilic homologue results are about 3 kcal/mol more favorable than that of the mesophilic one. All of the estimates of $\mu_{\text{ex}}^{\text{pw}}$ come with an error, estimated by block analysis, of about 0.2 kcal/mol each.

The broader shape of $P(\mu_{\text{ex}})$ is caused by two concomitant effects. First the Gaussian approximation cannot always properly account for the tails in the distribution $P(\Delta U)$ that extend at less-favorable gap values of ΔU ; second, the ensemble of water used for reconstructing the free energy is larger (see the last line in Table S1), and for the same water, the locations in several close sites are naturally included in the calculations. Note that, as reported in Figure S1, when separate states exist, the $P(\Delta U)$ value deviates from the unimodal Gaussian-like shape, and thus, we separated the states and carried out independent calculations.

Next, we turn our attention to the set of protein pairs investigated for the first time in this work. Because we are interested in comparing the two homologues, we focus only on the term $\mu_{\text{ex}}^{\text{pw}}$; in fact, the bulk term ($\mu_{\text{ex}}^{\text{bulk}}$) to be subtracted to obtain the hydration free energy is the same for the two species. In Figure 3, we report the distribution of the excess chemical potential ($\mu_{\text{ex}}^{\text{pw}} = \langle \Delta U \rangle + \frac{1}{2}\beta\sigma^2$) calculated for internal water for all of the protein pairs. In each panel, we compare the mesophilic (green) and the thermophilic (orange) variants. For the majority of the pairs, the distribution of $\mu_{\text{ex}}^{\text{pw}}$ extracted from the simulations of the mesophilic and of the thermophilic species are very similar and overlap. As a consequence, the per-molecule free-energy gain due to internal hydration is comparable between the two homologues; $\Delta\mu_{\text{ex}}^{\text{pw}} = \langle \mu_{\text{ex}}^{\text{pw}} \rangle^T - \langle \mu_{\text{ex}}^{\text{pw}} \rangle^M$ is -2 to 0 kcal/mol . However, in a few cases, the distributions of the thermophilic species is shifted toward more negative values and associated with a substantial favorable hydration free energy, such that $\Delta\mu_{\text{ex}}^{\text{pw}}$ equals -4 to -10 kcal/mol (see the pairs 1DZ3/3HIG, 1A5Z/3TL2, and 4CL3/1GV1). This gain is associated with a specific hydrophilic characters of the internal cavities of the thermophilic proteins; for example, for the pair 1DZ3/3H1G, the buried water molecules are systematically linked by a higher number of hydrogen bonds, and the cavities are depleted of hydrophobic contacts (see Figure S2). The 1DZ3/3HIG pairs is a specific case; these are the smallest proteins in our data set, the size being of 123 amino acids, and on average, only one single long-residence water molecule is buried in the protein matrix.

A separation of the enthalpic versus entropic contributions to $\mu_{\text{ex}}^{\text{pw}}$ is provided in the 2D plots presented in Figure 4. For each pair, the scattered distribution of the values $\langle U \rangle$ and $\frac{\sigma^2}{2k_b T}$ is plotted. We first note that the larger hydration free energy computed for the thermophilic proteins 1A5Z, 1DZ3, and 4CL3 with respect to their mesophilic counterparts is granted by more negative potential energy terms, $\langle U \rangle$, that quantifies a favorable environment for water hydrogen-bond connectivity in the protein interior. A second important result concerns the distribution of the entropic term $\frac{\sigma^2}{2k_b T}$ for the pair of tetrameric malate dehydrogenases, 1GV1/4CL3. The fluctuations of the interaction energy are systematically larger in the mesophilic

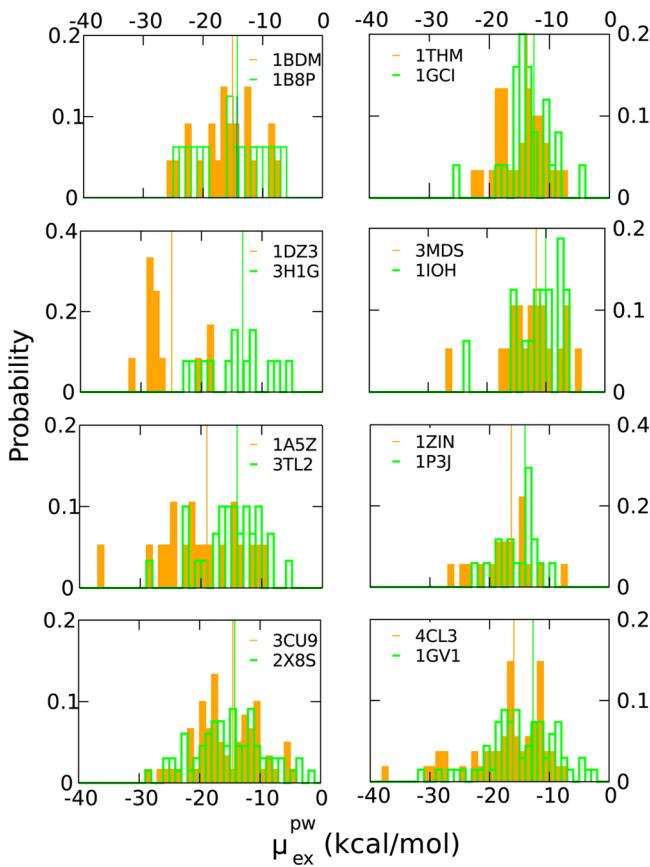


Figure 3. Probability distribution of the excess chemical potential due to the interactions between the water molecule and the protein as calculated via the Gaussian approximation. Each panel refers to a different pair of thermophilic and mesophilic homologues. Data for the mesophiles are in green, and data for the thermophiles are in orange. The average value $\langle \mu_{\text{ex}}^{\text{pw}} \rangle$ are indicated by the vertical lines. All averages are lower than the bulk estimate for $\mu_{\text{ex}}^{\text{bulk}} = -9$ kcal/mol, reporting an average favorable internal hydration.

malate than in the thermophilic one, $\frac{\sigma^2}{2k_b T}$, and takes values larger than 4 kcal/mol. This correlates to what was observed in a recent computational investigation, wherein the mesophilic tetrameric malate dehydrogenase (1GV1) was probed to be more flexible than the thermophilic species (4CL3) by monitoring several metrics describing protein conformation changes.⁵⁵ In fact, the conformational flexibility ensures larger changes in the confinement acting on internal water as visible in larger σ^2 values. As noted for conformational flexibility, the shift in water confinement between the two tetrameric homologues is an effect of interdomain interactions because for the single domains of the lactate and malate proteins of the pairs 1BDM/1B8P and 3TL2/1ASZ, no differences pop up. When focusing on the ensemble of systems all considered, we recover a neat trend: water–protein interactions are more favorable on average for thermophilic species than mesophilic ones ($\bar{U} = -16.4$ and -13.1 kcal/mol, respectively), while the entropic penalty is comparable (for the thermophilic proteins, $\frac{\sigma^2}{2k_b T} = 2.5$ kcal/mol, and for the mesophilic ones, $\frac{\sigma^2}{2k_b T} = 2.8$ kcal/mol). The values extracted from our simulations are comparable to what was computed with a different approach for internal water in proteins.^{36,63,65}

Stability Gain from Hydration. The per-molecule excess chemical potential $\mu_{\text{ex}}^{\text{pw}}$ measures how favorable a local internal protein site is for water. However, the protein internal cavities can be solvated by multiple molecules at once. This number is highly fluctuating and depends on conformational changes of the protein.^{20,35,60} A practical strategy for estimating this number is to count how many of the long-residence water molecules are located simultaneously in the interior of the protein. The average number obtained for the set of proteins is reported in Table 2.

For the G domain homologues discussed in the previous section,²⁰ we have attempted to estimate the total contribution stemming from these molecules to the overall stability of a protein. Mainly, we were interested in evaluating the energetic gain for a protein in solution from having the internal cavities wet instead of dry (see the scheme in Figure S7). On the basis of the hydration free-energy differences calculated for individual internal site with respect to bulk, the total contribution was derived by imposing the total occupation of the internal sites equal to the number of internal waters estimated by considering the kinetic cutoff on the exchange dynamics. Using a tight kinetic threshold ($\tau_c = 15$ ns), the cavities were found to host $\langle n_w \rangle = 4.7$ molecules in both homologues, and the total hydration free energy with respect to bulk was $\Delta\mu_{\text{ex}}^{\text{tot,pw}} = -15.7$ and -17.0 kcal/mol, respectively, for the mesophilic and the hyperthermophilic proteins. Therefore, a marginal contribution to the stability gap between the two proteins was associated with internal hydration, $\Delta\Delta\mu_{\text{ex}}^{\text{tot,pw}} = 1.3$ kcal/mol in favor of the hyperthermophilic one, and this was quantifiable in a shift of the melting temperature of about 6 K. For a less-strict condition on localization kinetics ($\tau_c = 4.5$ ns as used in this work), the gap in favor of the hyperthermophilic domain was larger, $\Delta\Delta\mu_{\text{ex}}^{\text{tot,pw}} = 2.5$ kcal/mol.

To address the issue for the new systems, we have selected two representative pairs, 1BDM/1B8P and 3MDS/1IOH. For these pairs, the per-particle excess chemical potentials calculated for the mesophilic and thermophilic variants are very close, and their folds are very similar structurally. Thus, we try to understand whether the collective internal hydration could give a different contribution to the stability of the highly similar protein folds. Here, by exploiting the Gaussian approximation, we follow a different and somehow simpler strategy. First, the interactions with the protein of the internal water molecules selected by the kinetic criterion are computed along the trajectory for the stretch of time that the water is buried in the protein. At time t , this interaction energy corresponds to $U^{\text{tot}} = \sum_{i=1}^{nw} U_i(t) \delta_i(t)$, where δ_i is a function taking value of 1 if the i molecule is inside the protein and of zero otherwise, and nw is the number of water molecules used for the free-energy calculations. The fluctuations of U^{tot} are controlled by the breathing mode of the proteins and the correlated escape and penetration of the long residence water. The probability distributions of U^{tot} extracted from the trajectories are reported in Figure 5. We appreciated, as already observed for the G domains, that the internal hydration is associated with multiple states; in fact, all of the distributions show a bimodal shape well fitted by a bi-Gaussian curve.

In a previous study, we have highlighted²⁰ the correlation among the distribution of internal hydration states and the number of conformational states visited by the proteins and directly measured by order parameters that quantify the fluctuations of the protein matrix only, i.e., RMSD, fraction of native contacts, and fraction of torsional angles. Preliminary results from cluster analysis⁵³ on the pairs 1BDM/1B8P and 3MDS/1IOH, confirm the correlation (see Figure S4). For instance, the mesophilic 1B8P is more flexible and visits a larger

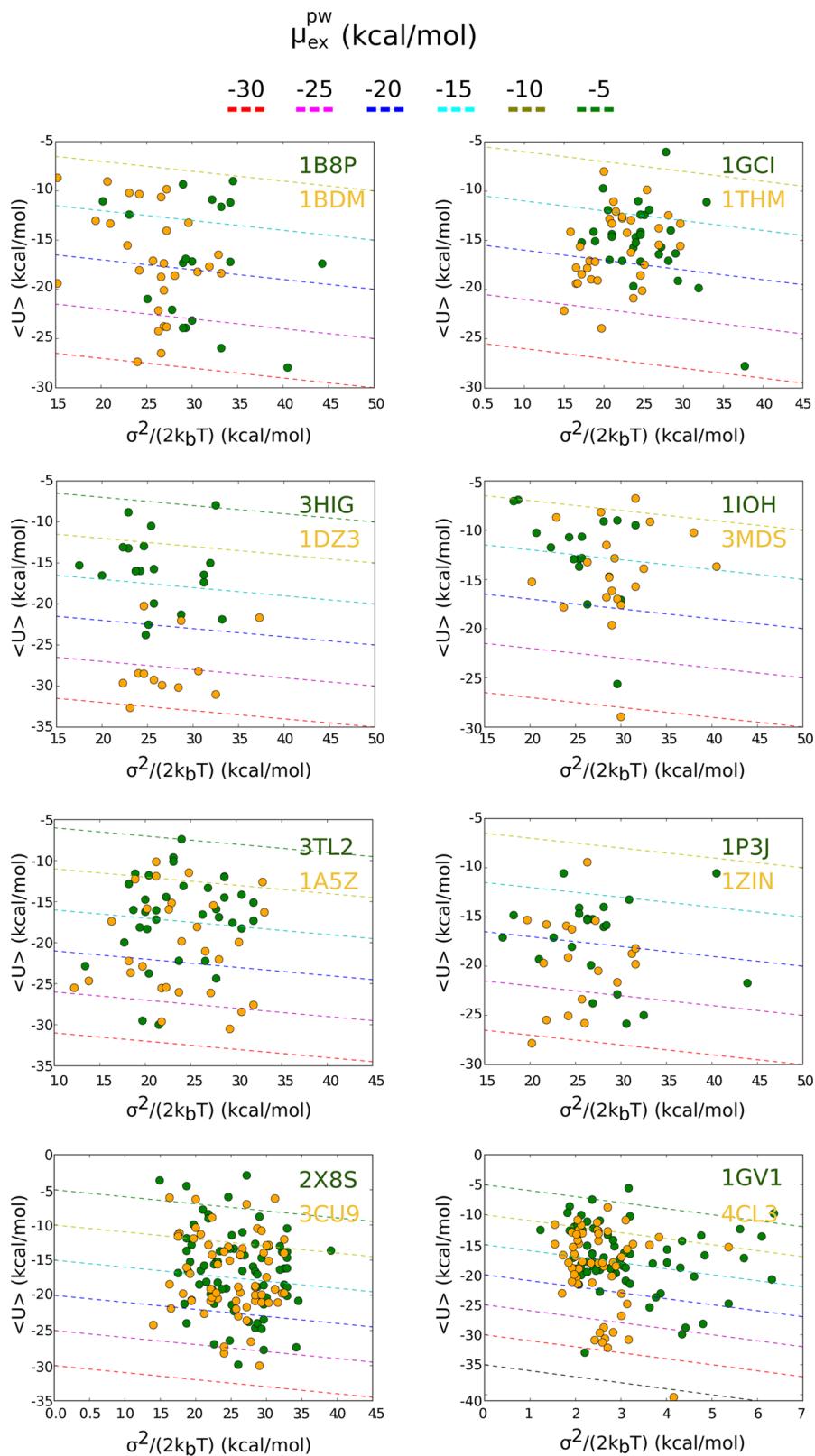


Figure 4. Plotted 2D representations of the enthalpic and entropic contributions to μ_{ex}^{pw} for each pair of thermophilic (orange) and mesophilic (green) species. Dashed lines represent iso- (free-energy) contours.

number of states than the thermophilic 1BDM; this flexibility is reflected in a broader distribution of internal hydration states (see Figure 5). For the pair 3MDS/1IOH, along the stretch of trajectory used for the study of internal hydration, the two

proteins visit a very similar number of states, and this agrees with a similar distribution of internal hydration states. A molecular representation of selected hydration states for the pair 1BDM/1B8P is given in Figure S5.

Table 2. Average Per-Molecule Excess Chemical Potential ($\mu_{\text{ex}}^{\text{pw}}$) and Hydration Free Energy^{a,b}

mesophile					thermophile				
protein	$\mu_{\text{ex}}^{\text{pw}}$ (kcal/mol)	$\Delta\mu_{\text{ex}}$ (kcal/mol)	$\langle n_w \rangle$	$\langle \Delta n_w \rangle$	protein	$\mu_{\text{ex}}^{\text{pw}}$ (kcal/mol)	$\Delta\mu_{\text{ex}}$ (kcal/mol)	$\langle n_w \rangle$	$\langle \Delta n_w \rangle$
1B8P	-14.3 (0.1)	-5.3 (0.2)	15	8	1BDM	-15.0 (0.1)	-6.0 (0.2)	16	10
1GCI	-12.6 (0.1)	-3.6 (0.2)	16	8	1THM	-13.7 (0.1)	-4.7 (0.2)	16	6
3H1G	-13.2 (0.1)	-4.2 (0.2)	1	0	1DZ3	-24.9 (0.1)	-15.9 (0.2)	1	0
1IOH	-10.0 (0.1)	-1.0 (0.2)	13	10	3MDS	-11.7 (0.1)	-2.7 (0.2)	13	8
3TL2	-14.0 (0.1)	-5.0 (0.2)	9	5	1A5Z	-19.0 (0.1)	-10.0 (0.2)	7	3
1P3J	-14.1 (0.1)	-5.1 (0.2)	5	4	1ZIN	-16.4 (0.1)	-7.4 (0.2)	5	3
2X8S	-14.2 (0.1)	-5.2 (0.2)	49	24	3CU9	-14.5 (0.1)	-5.5 (0.2)	25	13
1GV1	-12.7 (0.1)	-3.7 (0.2)	130	65	4CL3	-15.94 (0.1)	-3.7 (0.2)	135	67

^aHydration free energy: $\Delta\mu_{\text{ex}} = \mu_{\text{ex}}^{\text{pw}} - \mu^{\text{bulk}}$; $\mu^{\text{bulk}} = -9.0(0.1)$ kcal/mol. ^bThe number in parentheses indicates statistical error estimated from the propagation of standard deviations. The number of long-residence water molecules residing simultaneously in the interior of the proteins at $T = 300$ K are indicated by $\langle n_w \rangle$, and the drop of internal hydration due to temperature increase is referred as $\Delta n_w = \langle n_w \rangle^{300} - \langle n_w \rangle^{360}$.

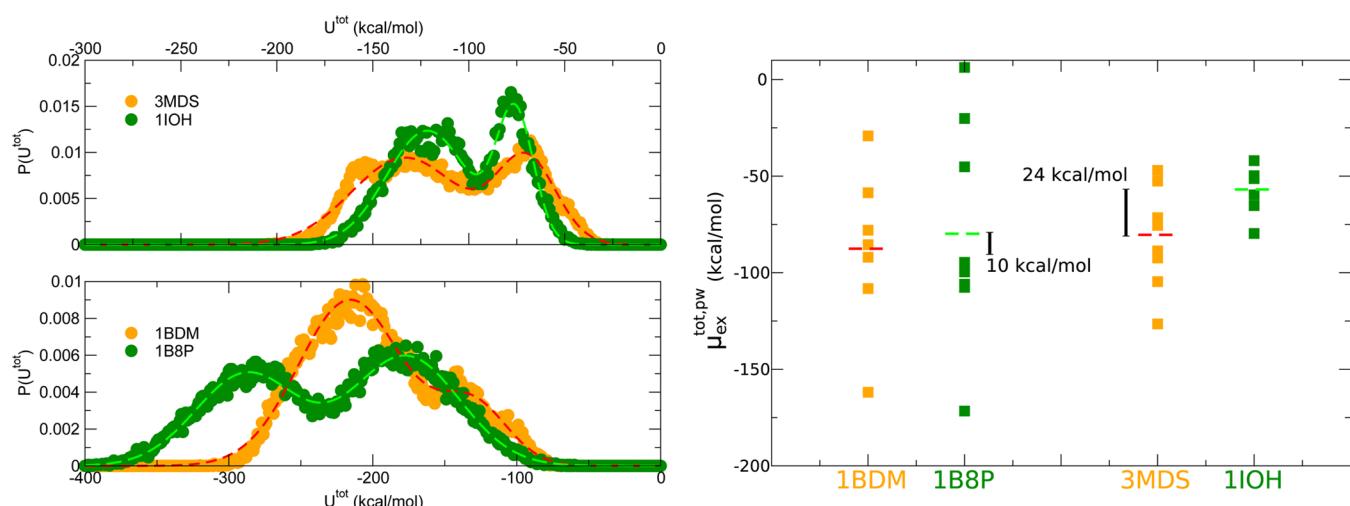


Figure 5. Left panel: probability distribution of the total interaction energy between internal water and the protein. A fit with a two-Gaussian function is represented as dashed lines. Right panel: spectra of values of the $\mu_{\text{ex}}^{\text{tot},\text{pw}}$ extracted using the Gaussian approximation for representative states of the internal hydration. Average values are represented as dashed horizontal lines, and the gap estimated between the thermophilic and mesophilic proteins are indicated.

For several stretches of the trajectory, where locally the behavior of U^{tot} is steady, we have computed the total excess chemical potential using the Gaussian framework, $\mu_{\text{ex}}^{\text{tot},\text{pw}} = \langle U^{\text{tot}} \rangle + \frac{1}{2}\beta(\sigma^{\text{tot}})^2$. This strategy allows us to focus on internal hydration states that are not altered by the in-and-out exchange events. The Gaussian calculations were performed for representative states that cover all the range of values experienced by U^{tot} (see Figure S3). In the right part of Figure 5, we report the spectrum of the extracted values of $\mu_{\text{ex}}^{\text{tot},\text{pw}}$ and of the average $\langle \mu_{\text{ex}}^{\text{tot},\text{pw}} \rangle$. For the two pairs, we probe that the internal hydration provides an important contribution to the protein stability, and that this contribution is larger for the thermophilic variants than for the mesophilic ones. Thus, even if the thermophilic and mesophilic variants host on average the same number of internal water (see Table 2), the thermophilic species have a larger gain from interior wetting. For the pair 1BDM/1B8P, we have verified that the contribution of the interaction energies between the internal water molecules themselves is of a few kcal/mol, about 1% of the total interaction energy of these water with the protein.

Water Retention. Thermophilic proteins are stable and functional at high temperatures; therefore, it is important to understand how internal wetting evolves upon thermal excitation approaching the optimal working conditions for these

extremophiles. For the thermophilic homologues considered in this work, the optimal working temperature of the hosting organism is between 330 and 360 K. For all pairs, we have performed supplemental simulations at $T = 360$ K. The length of the generated trajectories is approximately 100 ns for each system. Using the kinetic criterion already described and accounting for the thermal effect on the exchange kinetics, we calculated how many long-residence water molecules solvate, on average, the internal protein matrix at higher T values. For all pairs but one (1DZ3/3H1G), temperature increase favors internal dewetting, with a drop of the amount of internal water molecules of about 40–70%.

To consider in more detail how the dewetting occurs for the pairs 1BDM/1B8P and 3MDS/1IOH, we have explored the behavior at intermediate temperatures, including new simulations at $T = 320$ and 340 K. The results obtained are summarized in Figure 6. For the pair 1BDM/1B8P, the thermophilic enzyme exhibits a greater capacity of water retention upon thermal excitation and, therefore, can benefit from a systematic gain from internal hydration at a higher temperature window encompassing its experimental optimal working temperature, $T \approx 335$ K. For the pair 3MDS/1IOH, water retention in the thermophilic variant is less striking, although it is observed up to 360 K.

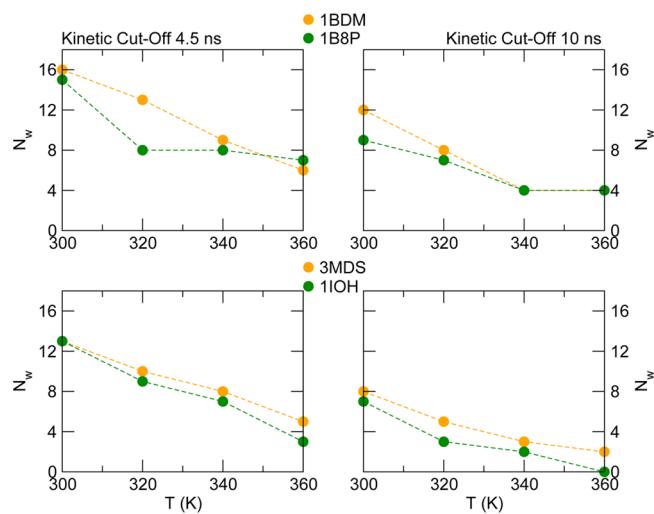


Figure 6. Variation of long-residence water molecules buried in the protein interior as a function of temperature. The charts in the top panel refer to the pair 1BDM/1B8P, and the charts in the bottom panel refer to the pair 3MDS/1IOH. For each pair, on the left and right charts we report the data using a kinetic cutoff of 4.5 and 10 ns, respectively.

Internal Water, Stability, And Function. As a final step, we inquired whether the localization of the long-residence water

molecules correlate to proteins functional sites. For this purpose, we have constructed a spatial density map of the long-residence water molecules that allows us to screen the most-occupied sites in the proteins' structure (Figure 7). The general finding is that some long-residence water can be in fact found in the active site, but most are scattered away. The presence of long-residence water in the active site is not surprising per se because on the basis of an excluded-volume effect, the exchange dynamics are expected to be slower than for the external hydration layer. However, what the mapping shows is that the localization of a potential stabilizing factor as buried water is preferential in peripheral regions of the proteins. This can be appreciated in the setting of Figure S6. For the sake of example, we discuss here two different cases. In the monomer of the malate dehydrogenase (pair 1BDM/1B8U), according to our kinetic cutoff, no long-residence water cluster is present in the binding site, implying that the binding site is regularly washed. In previous work on malate proteins,⁵⁵ it was shown that, when considered in its monomeric state, the protein loop gating the access to the binding site is highly flexible in both the mesophilic and thermophilic species, and this probably eases the turnover of water in the site. We can also observe that in the functional multimeric state (see Figure S6), the long-residence waters are largely distributed at the domain interfaces. This finding points out that for these tetrameric proteins, the stabilizing water

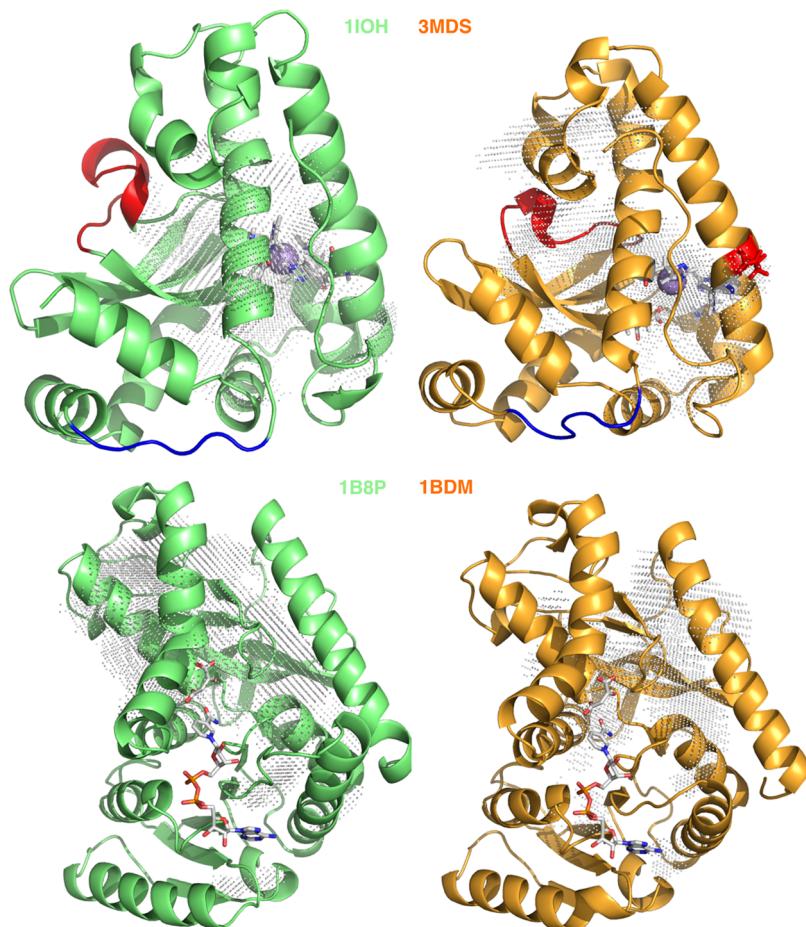


Figure 7. Density map of long-residence water molecule overlapped to the X-ray crystallographic structure of the mesophilic (green) and thermophilic (orange) proteins. Top panel: manganese superoxide dismutase highlighting the position of the manganese in the original Protein Data Bank structure (sphere), with the amino acids of the active site (licorice) and region individuated as a key for thermal stability and function (red and blue regions). Bottom panel: a single domain of the malate dehydrogenase, highlighting the ligand from the original Protein Data Bank structure (1B8U).

molecules play a role in the stability as well as in the internal allosteric sliding of the domains' interface. On the contrary for the manganese superoxide dismutases (pair IOH/3MDS), long-residence waters are found in the active site and in proximity of some residues considered key for activity as well as thermal stability.⁶⁹ Although the presence of structural water in the binding site is visible in some pairs, our finding shows, as expected, that a stability route by internal hydration could be better achieved by localizing the watery bricks in the region of the protein matrix not correlated to a function, thus challenging the stability and function trade-off. This rule of thumb adds to other empirical evidence showing, for example, that artificial thermal stabilization is more efficiently achieved by modifying charged residues at the protein surface because this avoids distorting the hydrophobic internal packing,⁷ or that it is safer to cumulate mutations far from the active site, thus avoiding function knockout.⁷⁰

CONCLUSIONS

We have performed a systematic investigation of the role of internal hydration on protein stability by considering an ensemble of mesophilic and thermophilic homologues. The study is based on a computationally affordable framework for the estimate of internal hydration free energy. When focusing on the hydration free energy of individual water molecules, we note that the internal cavities in thermophilic and mesophilic homologues offer, in most of the cases, an equally favorable environment to water. However, for some pairs, a marked larger and more favorable contribution is estimated for thermophilic variants, with differences among the homologues as high as 10 kcal/mol. Moreover, when considering the overall wetting of the internal sites, thus focusing on the ensemble of water simultaneously located in the interiors of the proteins, we appreciate that the contribution to the stability of thermophilic species can be rather important when compared to the results in the mesophilic variants. The analysis of the thermal perturbation of the internal wetting performed on selected cases in the study also highlights the capability of thermophilic proteins to retain to a larger extent the internal hydration state, thus making the favorable contribution to stability from internal water effective in a broader range of temperatures, up to their optimal working condition. Our findings suggest, in agreement with Yutani et al.,⁷¹ that the design of internal cavity hydration is a valuable strategy to be tested for the creation of mutants of enhanced thermal stability.

Finally, the highlighted correlation among internal wetting and the variability of protein conformational states opens a new perspective for investigating the relationship among mechanical and thermal stability, as was already proposed in the context of pressure-driven unfolding.^{38,72} For example, an NMR experiment could be designed^{60,73} to probe the correlation among exchange kinetics, mechanical fluctuations, and thermal stabilities in homologues.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.jpcb.5b05791](https://doi.org/10.1021/acs.jpcb.5b05791).

Figures showing probability distribution of the potential energy of a target water molecule in the hyperthermophilic G domain and resulting from the interaction of the internal water with proteins, local chemistry of the internal cavities samples by long-residence water, the number of

conformational states visited by the mesophilic and thermophilic malate dehydrogenase single domains, molecular representations of protein internal hydration, a density map of long-residence water with respect to the X-ray protein structures, and a schematic representation of internal hydration contribution to protein stability, and tables showing the ensemble of long-residence water molecules used for the free-energy calculations, data for interaction energy distributions, correlation among conformational and internal hydration states, and locations of buried hydrated sites. ([PDF](#))

AUTHOR INFORMATION

Corresponding Author

*E-mail: fabio.sterpone@ibpc.fr.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) grant agreement no. 258748. Part of this work was performed using HPC resources from GENCI [CINES and TGCC] (grants x201376818 and x201476818). We acknowledge financial support for infrastructures from ANR-11-LABX-0011-01 (Dynamics of Energy Transducing Membranes: Biogenesis and SupraMolecular Organization). We thank G. Villain for his contribution to the selection of the proteins studied in this work.

REFERENCES

- (1) Vieille, C.; Zeikus, G. J. Hyperthermophilic Enzymes: Sources, Uses, and Molecular Mechanisms for Thermostability. *Microbiol. Mol. Biol. Rev.* **2001**, *65*, 1–43.
- (2) Kumar, S.; Nussinov, R. How do Thermophilic Proteins Deal with Heat? *Cell. Mol. Life Sci.* **2001**, *58*, 1216–1233.
- (3) Sterpone, F.; Melchionna, S. Thermophilic Proteins: Insight and Perspective from In Silico Experiments. *Chem. Soc. Rev.* **2012**, *41*, 1665–1676.
- (4) Kumar, S.; Ma, B.; Tsai, C. J.; Nussinov, R. Electrostatic Strengths of Salt Bridges in Thermophilic and Mesophilic Glutamate Dehydrogenase Monomers. *Proteins: Struct., Funct., Genet.* **2000**, *38*, 368–83.
- (5) Karshikoff, A.; Ladenstein, R. Ion Pairs and the Thermotolerance of Proteins from Hyperthermophiles: a "Traffic Rule" for Hot Roads. *Trends Biochem. Sci.* **2001**, *26*, 550–6.
- (6) Chan, C.-H.; Yu, T.-H.; Wong, K.-B. Stabilizing Salt-Bridge Enhances Protein Thermostability by Reducing the Heat Capacity Change of Unfolding. *PLoS One* **2011**, *6*, e21624.
- (7) Sanchez-Ruiz, J. M.; Makhatadze, G. I. To Charge or not to Charge? *Trends Biotechnol.* **2001**, *19*, 132–5.
- (8) Gribenko, A. V.; Patel, M. M.; Liu, J.; McCallum, S. A.; Wang, C.; Makhatadze, G. I. Rational Stabilization of Enzymes by Computational Redesign of Surface Charge-Charge Interactions. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 2601–6.
- (9) Rathi, P. C.; Höffken, H. W.; Gohlke, H. Quality Matters: Extension of Clusters of Residues with Good Hydrophobic Contacts Stabilize (Hyper)Thermophilic Proteins. *J. Chem. Inf. Model.* **2014**, *54*, 355–361.
- (10) Nojima, H.; Ikai, A.; Oshima, T.; Noda, H. Reversible Thermal Unfolding of Thermostable Phosphoglycerate Kinase. Thermostability Associated With Mean Zero Enthalpy Change. *J. Mol. Biol.* **1977**, *116*, 429–442.
- (11) Liu, C.-C.; LiCata, V. J. The stability of Taq DNA Polymerase Results from a Reduced Entropic Folding Penalty; Identification of

- other Thermophilic Proteins with Similar Folding Thermodynamics. *Proteins: Struct., Funct., Genet.* **2014**, *82*, 785–793.
- (12) Razvi, A.; Scholtz, J. M. Lessons in Stability from Thermophilic Proteins. *Protein Sci.* **2006**, *15*, 1569–1578.
- (13) Samish, I.; MacDermaid, C. M.; Perez-Aguilar, J.; Saven, J. G. Theoretical and Computational Protein Design. *Annu. Rev. Phys. Chem.* **2011**, *62*, 129–149.
- (14) Koga, N.; Tatsumi-Koga, R.; Liu, G.; Xiao, R.; Acton, T. B.; Montelione, G.; Baker, D. Principles for designing ideal protein structures. *Nature* **2012**, *491*, 222–227.
- (15) Kiss, G.; Celebi-Olcum, N.; Moretti, R.; Baker, D.; Houk, K. N. Computational Enzyme Design. *Angew. Chem., Int. Ed.* **2013**, *52*, 5700–5725.
- (16) Chennamsetty, N.; Voynov, V.; Kayser, V.; Helk, B.; Trout, B. L. Design of therapeutic proteins with enhanced stability. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 11937–11942.
- (17) Xu, J.; Baase, W.; Quillin, M.; Baldwin, E.; Matthews, B. Structural and Thermodynamic Analysis of the Binding of Solvent at Internal Sites in T4 Lysozyme. *Protein Sci.* **2001**, *10*, 1067–1078.
- (18) Pedersen, J. T.; Olsen, O. H.; Betzel, C.; Eschenburg, S.; Branner, S.; Hastrup, S. Cavity Mutants of SavinaseTM: Crystal Structures and Differential Scanning Calorimetry Experiments Give Hints of the Function of the Buried Water Molecules in Subtilisins. *J. Mol. Biol.* **1994**, *242*, 193–202.
- (19) Berndt, K.; Beunink, J.; Schroeder, W.; Wuethrich, K. Designed Replacement of an Internal Hydration Water Molecule in BPTI: Structural and Functional Implications of a Glycine-to-Serine Mutation. *Biochemistry* **1993**, *32*, 4564–4570.
- (20) Rahaman, O.; Kalimeri, M.; Melchionna, S.; Hénin, J.; Sterpone, F. Role of Internal Water on Protein Thermal Stability: The Case of Homologous G Domains. *J. Phys. Chem. B* **2015**, *119*, 8939–8949.
- (21) Talon, R.; Coquelle, N.; Madern, D.; Girard, E. An Experimental Point of View on Hydration/Solvation in Halophilic Proteins. *Front. Microbiol.* **2014**, *5*, 66.
- (22) Ahmad, S.; Kamal, M. Z.; Sankaranarayanan, R.; Rao, N. M. Thermostable Bacillus Subtilis Lipases: In Vitro Evolution and Structural Insight. *J. Mol. Biol.* **2008**, *381*, 324–40.
- (23) Chi, Y. I.; Martinez-Cruz, L. A.; Jancarik, J.; Swanson, R. V.; Robertson, D. E.; Kim, S. H. Crystal Structure of the Beta-Glycosidase from the Hyperthermophile Thermosphaera Aggregans: Insights into its Activity and Thermostability. *FEBS Lett.* **1999**, *445*, 375–383.
- (24) Melchionna, S.; Sinibaldi, R.; Briganti, G. Explanation of the Stability of Thermophilic Proteins Based on Unique Micromorphology. *Biophys. J.* **2006**, *90*, 4204–12.
- (25) Sterpone, F.; Bertonati, C.; Briganti, G.; Melchionna, S. Key Role of Proximal Water in Regulating Thermostable Proteins. *J. Phys. Chem. B* **2009**, *113*, 131–7.
- (26) Yin, H.; Hummer, G.; Rasaiah, J. C. Metastable Water Clusters in the Nonpolar Cavities of the Thermostable Protein Tetrabrachion. *J. Am. Chem. Soc.* **2007**, *129*, 7369–77.
- (27) Mou, Z.; Ding, Y.; Wang, X.; Cai, Y. Comparison of Protein-water Interactions in Psychrophilic, Mesophilic, and Thermophilic Fe-SOD. *Protein Pept. Lett.* **2014**, *21*, 578–583.
- (28) Roche, J.; Caro, J. A.; Norberto, D. R.; Barthe, P.; Roumestand, C.; Schlessman, J. L.; Garcia, A. E.; García-Moreno, B. E.; Royer, C. A. Cavities Determine the Pressure Unfolding of Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 6945–6950.
- (29) Nisius, L.; Grzesiek, S. Key Stabilizing Elements of Protein Structure Identified through Pressure and Temperature Perturbation of its Hydrogen Bond Network. *Nat. Chem.* **2012**, *4*, 711–717.
- (30) Takano, K.; Funahashi, J.; Yamagata, Y.; Fujii, S.; Yutani, K. Contribution of Water Molecules in the Interior of a Protein to the Conformational Stability. *J. Mol. Biol.* **1997**, *274*, 132–142.
- (31) Hickey, D. R.; Berghuis, A. M.; Lafond, G.; Jaeger, J. A.; Cardillo, T. S.; McLendon, D.; Das, G.; Sherman, F.; Brayer, G. D.; McLendon, G. Enhanced Thermodynamic Stabilities of Yeast Iso-1-Cytochromes C with Amino Acid Replacements at Positions 52 and 102. *J. Biol. Chem.* **1991**, *266*, 11686–11694.
- (32) Vriend, G.; Berendsen, H.; van der Zee, J.; van den Burg, B.; Venema, G.; Eijsink, V. Stabilization of the Neutral Protease of Bacillus Stearothermophilus by Removal of a Buried Water Molecule. *Protein Eng. Des. Sel.* **1991**, *4*, 941–945.
- (33) Wade, R. C.; Mazor, M. H.; McCammon, J. A.; Quiocio, F. A. A Molecular Dynamics Study of Thermodynamic and Structural Aspects of the Hydration of Cavities in Proteins. *Biopolymers* **1991**, *31*, 919–931.
- (34) Roux, B.; Nina, M.; Pomes, R.; Smith, J. Thermodynamic Stability of Water Molecules in the Bacteriorhodopsin Proton Channel: a Molecular Dynamics Free Energy Perturbation Study. *Biophys. J.* **1996**, *71*, 670–681.
- (35) Sterpone, F.; Ceccarelli, M.; Marchi, M. Dynamics of Hydration in Hen Egg White Lysozyme. *J. Mol. Biol.* **2001**, *311*, 409–19.
- (36) Olano, L. R.; Rick, S. W. Hydration Free Energies and Entropies for Water in Protein Interiors. *J. Am. Chem. Soc.* **2004**, *126*, 7991–8000.
- (37) Rasaiah, J. C.; Garde, S.; Hummer, G. Water in Nonpolar Confinement: from Nanotubes to Proteins and beyond. *Annu. Rev. Phys. Chem.* **2008**, *59*, 713–40.
- (38) Damjanovic, A.; Schlessman, J.; Fitch, C.; Garcia, A.; Garcia-Moreno, E. Role of Flexibility and Polarity as Determinants of the Hydration of Internal Cavities and Pockets in Proteins. *Biophys. J.* **2007**, *93*, 2791–2804.
- (39) Chakrabarty, S.; Warshel, A. Capturing the Energetics of Water Insertion in Biological Systems: The Water Flooding Approach. *Proteins: Struct., Funct., Genet.* **2013**, *81*, 93–106.
- (40) Oikawa, M.; Yonetani, Y. Molecular Dynamics Free Energy Calculations to Assess the Possibility of Water Existence in Protein Nonpolar Cavities. *Biophys. J.* **2010**, *98*, 2974–2983.
- (41) De Simone, A.; Dodson, G. G.; Verma, C. S.; Zagari, A.; Fraternali, F. Prion and Water: Tight and Dynamical Hydration Sites Have a Key Role in Structural Stability. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 7535–7540.
- (42) Baudry, J.; Tajkhorshid, E.; Molnar, F.; Phillips, J.; Schulten, K. Molecular Dynamics Study of Bacteriorhodopsin and the Purple Membrane. *J. Phys. Chem. B* **2001**, *105*, 905–918.
- (43) Tashiro, M.; Stuchebrukhov, A. A. Thermodynamic Properties of Internal Water Molecules in the Hydrophobic Cavity around the Catalytic Center of Cytochrome c Oxidase. *J. Phys. Chem. B* **2005**, *109*, 1015–1022.
- (44) Pisliakov, A. V.; Sharma, P. K.; Chu, Z. T.; Haranczyk, M.; Warshel, A. Electrostatic Basis for the Unidirectionality of the Primary Proton Transfer in Cytochrome C Oxidase. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 7726–7731.
- (45) Lee, H. J.; Svahn, E.; Swanson, J. M. J.; Lepp, H.; Voth, G. A.; Brzezinski, P.; Gennis, R. B. Intricate Role of Water in Proton Transport through Cytochrome c Oxidase. *J. Am. Chem. Soc.* **2010**, *132*, 16225–16239.
- (46) Helms, V.; Wade, R. Thermodynamics of Water Mediating Protein-Ligand Interactions in Cytochrome P450cam: a Molecular Dynamics Study. *Biophys. J.* **1995**, *69*, 810–824.
- (47) Li, Z.; Lazaridis, T. Thermodynamic Contributions of the Ordered Water Molecule in HIV-1 Protease. *J. Am. Chem. Soc.* **2003**, *125*, 6636–6637.
- (48) Baron, R.; Setny, P.; McCammon, J. A. Water in Cavity-Ligand Recognition. *J. Am. Chem. Soc.* **2010**, *132*, 12091–12097.
- (49) Szep, S.; Park, S.; Boder, E. T.; van Duyne, G. D.; Saven, J. G. Structural Coupling between FKBP12 and Buried Water. *Proteins: Struct., Funct., Genet.* **2009**, *74*, 603–611.
- (50) Prakash, P.; Sayyed-Ahmad, A.; Gorfe, A. A. The Role of Conserved Waters in Conformational Transitions of Q61H K-ras. *PLoS Comput. Biol.* **2012**, *8*, e1002394.
- (51) Scorcipino, M. A.; Robertazzi, A.; Casu, M.; Ruggerone, P.; Ceccarelli, M. Heme proteins: the role of solvent in the dynamics of gates and portals. *J. Am. Chem. Soc.* **2010**, *132*, 5156–63.
- (52) Rahaman, O.; Melchionna, S.; Laage, D.; Sterpone, F. The Effect of Protein Composition on Hydration Dynamics. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3570–3576.

- (53) Kalimeri, M.; Rahaman, O.; Melchionna, S.; Sterpone, F. How Conformational Flexibility Stabilizes the Hyperthermophilic Elongation Factor G-Domain. *J. Phys. Chem. B* **2013**, *117*, 13775–13785.
- (54) Pechkova, E.; Sivozhelezov, V.; Nicolini, C. Protein Thermal Stability: the Role of Protein Structure and Aqueous Environment. *Arch. Biochem. Biophys.* **2007**, *466*, 40–8.
- (55) Kalimeri, M.; Girard, E.; Madern, D.; Sterpone, F. Interface Matters: The Stiffness Route to Stability of a Thermophilic Tetrameric Malate Dehydrogenase. *PLoS One* **2014**, *9*, e113895.
- (56) MacKerell, A. D.; Bashford, D.; Dunbrack, R.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins†. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (57) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (58) Garcia, A. E.; Hummer, G. Water Penetration and Escape in Proteins. *Proteins: Struct., Funct., Genet.* **2000**, *38*, 261–72.
- (59) Halle, B. Protein Hydration Dynamics in Solution: a Critical Survey. *Philos. Trans. R. Soc., B* **2004**, *359*, 1207–23.
- (60) Persson, F.; Halle, B. Transient Access to the Protein Interior: Simulation versus NMR. *J. Am. Chem. Soc.* **2013**, *135*, 8735–8748.
- (61) Tuckerman, M. *Statistical Mechanics. Theory and Molecular Simulation*; Oxford University Press: Oxford, England, 2010.
- (62) Cui, G.; Swails, J. M.; Manas, E. S. SPAM: A Simple Approach for Profiling Bound Water Molecules. *J. Chem. Theory Comput.* **2013**, *9*, 5539–5549.
- (63) Huggins, D. J. Quantifying the Entropy of Binding for Water Molecules in Protein Cavities by Computing Correlations. *Biophys. J.* **2015**, *108*, 928–936.
- (64) Hummer, G.; Szabo, A. Calculation of freeenergy differences from computer simulations of initial and final states. *J. Chem. Phys.* **1996**, *105*, 2004–2010.
- (65) Petrone, P. M.; Garcia, A. E. MHC-Peptide Binding is Assisted by Bound Water Molecules. *J. Mol. Biol.* **2004**, *338*, 419–435.
- (66) Day, R.; García, A. E. Water penetration in the low and high pressure native states of ubiquitin. *Proteins: Struct., Funct., Genet.* **2008**, *70*, 1175–84.
- (67) Tian, J.; García, A. E. Simulations of the confinement of ubiquitin in self-assembled reverse micelles. *J. Chem. Phys.* **2011**, *134*, 225101.
- (68) Widom, B. Some topics in theory of fluids. *J. Chem. Phys.* **1963**, *39*, 2802–2812.
- (69) Ludwig, M. L.; Metzger, A. L.; Patridge, K. A.; Stallings, W. C. Manganese superoxide dismutase from *Thermus thermophilus*. A structural model refined at 1.8 Å resolution. *J. Mol. Biol.* **1991**, *219*, 335–358.
- (70) Morley, K. L.; Kazlauskas, R. J. Improving enzyme properties: when are closer mutations better? *Trends Biotechnol.* **2005**, *23*, 231–237.
- (71) Takano, K.; Yamagata, Y.; Yutani, K. Buried water molecules contribute to the conformational stability of a protein. *Protein Eng., Des. Sel.* **2003**, *16*, 5–9.
- (72) Roche, J.; Dellarole, M.; Caro, J. A.; Norberto, D. R.; Garcia, A. E.; Garcia-Moreno, B. E.; Rou mestand, C.; Royer, C. A. Effect of Internal Cavities on Folding Rates and Routes Revealed by Real-Time Pressure-Jump NMR Spectroscopy. *J. Am. Chem. Soc.* **2013**, *135*, 14610–14618.
- (73) Denisov, V. P.; Peters, J.; Hörllein, H. D.; Halle, B. Using Buried Water Molecules to Explore the Energy Landscape of Proteins. *Nat. Struct. Biol.* **1996**, *3*, 505–509.