

Protein Cold Denaturation in Implicit Solvent Simulations: A Transfer Free Energy Approach

Andrea Arsiccio[†] and Joan-Emma Shea^{*,†,‡}

*†Department of Chemistry and Biochemistry, University of California, Santa Barbara,
California 93106, United States*

*‡Department of Physics, University of California, Santa Barbara, California 93106, United
States*

E-mail: shea@ucsb.edu

Abstract

Proteins are stable over a narrow temperature range, with either hot and cold denaturation occurring outside of this window, both of which adversely affect protein function. While hot unfolding is entropically-driven, cold denaturation, on the other hand, results from the more favorable interaction of water with apolar groups at low temperature. Because of the key role of water in this latter process, capturing cold-denaturation using implicit solvent models is challenging. We propose here a novel computational approach to develop an implicit solvent model that accounts for both hot and cold denaturation in simulations involving atomistically detailed protein representations. By mining a large number of protein structures solved by nuclear magnetic resonance, we derive transfer free energy contributions for the backbone and amino acids side chains representing the transfer of these moieties between water at two different temperatures. Using Trp-cage as a model system, we show that the implicit solvent model constructed using these temperature-dependent free energies of transfer recovers the parabolic temperature dependence of protein stability, capturing both hot and cold denaturation. The resulting cold-unfolded conformations show limited secondary structure content, but preserve most of their internal hydrogen-bonding network, in contrast to the extended configurations with no hydrogen-bonding populated during heat-induced denaturation.

Introduction

Proteins operate in an optimal manner within a given temperature range, and their stability curves display a maximum at a protein-specific value of temperature, with both higher and lower temperatures leading to a reduction in the fraction of folded protein. Heat denaturation is driven by an entropic mechanism, with high temperatures promoting the increase in conformational entropy observed upon unfolding. The resulting heat denatured conformations are very expanded, with a substantial (or complete) loss in secondary structure. Cold denaturation, on the contrary, is enthalpically driven, and caused by the reduced penalty

involved in hydrating apolar groups at low temperature.¹⁻⁶ Cold-denatured structures are generally quite compact, often with a significant level of residual secondary structure, and a high degree of solvent penetration.⁷⁻¹⁰ The interaction between water and the protein plays a key role in cold unfolding, and for this reason, computational studies investigating this process have primarily relied on explicit solvent simulations, which give the most detailed representation of the solvent.¹⁰⁻¹⁴ The limitation of explicit solvent simulations, however, is that despite the increase in computational power that we have seen over the last decade and the emergence of efficient sampling algorithms, it is computationally prohibitive to obtain converged trajectories at low temperatures, as the reduced mobility under these conditions demands very long equilibration times. This obstacle to characterizing cold unfolding can in principle be addressed by adopting an implicit description of the solvent. Conformational transitions are indeed faster in implicit-solvent simulations as the computational cost is considerably alleviated by the absence of the degrees of freedom corresponding to the water molecules. Moreover, the absence of viscosity when the solvent is removed accelerates the exploration of the conformational phase space, especially at low temperature.

The majority of existing implicit solvent models are not able to account for cold denaturation, and efforts in developing implicit models that capture cold denaturation have focused thus far on models that interface with a coarse-grained description of the protein. For example Sirovetz et al.¹⁵ added a linear temperature-dependent perturbation term as well as a pressure-dependent perturbation to the coarse-grained AWSEM force field¹⁶ and were able to predict cold- and pressure-denaturation of ubiquitin and λ -repressor. van Dijk et al.¹⁷ added an effective potential for hydrophobe-water interactions to a protein lattice model, and they were able to reproduce cold-denaturation and heat capacity measurements of solvated proteins. This added potential consisted of a second-order approximation to the free energy of transfer of hydrophobic particles from an oily environment to water.

Here, we propose a computational approach in which we introduce a temperature-dependent non-polar contribution to the free energy of hydration and combine our approach with a

fully-atomistic representation of the protein structure, allowing for greater molecular detail and closer comparison to experiment compared to coarse-grained or cubic lattice models used in previous studies.^{15,17} The introduction of a temperature dependent term is critical in order to capture cold denaturation. Importantly, the parameters involved in the temperature-dependent contribution in our proposed model do not depend on the protein model considered, and are hence generally applicable.

Our approach is briefly described below, with details given in the section "Theoretical Background". We consider the free energy of hydration of a peptide as the sum of a non-polar (G^{np}) and an electrostatic (G^{el}) contributions, and consider a solvent accessible surface area (*SASA*) description of the non-polar solvation term as $G^{np} = \gamma SASA$, where γ is the surface tension. We note that the non-polar solvation term could also be described by adding a volume term¹⁸ and/or decomposing it into repulsive and attractive components,^{18,19} however, for the sake of computational efficiency and model simplicity, we will use the most commonly used *SASA*-based description of the non-polar solvation term which is based on arguments from scaled particle theory.^{20,21} In most implicit solvent models, γ is given a single value, independent of temperature and atom type. As we will show, the use of a single valued γ does not allow for the description of cold denaturation. Some implicit models have moved beyond this simplistic approach, for example with atom-type-specific γ parameters obtained by fitting the free energies of transfer of side chain analogs from *n*-octanol to water,²² or from vapor to water.²³ Other approaches that move beyond a single value for the surface tension include a simple approach in which two γ values, determined through a trial-and-error approach, are used to represent either the hydrophobic or the hydrophilic effect.^{24,25} A more sophisticated approach that uses hydration free energies of side chain analogs as starting experimental data to obtain temperature-dependent values of γ ²⁶ has been proposed, however, this approach assumes the property of additivity, i.e., the possibility to compute the hydration free energy of a molecule as a sum over the separate values of its components, a hypothesis that has been challenged.^{27,28} We propose to use a

different means of obtaining temperature-dependent values of γ by selecting a different set of starting experimental data. In our approach, we derive the temperature-dependent term by mining a large number of protein structures resolved by nuclear magnetic resonance (NMR) in the range 265-335 K, an approach that has been used to extract temperature-dependent hydrophobicity propensities of different amino acids.²⁹

The newly proposed implicit description of protein hydration is tested on the α -helical Trp-cage (NLYIQWLKDGGPSSGRPPPS, pdb 1L2Y³⁰). This 20-residue peptide has been selected because it has been extensively characterized both computationally and experimentally, has secondary and tertiary structure despite its small size, and is known to be a fast folder.³¹⁻⁴⁰ Furthermore, its cold denaturation has also been studied computationally in explicit solvent, offering a means of comparison.^{10,13} We show that the addition of the proposed temperature-dependent term allows the description of cold denaturation in implicit solvent simulations providing an efficient and accurate alternative to costly explicit solvent simulations.

Materials and Methods

Theoretical Background

The following expression can be used to describe the free energy of a peptide dissolved in water,

$$G^{tot} = E^{vac} + G^{el} + G^{mp} \tag{1}$$

where E^{vac} is the peptide energy in vacuum, which results from both internal bonded contributions (bonds, angles, dihedrals) and non-bonded van der Waals interactions. G^{el} and G^{mp} represent instead the polar and non-polar contributions, respectively. The non-polar contribution G^{mp} to the free energy of hydration is generally expressed as a SASA-based

term. We here aim at improving such term, so as to eventually recover the description of cold unfolding in our implicit solvent simulations. For this purpose, we will compare three different approaches. Approach 1 corresponds to the standard implementation, in which there is a single temperature independent value for γ . As we will show, this standard implementation is unable to capture cold denaturation, and we use this model as a reference to evaluate how our new models improve over the standard implementation. Approach 2 preserves the standard implementation only at $T_0 = 298$ K, and employs different values of γ at $T \neq T_0$. Finally, approach 3 introduces a completely new set of γ values. The relative merits of approaches 2 and 3 will be assessed in terms of their ability to successfully describe the free energy landscape of Trp-cage.

Approach 1 - Standard implementation

In this case we will simply compute G^{np} as,

$$G^{np} = \gamma_0 \text{SASA} \tag{2}$$

where γ_0 is a surface tension, independent of temperature and residue-type, while SASA is the total solvent accessibility of the protein. This is the approach that is commonly implemented in molecular dynamics software, like the AMBER 20 simulation suite employed in this work, where the default γ_0 value is $5 \text{ cal mol}^{-1} \text{ \AA}^{-2}$.⁴¹

Approach 2 - New γ values for $T \neq T_0$

In this approach, we seek to reproduce the difference in the non-polar solvation term between a specific temperature T and a reference temperature $T_0 = 298$ K. At T_0 , the default γ_0 value of $5 \text{ cal mol}^{-1} \text{ \AA}^{-2}$, independent of residue type, will still be used. In other words, we will keep the implicit solvent description unaltered at T_0 , and search different γ values only at $T \neq T_0$.

Consider a set Ω of protein structures, comprising N total residues, of which N_i belong

to residue type i . We aim to extract from this set proper values of surface tension (γ) parameters to be used in the framework of our implicit solvent description. Considering the dominant contribution of the backbone in driving protein folding,^{42,43} we aim to obtain separate values of surface tensions for the different side chains, and for the backbone. The first step for obtaining such γ parameters is to compute the probability P_{nb} (the subscript nb is here used as the abbreviation of "not buried") for a given side chain, or for the backbone, to be surface exposed.

The probability $P_{nb}^{sc,i}$ of residue i to expose its side chain (superscript sc) to the aqueous environment can be computed as,

$$P_{nb}^{sc,i} = \frac{1}{N_i} \sum_{k=1}^{N_i} \left(\frac{SASA_k^{sc,i}}{SASA_{max}^{sc,i}} \right) \quad (3)$$

where $SASA_k^{sc,i}$ is the side chain solvent accessibility of the k^{th} residue type i in the set Ω , and $SASA_{max}^{sc,i}$ is the maximum solvent accessibility of side chain type i . The values of $SASA_{max}^{sc,i}$ used in this work were extracted from Creamer et al.⁴⁴ Creamer et al. developed two models that bracket the surface area of the unfolded state between limiting extremes. Here, we made use of the upper bound model, corresponding to an extended conformation. This is more physically grounded than using the tripeptide representation, GLY- i -GLY, for computing $SASA_{max}^{sc,i}$, as the tripeptide model was shown to overestimate the surface area of side chains in the unfolded state.⁴⁵

The probability P_{nb}^{bb} of the backbone (superscript bb) to be surface-exposed can be extracted from the set Ω as,

$$P_{nb}^{bb} = \frac{1}{N} \sum_{k=1}^N \left(\frac{SASA_k^{bb}}{SASA_{max}^{bb}} \right) \quad (4)$$

where the summation runs over all residues N is the set Ω , and the maximum solvent accessibility of the backbone $SASA_{max}^{bb}$ was again extracted from Creamer et al.⁴⁴

Once the probability P_{nb} is known, the corresponding energy can be computed,

$$E_{nb}^{sc,i \text{ or } bb} = -RT \ln P_{nb}^{sc,i \text{ or } bb} \quad (5)$$

$E_{nb}^{sc,i \text{ or } bb}$ is the energy of side chain type i /peptide backbone when surface exposed. In contrast, the energy of buried (subscript b) side chain type i /peptide backbone can be calculated as,

$$E_b^{sc,i \text{ or } bb} = -RT \ln(1 - P_{nb}^{sc,i \text{ or } bb}) \quad (6)$$

The energetic cost for transferring side chain type i , or the peptide backbone, from the protein core to the solvent accessible surface can hence be computed as,

$$E_{nb}^{sc,i \text{ or } bb} - E_b^{sc,i \text{ or } bb} = -RT \ln \left(\frac{P_{nb}^{sc,i \text{ or } bb}}{1 - P_{nb}^{sc,i \text{ or } bb}} \right) \quad (7)$$

The following expression can then be used to obtain γ_2 values (the subscript 2 here refers to approach number 2) at $T \neq T_0$,

$$\left(E_{nb}^{sc,i \text{ or } bb} - E_b^{sc,i \text{ or } bb} \right)_T - \left(E_{nb}^{sc,i \text{ or } bb} - E_b^{sc,i \text{ or } bb} \right)_{T_0} = [\gamma_2^{sc,i \text{ or } bb}(T) - \gamma_0] SASA_{max}^{sc,i \text{ or } bb} \quad (8)$$

Eq. 8 states that the difference in the energetic cost for exposing side chain type i or the peptide backbone between two different temperatures equals the corresponding difference in surface tension, multiplied by the change in surface area (the solvent accessible surface area is assumed to be zero in the buried state, and equal to $SASA_{max}$ in the unfolded state). In summary, the non polar contribution G^{np} for a peptide comprising n residues can be expressed as,

$$G^{np} = \begin{cases} \gamma_0 SASA, & T = T_0 \\ \gamma_0 SASA + \sum_{k=1}^n (\gamma_2^{sc,k}(T) - \gamma_0) SASA^{sc,k} + \sum_{k=1}^n (\gamma_2^{bb}(T) - \gamma_0) SASA^{bb,k}, & T \neq T_0 \end{cases} \quad (9)$$

We can further define a fractional solvent accessibility as the ratio between the *SASA* of each side chain or backbone, and the corresponding *SASA* in the tripeptide *GLY* – *k* – *GLY*,

$$\alpha^{sc,k \text{ or } bb,k} = \frac{SASA^{sc,k \text{ or } bb,k}}{SASA_{GLY-k-GLY}^{sc,k \text{ or } bb,k}} \quad (10)$$

Using the fractional solvent accessibility we can then rewrite Eq. 9 as,

$$G^{np} = \begin{cases} \gamma_0 SASA, & T = T_0 \\ \gamma_0 SASA + \sum_{k=1}^n \Delta g_{tr,2}^{sc,k}(T) \alpha^{sc,k} + \Delta g_{tr,2}^{bb}(T) \sum_{k=1}^n \alpha^{bb,k}, & T \neq T_0 \end{cases} \quad (11)$$

where $\Delta g_{tr,2}^{sc,k \text{ or } bb} = (\gamma_2^{sc,k \text{ or } bb}(T) - \gamma_0) SASA_{GLY-k-GLY}^{sc,k \text{ or } bb,k}$.

The rationale for rewriting Eq. 9 into Eq. 11 is that Eq. 11 shares the same functional form already used for osmolytes in past work,^{46–48} with the only difference that the free energies of transfer $\Delta g_{tr,2}(T)$ values now depend on temperature, as they do not represent any more the transfer between water and an osmolyte solution, but between water at temperature T_0 and water at temperature $T \neq T_0$. This allows an unprecedented parallel between the effect of osmolytes and temperature on protein stability.

Approach 3 - A new set of γ values developed at all temperatures

In this approach, we aim at developing a completely new set of γ values, that depend on residue-type already at T_0 . For this purpose, the same procedure described in Approach 2 can be used, but Eq. 8 should be modified as,

$$\left(E_{nb}^{sc,i \text{ or } bb} - E_b^{sc,i \text{ or } bb} \right)_T = \gamma_3^{sc,i \text{ or } bb}(T) SASA_{max}^{sc,i \text{ or } bb} \quad (12)$$

and this equation could be used to obtain γ_3 values at any temperature T .

According to approach 3, the final expression for G^{np} would be,

$$G^{np} = \sum_{k=1}^n \gamma_3^{sc,k}(T) SASA^{sc,k} + \sum_{k=1}^n \gamma_3^{bb}(T) SASA^{bb,k} \quad (13)$$

and Eq. 13 could again be transformed to introduce free energy of transfer values,

$$G^{np} = \sum_{k=1}^n \Delta g_{tr,3}^{sc,k}(T) \alpha^{sc,k} + \Delta g_{tr,3}^{bb}(T) \sum_{k=1}^n \alpha^{bb,k} \quad (14)$$

where $\Delta g_{tr,3}^{sc,k \text{ or } bb} = \gamma_3^{sc,k \text{ or } bb}(T) SASA_{GLY-k-GLY}^{sc,k \text{ or } bb,k}$.

In summary, the following expression will be used in our implicit solvent simulations,

$$G^{tot} = E^{vac} + G^{el} + G_0^{np} + G^{tr}(T) \quad (15)$$

where G_0^{np} and $G^{tr}(T)$ assume the following form depending on the approach considered,

- Approach 1:

$$G_0^{np} = \gamma_0 SASA \quad (16)$$

$$G^{tr}(T) = 0 \quad \forall T \quad (17)$$

- Approach 2:

$$G_0^{np} = \gamma_0 SASA \quad (18)$$

$$G^{tr}(T) = \begin{cases} 0, & T = T_0 \\ \sum_{k=1}^n \Delta g_{tr,2}^{sc,k}(T) \alpha^{sc,k} + \Delta g_{tr,2}^{bb}(T) \sum_{k=1}^n \alpha^{bb,k}, & T \neq T_0 \end{cases} \quad (19)$$

- Approach 3:

$$G_0^{np} = 0 \quad (20)$$

$$G^{tr}(T) = \sum_{k=1}^n \Delta g_{tr,3}^{sc,k}(T) \alpha^{sc,k} + \Delta g_{tr,3}^{bb}(T) \sum_{k=1}^n \alpha^{bb,k} \quad \forall T \quad (21)$$

Simulation Details

The three approaches proposed in the Theoretical Background section were tested on Trp-cage as model protein. The initial configuration of Trp-cage was downloaded from the pdb database (pdb 1L2Y³⁰). The Amber ff99SB-ILDN force field⁴⁹ was used for the peptide. Simulations were performed using the AMBER 20 simulation suite,⁵⁰ in combination with Plumed 2.4.7.⁵¹ The generalized Born/surface area model of AMBER 20 was used to simulate the first three terms of Eq. 15, while the free energy of transfer term $G^{tr}(T)$ was added as an external bias using Plumed. The OBC(II) model⁵² was used to estimate the Born radii (IGB = 5). The effect of temperature on the dielectric constant was taken into account. For this purpose, the equations proposed in^{53,54} were used. For the lowest temperature values considered in our simulations, the dielectric constant was extrapolated from the correlation described in.⁵⁴

Trp-cage was first energy minimized for 3000 steps using the steepest descent algorithm, and then simulated for 500 ns at different temperatures spanning the range 180 - 450 K. The first 200 ns were considered as an equilibration, and only the last 300 ns were used for the subsequent analyses. Langevin dynamics was used to control the temperature, with a collision frequency of 1.0 ps^{-1} . The SHAKE algorithm⁵⁵ was applied to constrain all bonds linking to hydrogen atoms, and a time step of 2.0 fs was used. Configurations were saved every 2 ps, and the center of mass translation and rotation were removed every 500 steps (1 ps). No cut-off was used for the Coulombic and Lennard-Jones interactions.

Analysis of the Trajectories

α -helix content

The α -helix content α was defined as the number of residue sections having an α -helical configuration,⁵⁶

$$\alpha = \sum_{\mu} g[r_{dist}(\{R_i\}_{i \in \Omega_{\mu}}, \{R^0\})] \tag{22}$$

where the summation runs over all possible segments involved in the α -helix, $\{R_i\}_{i \in \Omega_{\mu}}$ are the atomic coordinates of a set Ω_{μ} of 6 residues of the protein, and $g(r_{dist})$ is a switching function

$$g(r_{dist}) = \frac{1 - \left(\frac{r_{dist}}{r_0}\right)^8}{1 - \left(\frac{r_{dist}}{r_0}\right)^{12}} \tag{23}$$

A cutoff distance of $r_0 = 0.08 \text{ nm}$ was used, and r_{dist} is the distance RMSD with respect to a reference α -helix configuration $\{R^0\}$.

Cluster analysis

The conformations assumed by the peptides during the last 300 ns of the simulation time were grouped together by means of a cluster analysis, according to the Daura algorithm.⁵⁷

The structures were grouped together if the root mean square deviations of the N-C $_{\alpha}$ -C atoms were less than 0.2 nm compared to each other. The most probable conformations were subsequently visualized using VMD (Visual Molecular Dynamics).⁵⁸

Hydrogen bond analysis

The number of intra-peptide hydrogen bonds was also measured. To determine the presence of a hydrogen bond, a geometric criterion was used, requiring that the distance between donor and acceptor was less than 0.30 nm, and that the angle formed between the acceptor, hydrogen and donor atoms was greater than 135°.

Results and Discussion

Determination of Free Energies of Transfer: Characterization of the Dataset

We aimed to use a dataset of resolved protein structures to extract suitable values of surface tension, and hence free energies of transfer, to be used in our implicit solvent simulations (approaches 2 and 3). For this purpose, we explored the PDB database, and downloaded a set of structures with more than 20 amino acids, obtained by NMR, and with a 30% level of redundancy. Only NMR structures were considered, as already done in previous work,²⁹ because they were experimentally resolved at the temperature range of interest for our next investigation (265-335 K). ccPDB 2.0⁵⁹ was used to generate the dataset (available as a Supporting Information file), which included 4120 protein structures. Of these structures, 40 were resolved at $T < 265$ K and 1 at $T > 335$ K, and were therefore discarded. Moreover, 156 structures had no temperature value indicated, while 338 listed multiple values and had therefore to be excluded from our analysis.

3585 structures remained after these preliminary considerations, and we divided them into bins based on their temperature, similarly to what was done already in previous work.²⁹

The bins considered, and the number of structures in each of them, are shown in Table 1. Figure 1 shows the distribution of the structures over the selected temperature range of 265 K-335 K.

Table 1: Number of structures in each temperature bin considered in this work, and average temperature \bar{T} of the bin.

Bin	# of structures	\bar{T} , K
265 K < T \leq 290 K	237	283.5
290 K < T \leq 297 K	469	294.0
297 K < T \leq 299 K	1853	298.0
299 K < T \leq 305 K	582	302.2
305 K < T \leq 335 K	444	311.6

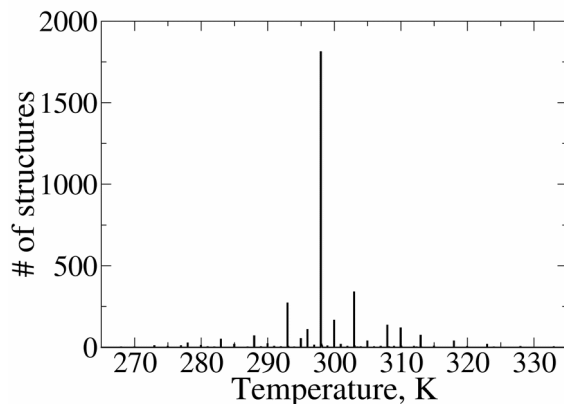


Figure 1: Distribution of protein structures over the considered temperature range (265 K-335 K).

After the subdivision into bins, we had therefore 5 sets Ω containing N_p protein structures, and whose average temperature (listed in the last column of Table 1) was computed as,

$$\bar{T} = \frac{\sum_{k=1}^{N_p} T_K}{N_p} \quad (24)$$

Extraction of Free Energies of Transfer Values

The *SASA* of each residue within the sets Ω was computed using the algorithm by Lee and Richards,^{60,61} with a probe size equal to 1.4 Å. These values were then inserted into Equations 3 and 4 to obtain the probabilities P_{nb} of each residue type to be surface exposed. These probabilities, as computed for each temperature bin, are shown in Figures S1-S5, together with the total number of residues in each dataset.

Once the probabilities P_{nb} are known, it is then possible to compute the free energies of transfer $\Delta g_{tr}^{sc,k}$ or bb using Equations 5-14. These free energies of transfer depend on temperature, as shown in Figures 2 and 3, where the values of Δg_{tr} for the different side chains, and for the backbone (BB), are shown for the different temperature bins considered.

For all side chains, with the only exception of serine, water at ambient temperature represents the 'poorest' solvent, and both lowering or increasing the temperature facilitates their exposure to the solvent. This is particularly true for hydrophobic (Figures 2A and 3A) and aromatic (Figures 2B and 3B) side chains, while the effect of temperature is marginal for charged (Figures 2C and 3C) and polar (Figures 2D and 3D) ones. The parabolic trend observed in these figures is in line with the well known temperature dependence of protein stability, that displays a maximum (generally) around ambient temperature, with both hot and cold unfolding occurring.

Each curve shown in Figures 2 and 3 was fitted to a parabola, and the resulting equations and coefficients of determination (R^2) are shown in Tables S1-S2. Using these fitting parabolic equations, it is now possible to compute $\Delta g_{tr}^{sc,k}$ or $bb(T)$ at any possible value of T , and the $G^{tr}(T)$ term (Eq. 15) can hence be calculated and implemented in an implicit solvent simulation. Plumed 2.4.7 was used for this purpose, and the fast algorithm by Hasel et al.⁶² was employed to compute the fractional solvent accessibilities $\alpha^{sc,k}$ or bb .

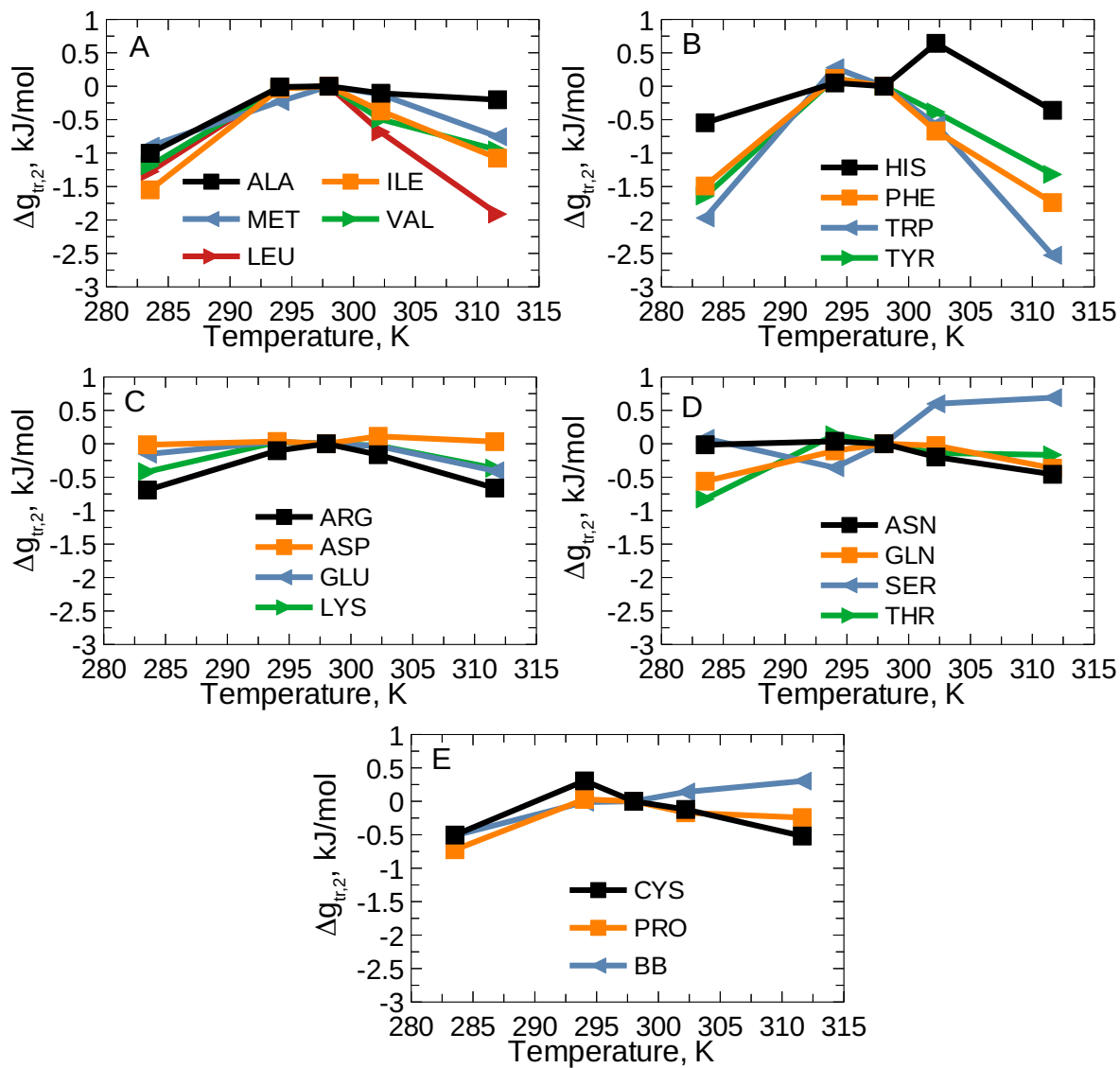


Figure 2: Values of $\Delta g_{tr,2}^{sc,k}$ or bb for the different side chains, and for the backbone (BB), for the case of approach 2. The average temperature \bar{T} for the different bins considered in this work (Table 1) is shown on the x-axis.

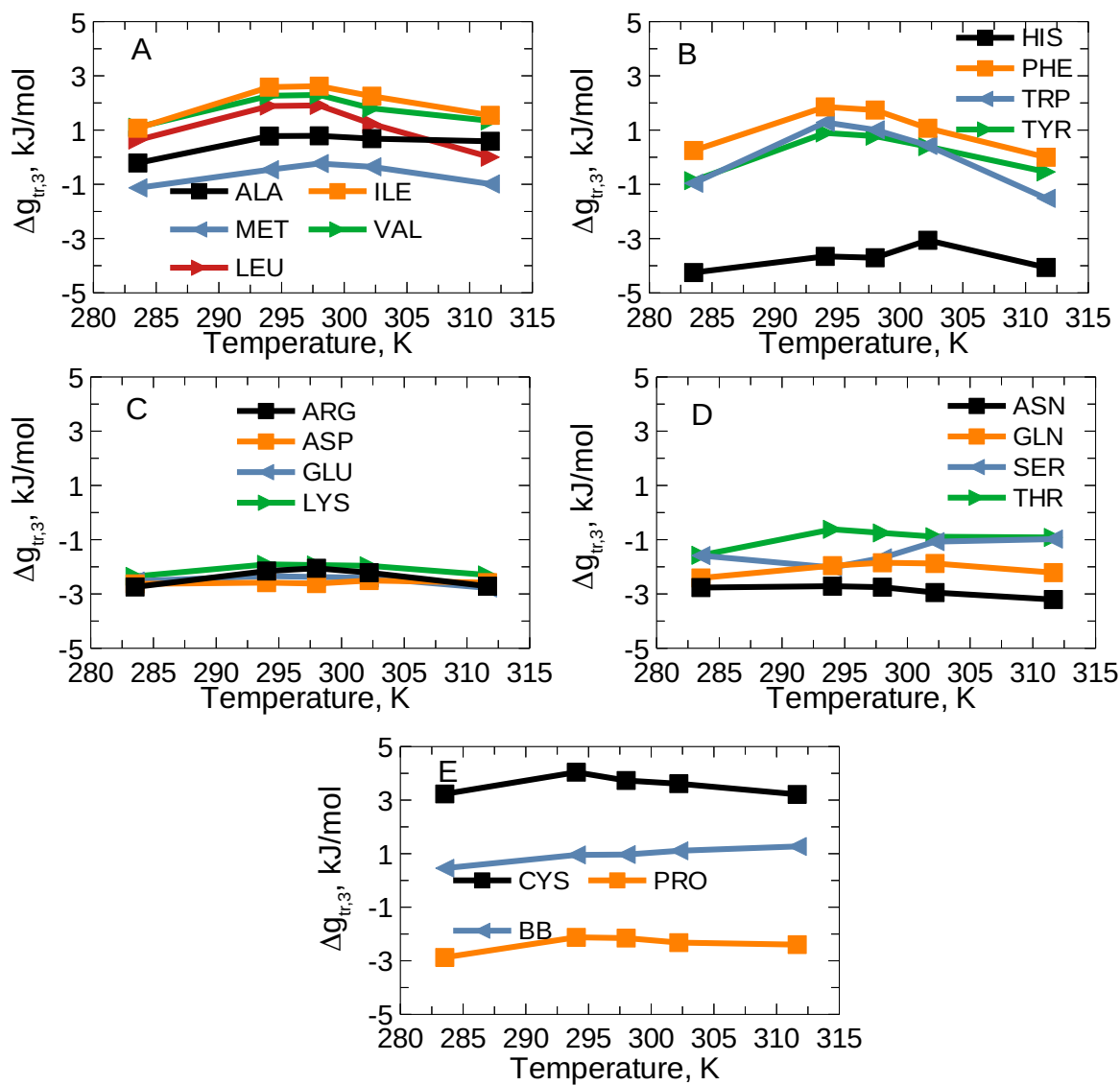


Figure 3: Values of $\Delta g_{tr,3}^{sc,k}$ or bb for the different side chains, and for the backbone (BB), for the case of approach 3. The average temperature \bar{T} for the different bins considered in this work (Table 1) is shown on the x-axis.

The Cold Denaturation of Trp-cage can be Recovered in Implicit Solvent Simulations

After having extracted the free energy of transfer values, the three approaches described in the Theoretical Background section were tested on Trp-cage as model protein.

The protein was simulated at different temperature values, and the evolution of the radius of gyration R_g , end-to-end distance R_{ee} , α -helix content, solvent accessible surface area SASA, number of internal hydrogen-bonds and folded fraction over temperature are shown in Figure 4. The end-to-end distance R_{ee} was computed between the centers of mass of the first and last residues of Trp-cage. The distributions of the α -helix content and end-to-end distance as function of the radius of gyration at different temperature values are shown in Figures S6-S8 for approaches 1, 2 and 3. Figures S6-S8 also display the most sampled protein conformations in each condition, as obtained using the Daura algorithm.⁵⁷

In the case of approach 1 (black line in Figure 4), in which no temperature-dependence was introduced, no cold unfolding was observed. Trp-cage was very stable at low temperatures, displaying a high α -helix content ($\alpha = 5.94 \pm 0.14$ at 190 K), and a compact conformation ($R_g = 0.726 \pm 0.004$ nm, $SASA = 18.86 \pm 0.16$ nm² at 190 K). The most sampled protein structure at 190 K (Figure S6) was completely folded, and was used as reference to compute the folded fraction in Figure 4F. More specifically, Trp-cage was considered to be folded when the N-C $_{\alpha}$ -C atoms RMSD compared to the reference structure was < 0.35 nm. Although hot unfolding is predicted by approach 1, the protein conformations obtained in this condition are not very expanded. The maximum radius of gyration and end-to-end distance observed at 400 K are only 0.915 ± 0.007 nm and 2.09 ± 0.04 nm, respectively. As evident from the cartoon structures shown in Figure S6, Trp-cage mostly conserves its compact structure even at 400 K when approach 1 is used.

The situation changes remarkably when moving to approach 2 (red line in Figure 4). While substantial overlap with approach 1 is observed between 270 K and 320 K, the behavior changes both at high and low temperature values. Approach 2 predicts considerably more

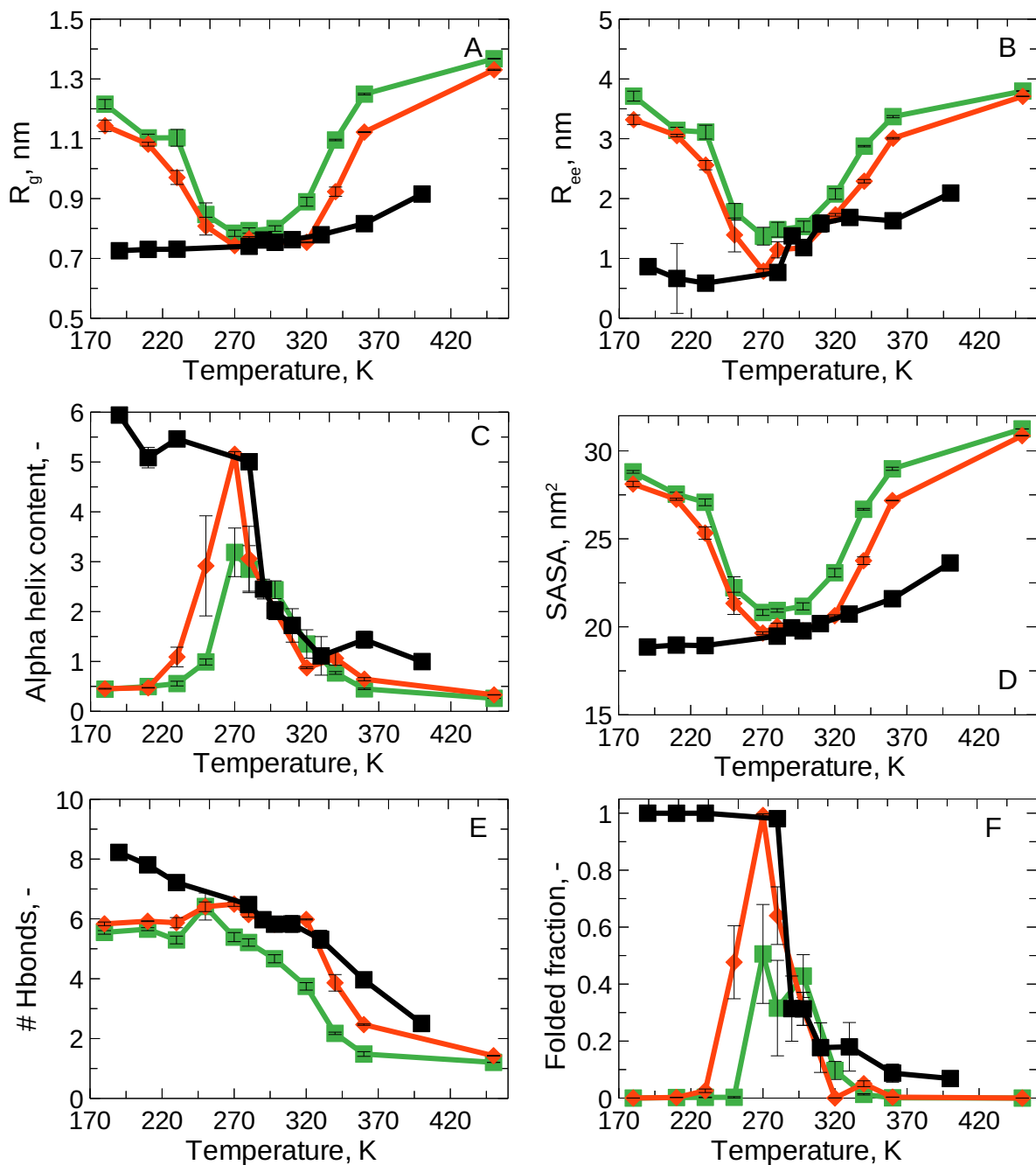


Figure 4: Evolution of radius of gyration R_g , end-to-end distance R_{ee} , α -helix content, solvent accessible surface area SASA, number of internal hydrogen-bonds and folded fraction over temperature. Error bars show one standard deviation as calculated from block averaging. Black line: approach 1, red line: approach 2, green line: approach 3.

expanded hot-unfolded structures ($R_g = 1.122 \pm 0.001$ nm, $R_{ee} = 3.01 \pm 0.01$ nm and $SASA = 27.18 \pm 0.01$ nm² at 360 K). This is also evident from the extremely extended structures sampled at 360 K and shown in Figure S7. What is, nevertheless, most remarkable is the ability of approach 2 to predict also cold unfolding. Trp-cage is completely unfolded at 210 K according to approach 2, in line with previous studies that located cold unfolding of this protein at 224 K¹⁰ or 231 K.¹³

At this extreme temperature (210 K), the cold unfolded structures are quite extended ($R_g = 1.082 \pm 0.006$ nm and $SASA = 27.25 \pm 0.06$ nm² at 210 K), and characterized by a dramatically reduced α -helix content (0.472 ± 0.005). Approach 2 predicts cold denaturation to occur in the range 250-230 K, as can be observed looking at Figure 4. Indeed, most properties show a sudden change in this temperature range, and then tend to level out at lower temperatures. Hot denaturation occurs in a wider range of temperatures compared to cold unfolding, according to approach 2. In this case, the radius of gyration, end-to-end distance, SASA and internal hydrogen-bonding network show an important variation in the range 320-360 K, and then the computed properties level out between 360 and 450 K (Figure 4). A comparison of Trp-cage structures during cold (250-230 K) and hot (320-360 K) denaturation is given in Figure 5. The protein representations indicate that hot unfolding results in a sudden and complete loss of secondary structure. The α -helix content of Trp-cage is almost completely lost already at 320 K. In contrast, the most stable protein conformation at 250 K (37.1 % probability) preserves most of the native secondary structure content, and the helix formed by residues Y3-D9 is still partially present at 230 K.

Even at the extreme temperature values simulated in this work, we still observe marked differences between hot and cold-unfolded conformations. For instance, the internal hydrogen bonding network (Figure 4E), mostly disrupted at high temperature, is instead preserved during cold unfolding (the number of internal hydrogen bonds is 5.83 ± 0.06 at 180 K, but only 1.43 ± 0.01 at 450 K). This represents a substantial difference between hot and cold-unfolded structures in our simulations, as further evidenced by Figures 6A,B. This substantial

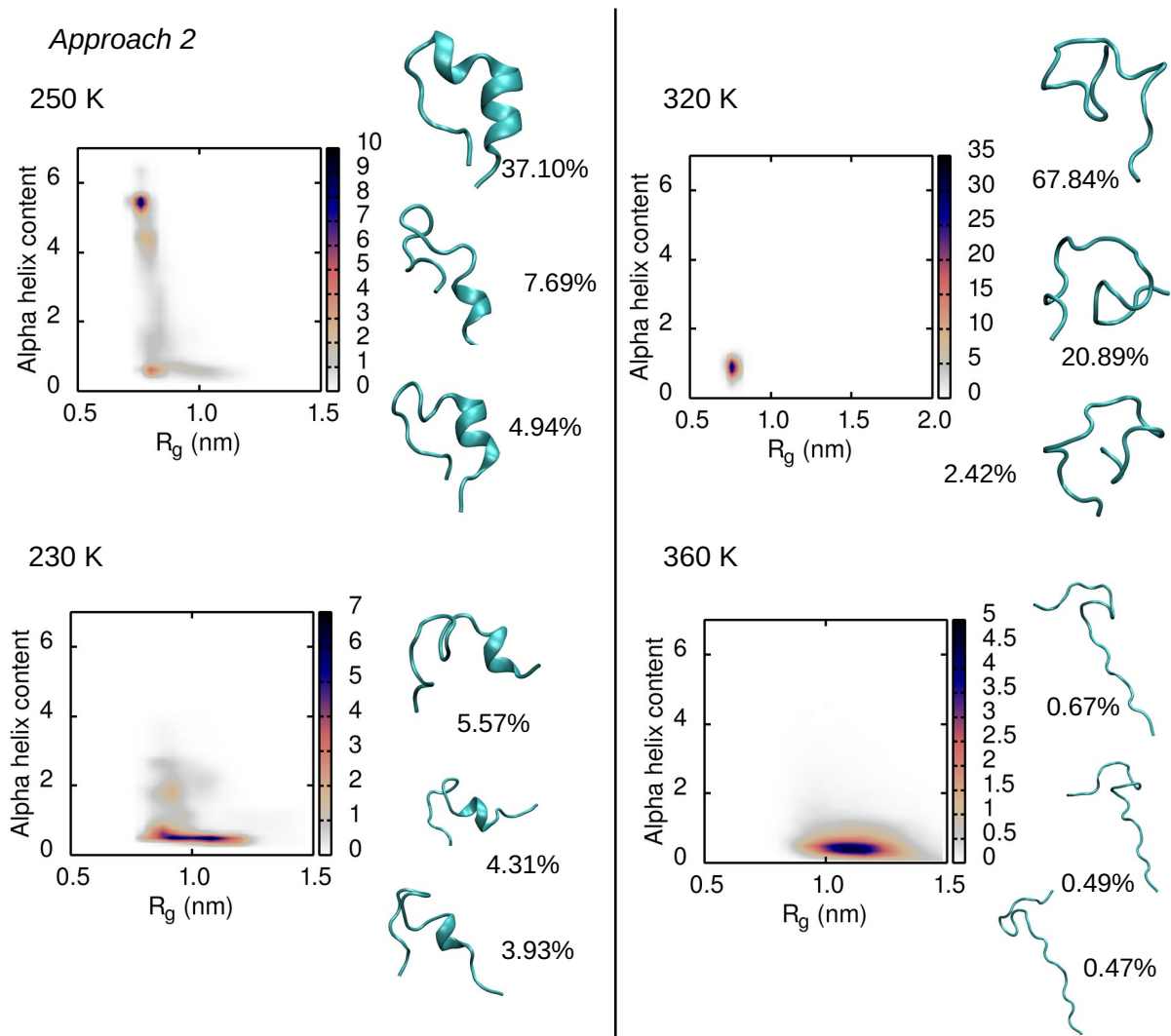


Figure 5: Comparison between the simulation results for approach 2 at the onset of cold (250 K - 230 K) and hot (320 K - 360 K) unfolding. The distribution of α -helix content as function of the radius of gyration R_g , and the most probable protein conformations, with corresponding probabilities, are shown.

preservation of the internal hydrogen-bonding network during cold unfolding observed in our simulations is also in line with previous results obtained for Trp-cage in explicit solvent.¹³

While Trp-cage extends almost completely at both low and high temperatures, a turn structure is nevertheless formed between residues G10-G15. This structure is stabilized by a significant (≈ 5) number of hydrogen bonds at 180 K, while only one bond is observed at 450 K (Figure 6A). This remarkable difference between hot and cold-unfolded configurations is also clearly illustrated by the time evolution of internal hydrogen bonds in Figure 6B.

We also computed the distance between the centers of masses of residues W6 and S14 (Figure 6C), which are located in Trp-cage’s hydrophobic core or on the C-terminus, respectively. These two residues are close in the folded structure, because of Trp-cage’s shape. An increase in the W6-S14 distance means that the hydrophobic core of Trp-cage becomes exposed to the solvent. In approach 1, this occurs only during hot unfolding ($T > 320$ K), while according to approaches 2 and 3 the W6-S14 distance increases also at low temperature. In particular, the largest increase is again observed between 250 and 230 K (cold-unfolding) or between 320 and 360 K (hot-unfolding) according to approach 2. The increase in the W6-S14 distance during cold unfolding is in line with previous explicit solvent simulations of Trp-cage.^{10,13}

The parabolic profile of protein stability is also predicted by approach 3 (green line in Figure 4). It is interesting and important to note that the completely new set of surface tension values γ_3 used in approach 3 can fairly well predict the folding behavior of Trp-cage, although the protein is slightly less stable in these conditions compared to approaches 1 and 2.

The values of $\gamma_3^{sc,i}$ or bb at $T_0 = 298$ K are shown in Table S3. The values are positive (i.e., unfavorable exposure to the solvent) for apolar and aromatic side chains and for the backbone, and negative for charged and polar side chains. The value for the backbone ($5.4 \text{ cal mol}^{-1} \text{ \AA}^{-2}$), in particular, is also very close to the default value of $\gamma_0 = 5 \text{ cal mol}^{-1} \text{ \AA}^{-2}$ used in AMBER. This explains the substantial overlap between approach 3 and approach 1

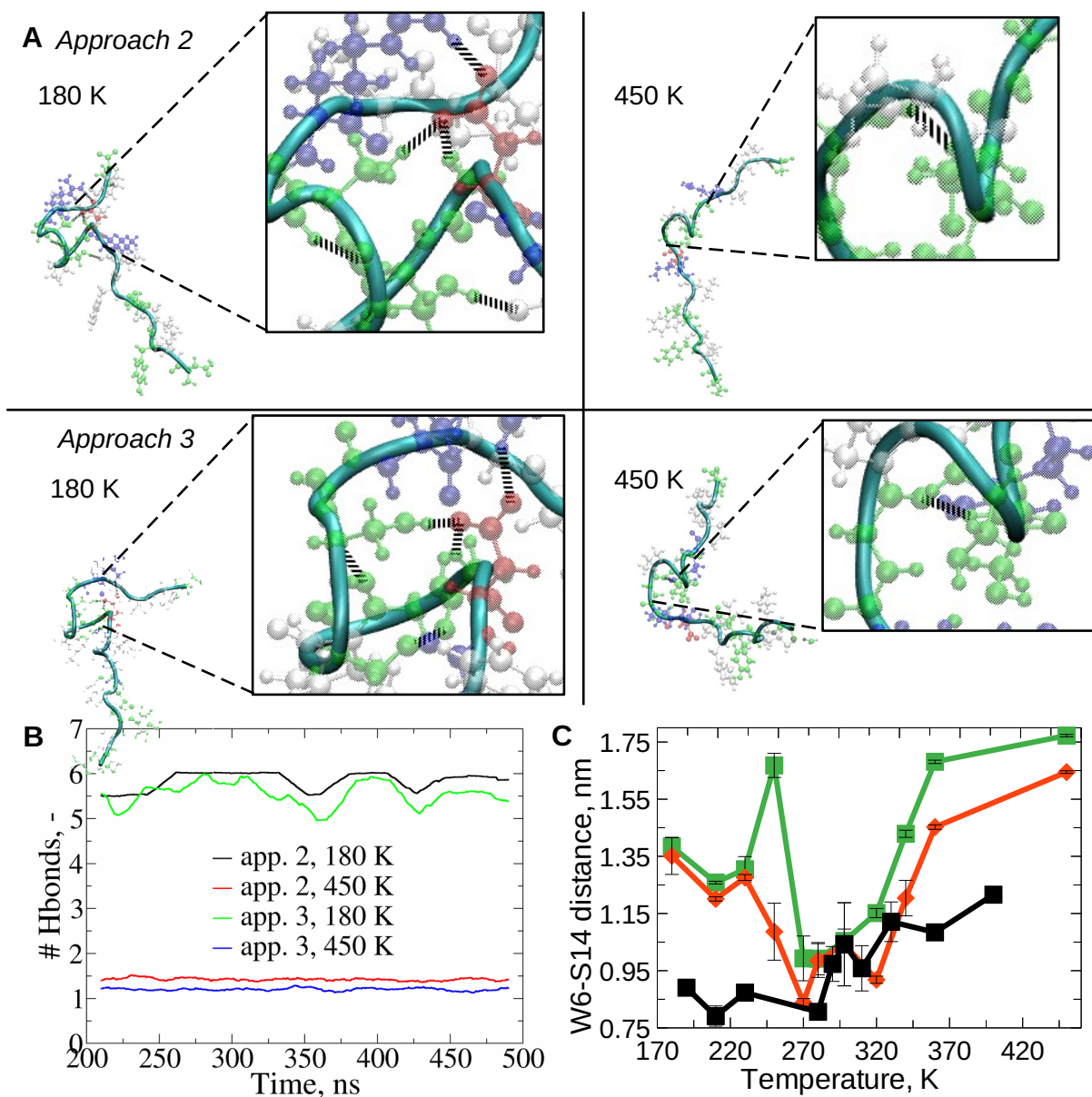


Figure 6: (A) Comparison of cold (180 K) and hot (450 K) unfolded structures, as obtained by approaches 2 or 3. The turn formed by residues G10-G15 is shown in the insets, and the hydrogen bonds that stabilize this turn are evidenced in black. The images were obtained using Visual Molecular Dynamics (VMD).⁵⁸ Polar uncharged residues are in green, nonpolar ones in white, positively charged in blue and negatively charged in red. (B) Evolution of the number of internal hydrogen bonds over time (last 300 ns) as predicted by approaches (app.) 2 or 3, at 180 and 450 K. (C) Distance between residues W6 and S14. Black line: approach 1, red line: approach 2, green line: approach 3.

around ambient temperature (black and green lines in Figure 4).

The unfolded structures obtained at 210 K ($R_g = 1.10 \pm 0.01$ nm, $R_{ee} = 3.14 \pm 0.05$ nm) in the case of approach 3 are slightly less extended than the hot-unfolded ones at 360 K ($R_g = 1.249 \pm 0.003$ nm, $R_{ee} = 3.37 \pm 0.02$ nm), but in both cases the secondary structure is almost completely lost (Figure S8). The substantial preservation of internal hydrogen bonds in cold-unfolded structures is confirmed in the case of approach 3 (the number of internal hydrogen bonds is 5.55 ± 0.06 at 180 K, compared to 1.21 ± 0.002 at 450 K, see Figure 6).

The well-known parabolic profile of protein stability can hence be recovered also in implicit solvent simulations, and the reduced protein stability at low temperatures can be predicted. In the modelling approach here proposed, this decrease in stability is driven by the increased compatibility of water with the peptide backbone and most sidechains, especially apolar and aromatic ones, at low temperature. We emphasize that the advantages of an implicit solvent model over an explicit solvent model are not only associated with the reduced computational cost and faster equilibration at low temperatures. Rather, the use of implicit solvation models makes it possible to better understand the energetic contributions and essential physical requirements for cold denaturation.

Conclusions

In this work we have developed a new approach to deal with the process of cold unfolding in implicit solvent simulations. Compared to previous studies that added temperature-dependent potentials and recovered cold-denaturation in absence of explicit water,^{15,17} our approach has the advantage of being fully compatible with an atomistic representation of the protein, and does not make use of protein-dependent parameters. On the contrary, the free energy of transfer values involved in the temperature-dependent term are obtained by mining a large set of PDB files resolved by NMR. These free energy contributions represent the transfer between water at two different temperature values. The two transfer free energy

sets (approaches 2 and 3) obtained tell us that water at ambient temperature is the 'poorest' solvent for the peptide backbone and most side chains, and that an increase or decrease in temperature results in improved solvation, especially for apolar groups. When deriving the first set (approach 2), we did not alter the description of protein hydration at 298 K compared to previously implemented algorithms. The second set (approach 3) represents, instead, a completely new description of non-polar hydration, even at ambient temperature.

We have tested the two sets using Trp-cage as model protein. While the previously implemented algorithm (approach 1) without temperature-dependent term fails in predicting cold-unfolding, both the proposed sets of transfer free energies can describe the parabolic profile of protein stability, with denaturation being observed both at high and low temperatures. Overall, approach 2 has the advantage of predicting cold unfolding, without affecting protein stability around ambient temperature, and is therefore recommended for future applications. Approach 3 results in a slightly reduced stability of Trp-cage, although the substantial overlap with approaches 1 and 2 remains remarkable, especially considering that approach 3 was developed in a completely independent manner.

In our implicit solvent simulations, cold unfolding results from the improved solvation of protein moieties, especially apolar and aromatic ones, at low temperature. Hot denaturation, mostly driven by the increased entropy of unfolded states, can also be predicted by previous implementations of implicit solvation, but the transfer free energies proposed in this work promote the sampling of significantly more extended hot-unfolded conformations. This indicates, once again, the need for a temperature-dependent term in implicit solvent simulations. Such term, as shown in the present work, not only allows for the description of cold unfolding, but also improves the modelling of hot denatured structures.

Through the addition of a temperature-dependent term, we find that hot denaturation results in a sudden and complete loss of secondary structure, while cold unfolding is more gradual, and a large fraction of the internal hydrogen-bonding network is retained even at very low temperatures. This represents a substantial difference with hot-unfolded conforma-

tions of Trp-cage, where the native hydrogen bonds are mostly disrupted.

Supporting Information

Probability of residues to be surface exposed and total number of residues considered in different temperature ranges, fitting equations of the temperature-dependent term for approaches 2 and 3, distribution of α -helix content and end-to-end distance as function of the radius of gyration and most probable Trp-cage conformations for approaches 1, 2 and 3, values of γ_3 at 298 K, dataset of PDB structures obtained by NMR.

Acknowledgement

The authors acknowledge support from the Center for Scientific Computing at the California Nanosystems Institute (CNSI, NSF grant CNS-1725797) for the availability of high performance computing resources and support. This work used the Extreme Science and Engineering Discovery Environment, which is supported by the National Science Foundation grant number ACI-1548562 (MCA05S027). The authors acknowledge support from the NSF (MCB-1716956) and the NIH (R01-GM118560-01A).

References

- (1) Privalov, P. L. Cold Denaturation of Proteins. *Crit. Rev. Biochem. Mol. Biol.* **1990**, *25*, 281–305.
- (2) Franks, F. Protein Destabilization at Low Temperatures. *Adv. Protein Chem.* **1995**, *46*, 105–139.
- (3) Graziano, G. On the mechanism of cold denaturation. *Phys. Chem. Chem. Phys.* **2014**, *16*, 21755–21767.

- (4) Dill, K. A.; Alonso, D. O. V.; Hutchinson, K. Thermal Stabilities of Globular Proteins. *Biochemistry* **1989**, *28*, 5439–5449.
- (5) Graziano, G.; Catanzano, F.; Riccio, A.; Barone, G. A Reassessment of the Molecular Origin of Cold Denaturation. *J. Biochem.* **1997**, *122*, 395–401.
- (6) Lopez, C. F.; Darst, R. K.; Rossky, P. J. Mechanistic Elements of Protein Cold Denaturation. *J. Phys. Chem. B* **2008**, *112*, 5961–5967.
- (7) Davidovic, M.; Mattea, C.; Qvist, J.; Halle, B. Protein Cold Denaturation as Seen From the Solvent. *J. Am. Chem. Soc.* **2009**, *131*, 1025–1036.
- (8) Babu, C. R.; Hilser, V. J.; Wand, A. J. Direct Access to the Cooperative Substructure of Proteins and the Protein Ensemble via Cold Denaturation. *Nat. Struct. Mol. Biol.* **2004**, *11*, 352–357.
- (9) Matysiak, S.; Debenedetti, P. G.; Rossky, P. J. Role of Hydrophobic Hydration in Protein Stability: A 3D Water-Explicit Protein Model Exhibiting Cold and Heat Denaturation. *J. Phys. Chem. B* **2012**, *116*, 8095–8104.
- (10) Kozuch, D. J.; Stillinger, F. H.; Debenedetti, P. G. Low temperature protein refolding suggested by molecular simulation. *J. Chem. Phys.* **2019**, *151*, 185101.
- (11) Patel, B. A.; Debenedetti, P. G.; Stillinger, F. H.; Rossky, P. J. The Effect of Sequence on the Conformational Stability of a Model Heteropolymer in Explicit Water. *J. Chem. Phys.* **2008**, *128*, 175102.
- (12) Dias, C. L.; Ala-Nissila, T.; Karttunen, M.; Vattulainen, I.; Grant, M. Microscopic Mechanism for Cold Denaturation. *Phys. Rev. Lett.* **2008**, *100*, 118101.
- (13) Kim, S. B.; Palmer, J. C.; Debenedetti, P. G. Computational investigation of cold denaturation in the Trp-cage miniprotein. *Proc. Natl. Acad. Sci.* **2016**, *113*, 8991–8996.

- (14) Arsiccio, A.; McCarty, J.; Pisano, R.; Shea, J.-E. Heightened Cold-Denaturation of Proteins at the Ice–Water Interface. *J. Am. Chem. Soc.* **2020**, *142*, 5722–5730.
- (15) Sirovetz, B. J.; Schafer, N. P.; Wolynes, P. G. Water Mediated Interactions and the Protein Folding Phase Diagram in the Temperature–Pressure Plane. *J. Phys. Chem. B* **2015**, *119*, 11416–11427.
- (16) Davtyan, A.; Schafer, N. P.; Zheng, W.; Clementi, C.; Wolynes, P. G.; Papoian, G. A. AWSEM-MD: Protein Structure Prediction Using Coarse-Grained Physical Potentials and Bioinformatically Based Local Structure Biasing. *J. Phys. Chem. B* **2012**, *116*, 8494–8503.
- (17) van Dijk, E.; Varilly, P.; Knowles, T. P. J.; Frenkel, D.; Abeln, S. Consistent Treatment of Hydrophobicity in Protein Lattice Models Accounts for Cold Denaturation. *Phys. Rev. Lett.* **2016**, *116*, 078101.
- (18) Wagoner, J. A.; Baker, N. A. Assessing implicit models for nonpolar mean solvation forces: The importance of dispersion and volume terms. *Proc. Natl. Acad. Sci.* **2006**, *103*, 8331–8336.
- (19) Gallicchio, E.; Levy, R. M. AGBNP: An analytic implicit solvent model suitable for molecular dynamics simulations and high-resolution modeling. *J. Comput. Chem.* **2004**, *25*, 479–499.
- (20) Lebowitz, J. L.; Helfand, E.; Praestgaard, E. Scaled Particle Theory of Fluid Mixtures. *J. Chem. Phys.* **1965**, *43*, 774–779.
- (21) Reiss, H.; Frisch, H. L.; Lebowitz, J. L. Statistical Mechanics of Rigid Spheres. *J. Chem. Phys.* **1959**, *31*, 369–380.
- (22) Eisenberg, D.; McLachlan, A. D. Solvation energy in protein folding and binding. *Nature* **1986**, *319*, 199–203.

- (23) Wesson, L.; Eisenberg, D. Atomic solvation parameters applied to molecular dynamics of proteins in solution. *Protein Sci.* **1992**, *1*, 227–235.
- (24) Fraternali, F.; van Gunsteren, W. An Efficient Mean Solvation Force Model for Use in Molecular Dynamics Simulations of Proteins in Aqueous Solution. *J. Mol. Biol.* **1996**, *256*, 939 – 948.
- (25) Ferrara, P.; Apostolakis, J.; Caflisch, A. Evaluation of a fast implicit solvent model for molecular dynamics simulations. *Proteins* **2002**, *46*, 24–33.
- (26) Elcock, A. H.; McCammon, J. A. Continuum Solvation Model for Studying Protein Hydration Thermodynamics at High Temperatures. *J. Phys. Chem. B* **1997**, *101*, 9624–9634.
- (27) Kónig, G.; Bruckner, S.; Boresch, S. Absolute Hydration Free Energies of Blocked Amino Acids: Implications for Protein Solvation and Stability. *Biophys. J.* **2013**, *104*, 453–462.
- (28) Chang, J.; Lenhoff, A. M.; Sandler, S. I. Solvation Free Energy of Amino Acids and Side-Chain Analogues. *J. Phys. Chem. B* **2007**, *111*, 2098–2106.
- (29) van Dijk, E.; Hoogeveen, A.; Abeln, S. The Hydrophobic Temperature Dependence of Amino Acids Directly Calculated from Protein Structures. *PLOS Comput. Biol.* **2015**, 1–17.
- (30) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. Designing a 20-residue protein. *Nat. Struct. Biol.* **2002**, *9*, 425–430.
- (31) Qiu, L.; Pabit, S. A.; Roitberg, A. E.; Hagen, S. J. Smaller and Faster: The 20-Residue Trp-Cage Protein Folds in 4 μ s. *J. Am. Chem. Soc.* **2002**, *124*, 12952–12953.
- (32) Snow, C. D.; Zagrovic, B.; Pande, V. S. The Trp Cage: Folding Kinetics and Unfolded

- State Topology via Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2002**, *124*, 14548–14549.
- (33) Pitera, J. W.; Swope, W. Understanding folding and design: Replica-exchange simulations of “Trp-cage” miniproteins. *Proc. Natl. Acad. Sci.* **2003**, *100*, 7587–7592.
- (34) Levine, Z. A.; Larini, L.; LaPointe, N. E.; Feinstein, S. C.; Shea, J.-E. Regulation and aggregation of intrinsically disordered peptides. *Proc. Natl. Acad. Sci.* **2015**, *112*, 2758–2763.
- (35) Day, R.; Paschek, D.; Garcia, A. E. Microsecond simulations of the folding/unfolding thermodynamics of the Trp-cage miniprotein. *Proteins* **2010**, *78*, 1889–1899.
- (36) Ganguly, P.; Shea, J.-E. Distinct and Nonadditive Effects of Urea and Guanidinium Chloride on Peptide Solvation. *J. Phys. Chem. Lett.* **2019**, *10*, 7406–7413.
- (37) Zhou, R. Trp-cage: Folding free energy landscape in explicit water. *Proc. Natl. Acad. Sci.* **2003**, *100*, 13280–13285.
- (38) Marinelli, F.; Pietrucci, F.; Laio, A.; Piana, S. A Kinetic Model of Trp-Cage Folding from Multiple Biased Molecular Dynamics Simulations. *PLOS Comput. Biol.* **2009**, *5*, 1–18.
- (39) Juraszek, J.; Bolhuis, P. G. Sampling the multiple folding mechanisms of Trp-cage in explicit solvent. *Proc. Natl. Acad. Sci.* **2006**, *103*, 15859–15864.
- (40) Peter, E. K.; Shea, J.-E.; Schug, A. CORE-MD, a path correlated molecular dynamics simulation method. *J. Chem. Phys.* **2020**, *153*, 084114.
- (41) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. *J. Phys. Chem.* **1994**, *98*, 1978–1988.
- (42) Rose, G. D.; Fleming, P. J.; Banavar, J. R.; Maritan, A. A backbone-based theory of protein folding. *Proc. Natl. Acad. Sci.* **2006**, *103*, 16623–16633.

- (43) Liu, Y.; Bolen, D. W. The peptide backbone plays a dominant role in protein stabilization by naturally occurring osmolytes. *Biochemistry* **1995**, *34*, 12884–12891.
- (44) Creamer, T. P.; Srinivasan, R.; Rose, G. D. Modeling unfolded states of proteins and peptides. II. Backbone solvent accessibility. *Biochemistry* **1997**, *36*, 2832–2835.
- (45) Creamer, T. P.; Srinivasan, R.; Rose, G. D. Modeling unfolded states of peptides and proteins. *Biochemistry* **1995**, *34*, 16245–16250.
- (46) Auton, M.; Bolen, D. W. Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proc. Natl. Acad. Sci.* **2005**, *102*, 15065–15068.
- (47) Auton, M.; Holthausen, L. M. F.; Bolen, D. W. Anatomy of energetic changes accompanying urea-induced protein denaturation. *Proc. Natl. Acad. Sci.* **2007**, *104*, 15317–15322.
- (48) Auton, M.; Bolen, D. W.; Rösgen, J. Structural thermodynamics of protein preferential solvation: Osmolyte solvation of proteins, aminoacids, and peptides. *Proteins* **2008**, *73*, 802–813.
- (49) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function, and Bioinformatics* **2010**, *78*, 1950–1958.
- (50) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.* **1995**, *91*, 1 – 41.
- (51) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604 – 613.

- (52) Onufriev, A.; Bashford, D.; Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* **2004**, *55*, 383–394.
- (53) Bradley, D. J.; Pitzer, K. S. Thermodynamics of electrolytes. 12. Dielectric properties of water and Debye-Hueckel parameters to 350.degree.C and 1 kbar. *J. Phys. Chem.* **1979**, *83*, 1599–1603.
- (54) Hodge, I. M.; Angell, C. A. The relative permittivity of supercooled water. *J. Chem. Phys.* **1978**, *68*, 1363–1368.
- (55) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327 – 341.
- (56) Pietrucci, F.; Laio, A. A Collective Variable for the Efficient Exploration of Protein Beta-Sheet Structures: Application to SH3 and GB1. *J. Chem. Theory .Comput.* **2009**, *5*, 2197–2201.
- (57) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide Folding: When Simulation Meets Experiment. *Angew. Chem. Int. Ed.* **1999**, *38*, 236–240.
- (58) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33 – 38.
- (59) Agrawal, P.; Patiyal, S.; Kumar, R.; Kumar, V.; Singh, H.; Raghav, P. K.; Raghava, G. P. S. ccPDB 2.0: an updated version of datasets created and compiled from Protein Data Bank. *Database* **2019**, *2019*.
- (60) Lee, B.; Richards, F. The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.* **1971**, *55*, 379 – IN4.

- (61) Richards, F. M. Areas, volumes, packing, and protein structure. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151–176.
- (62) Hasel, W.; Hendrickson, T. F.; Still, W. C. A rapid approximation to the solvent accessible surface areas of atoms. *Tetrahedron Comput. Methodol.* **1988**, *1*, 103–116.

Graphical TOC Entry

