

Supplementary Material of manuscript: “ Φ -Value analysis by molecular dynamics simulations of reversible folding”, by G. Settanni *et al.*

1 Evaluation of **cluster- P_{fold}**

The value of **cluster- P_{fold}** obtained with the procedure exposed in the Methods section does not depend strongly on the criterion used to define commitment to fold (unfold) [15], i.e., $Q > 0.85$ ($Q < 0.15$), because in the near-equilibrium trajectories the relaxation time within the folded (or unfolded) basin is two order of magnitude shorter than the average unfolding (or folding) time (data not shown). Thus, instead of searching for observables that correlate with P_{fold} , we just identify a feature that unequivocally belongs to folded (unfolded) conformations, i.e., very high (low) Q . When the system undergoes a folding (unfolding) transition, it explores rapidly, with respect to folding/unfolding times, the folded (unfolded) basin and reaches the conformations with high (low) Q . This unequivocally signals the occurrence of the transition event. Although Q can distinguish fully folded and fully denatured structures, it does not help distinguishing structures with properties intermediate between the native and denatured state. Indeed, structures with Q as large as 0.7 may have $\text{cluster-}P_{fold} = 0$ and, vice versa, Q as low as 0.3 may correspond to structures with $\text{cluster-}P_{fold} = 1$ (See Fig. 1 in Supplementary Material). The procedure to compute standard P_{fold} with multiple runs starting from the same structure is extremely CPU expensive whereas the **cluster- P_{fold}** calculations requires only a few seconds if near-equilibrium trajectories and an efficient clustering algorithm are available.

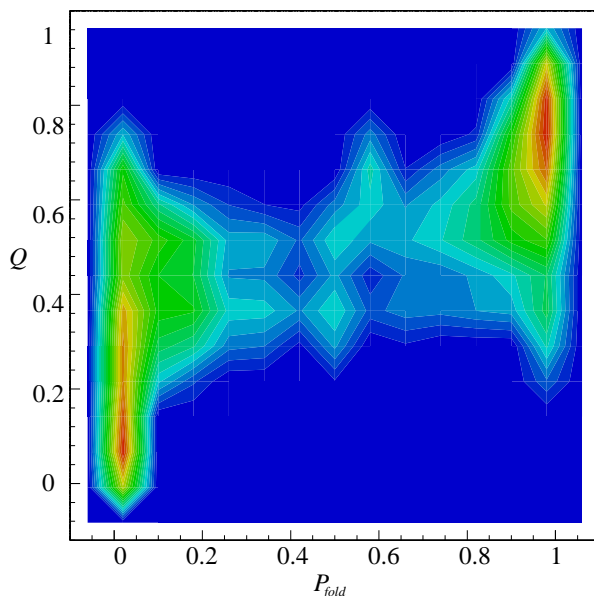


Figure 1: Distribution of **fraction of native contacts (Q)** and **cluster- P_{fold}** in the wild-type peptide simulations. The color scale ranging from red to blue corresponds to high and low density, respectively. While structures with very large Q ($Q > 0.8$) or very low Q ($Q < 0.2$) have cluster- P_{fold} close to 1 or 0, respectively, conformations with intermediate values of Q span all the allowed spectrum of cluster- P_{fold} values.

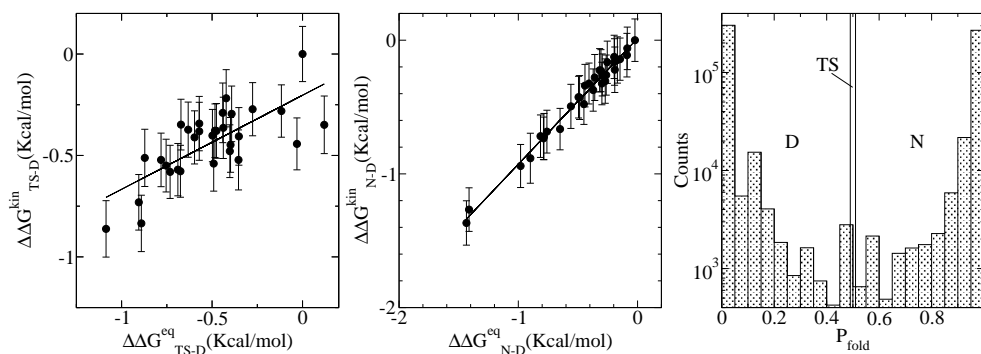


Figure 2: Left and middle graph: comparison between free energy changes calculated with the kinetic or **cluster- P_{fold}** data. The correlation coefficient is 0.73 and 0.99 for $\Delta\Delta G_{TS-D}$ and $\Delta\Delta G_{N-D}$, respectively. Right graph: distribution of cluster- P_{fold} for the wild-type simulations and identification of the native, TS, and denatured state ensemble.