Supporting Information

De Simone et al. 10.1073/pnas.1118048109

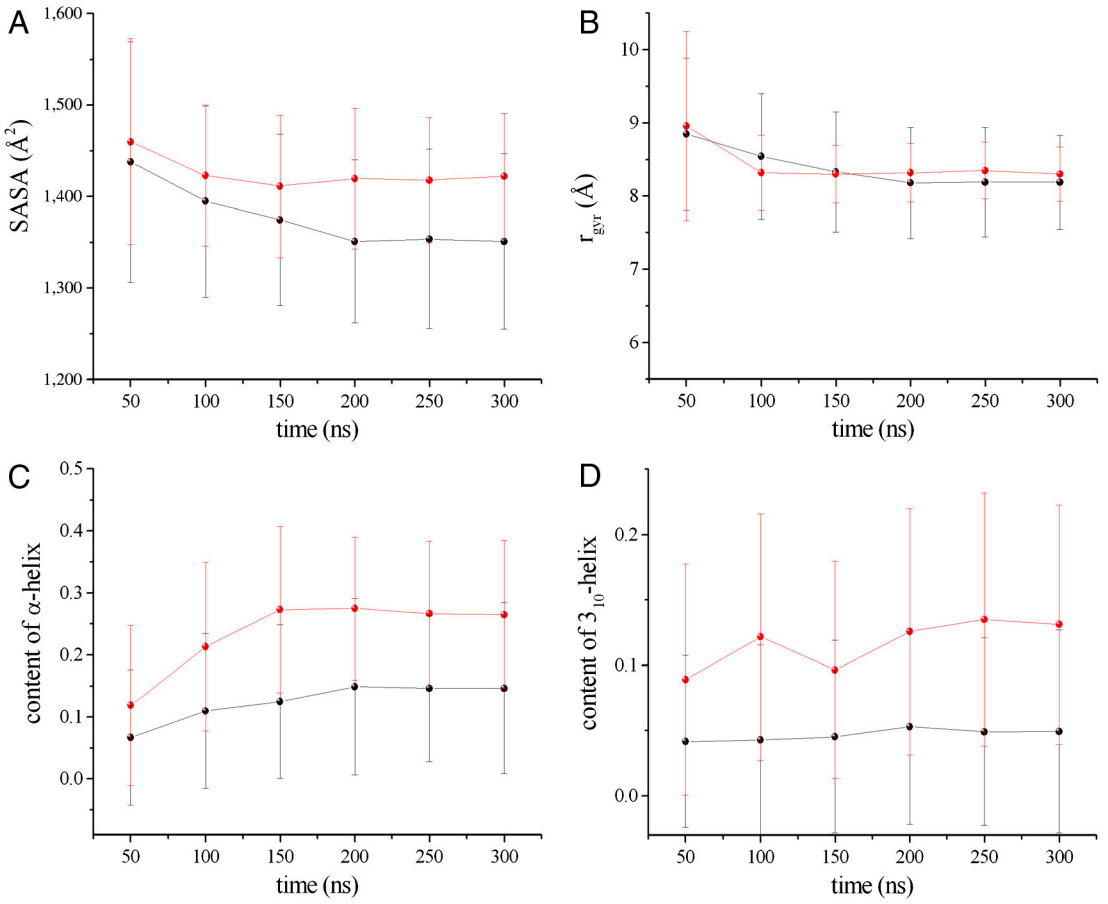

Fig. S1. Time evolution of structural parameters in the water (black) and air-water interface REMD simulations (red). The values have been averaged in segments of 50 ns. Standard deviations are reported as error bars. (*A*) Hydrophobic surface accessible area. (*B*) Radius of gyration. (*C*) Content of alpha helix. (*D*) Content of 3_{10} helix.

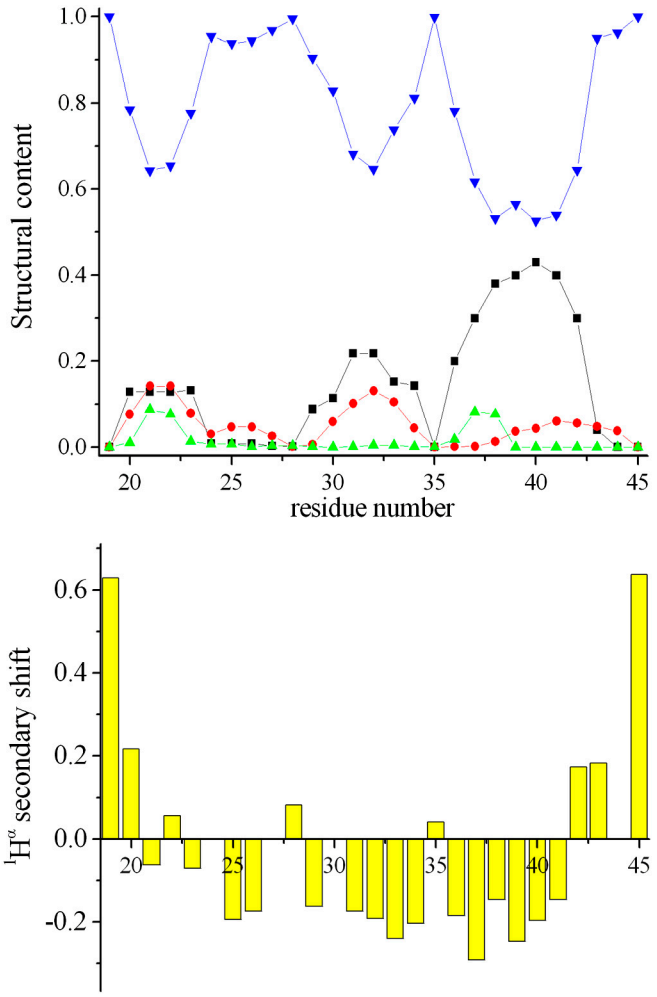

Fig. S2. Secondary structure propensity of the EAS loop in solution from REMD simulations and NMR measurements. (*Upper*) Secondary structure profiles averaged over the set of conformations from the converged part of the REMD sampling and calculated by using the DSSP program. Color codes: α-helix (black), 3_{10} helix (red), β-sheet (green) and coil (blue). (*Lower*) Secondary chemical shifts of $^1H^\alpha$ atoms reveal the presence of residual α-helical structure for regions 29 to 34 and 36 to 41 (negative secondary chemical shifts). These are calculated by subtracting the random coil reference values, calculated by means of the CamCoil method (1), from the measured chemical shifts of the full length EAS (2). As the measured chemical shifts refer to the full length protein, the termini region of the loop show positive secondary shifts, indicative of a β-sheet conformation. This finding is consistent with the full length structure in which the loop connects the strands S2 and S3 of the protein. The regions 29 to 34 and 36 to 41, however, show negative secondary shifts indicating residual α-helical structure.

1. De Simone A, Cavalli A, Hsu ST, Vranken W, Vendruscolo M (2009) Accurate random coil chemical shifts from an analysis of loop regions in native states of proteins. *J Am Chem Soc* 131:16332–16333.
2. Kwan AH, et al. (2006) Structural basis for rodlet assembly in fungal hydrophobins. *Proc Natl Acad Sci USA* 103:3621–3626.

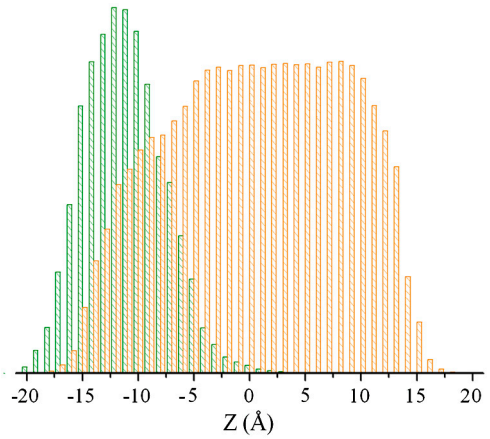

Fig. S3. Distribution of the loop and water atoms along the z direction running perpendicular to the water layer. For each frame the atomic coordinates have been translated in order to align the center of masses of the water layer at a z value of 0. Green and orange histograms show the distribution along the Z direction (i.e., orthogonal to the water layer) of the peptide and water atoms, respectively.

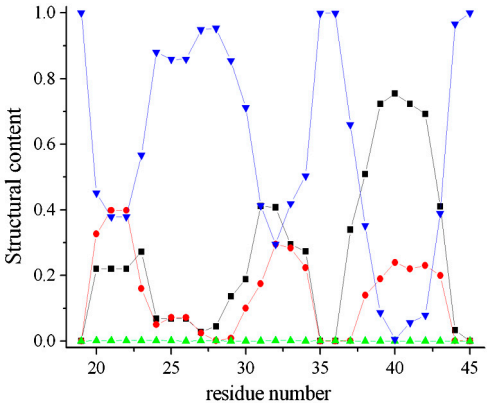

Fig. S4. Secondary structure propensity of the $EAS_{19–45}$ loop at the air–water interface from the REMD simulations. Color codes: α-helix (black), 3_{10} helix (red), β-sheet (green) and coil (blue).

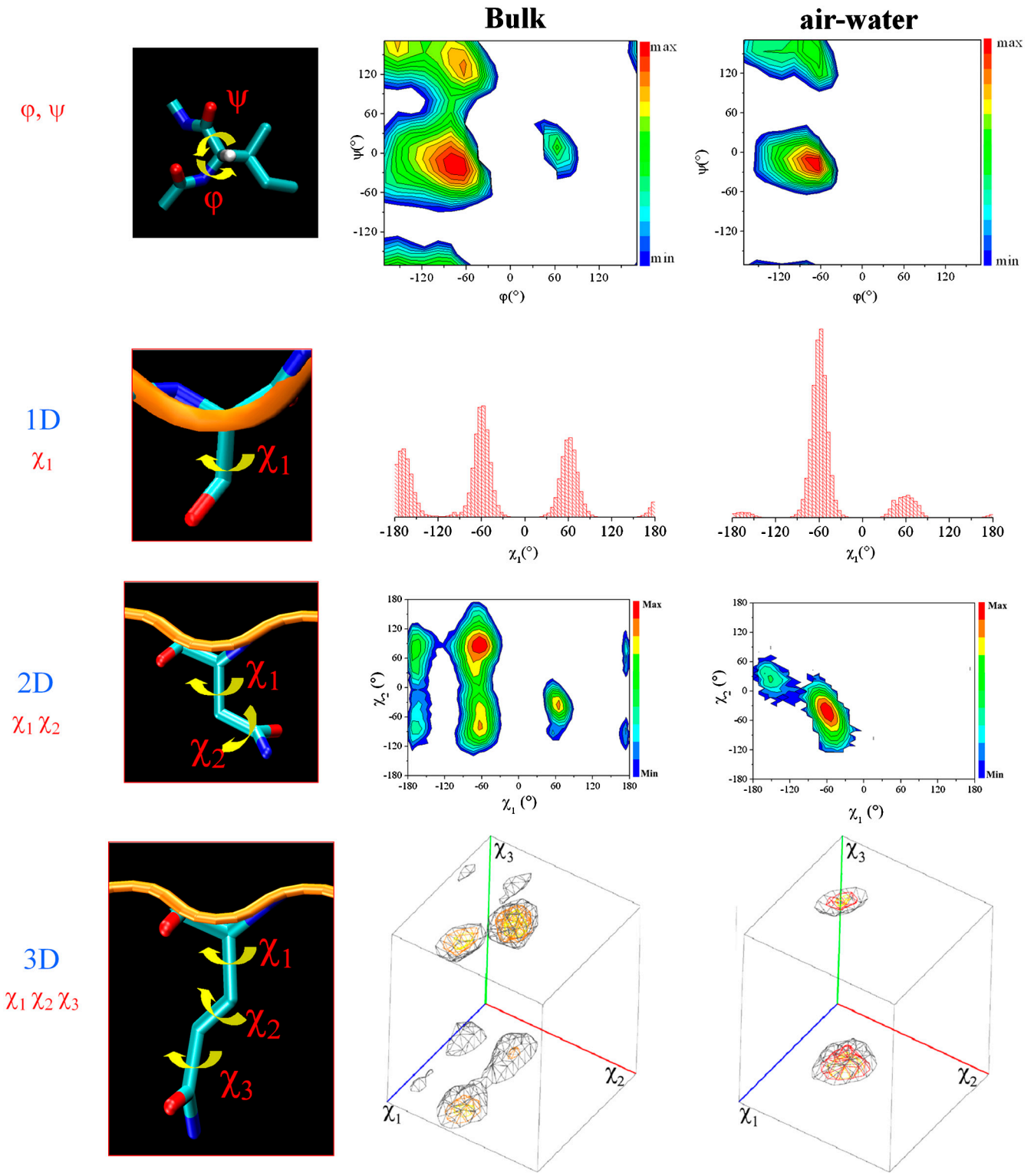

Fig. S5. Conformational entropy. We calculated the conformational entropy of the loop in a given sampling from the distributions of the ensembles in 1D, 2D, and 3D phase spaces of dihedral angles. For the main chain, the Ramachandran space is employed (ϕ and ψ angles). For the side chains, 1D (i.e., Ser), 2D (i.e., Asn), and 3D (i.e., Gln) spaces of the χ angles are considered. The total entropy content of a residue of the polypeptide chain is considered as $S_{tot} = S_{mainchain} + S_{sidechain}$. For Gly and Ala residues, only the main chain term is considered. The general formula for calculating the entropy from a distribution is $S = -R\sum_{i,j\ldots} P_{i,j\ldots} \ln P_{i,j\ldots}$.

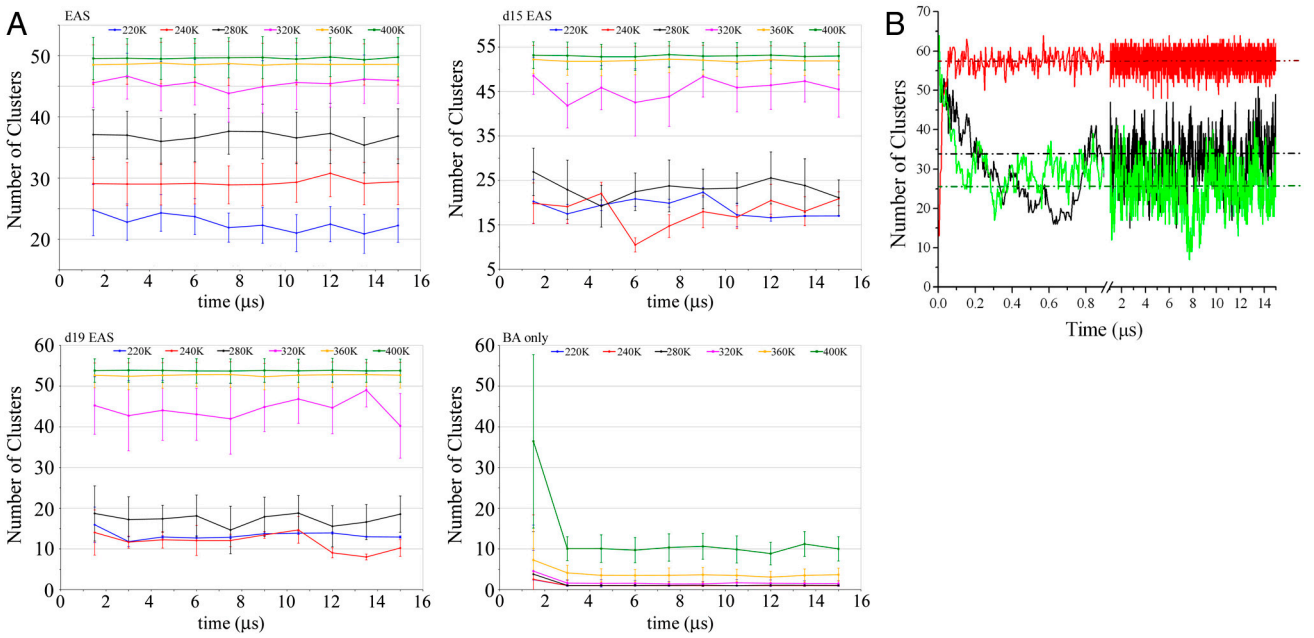

Fig. S6. (*A*) Time evolution of the clustering index in the coarse-grained simulations. (*B*) Control simulations have been performed for all the systems (EAS WT, Δ15 EAS, Δ19 EAS, BA) by starting from configurations that were equilibrated at different temperatures. This control analysis showed that 5 μs is a sufficient equilibration time. Three control simulations are shown: EAS WT at 400 K starting from a configuration equilibrated at 200 K (red lines). Δ15 EAS at 300 K starting from a configuration equilibrated at 400 K (black lines). Δ19 EAS at 300 K starting from a configuration equilibrated at 400 K (green lines).

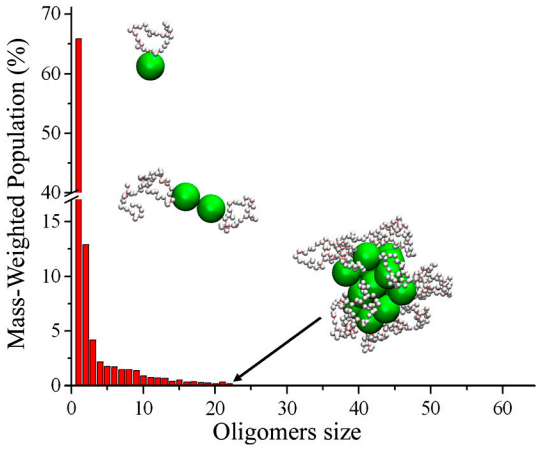

Fig. S7. Mass weighted distribution of WT EAS oligomers in the CG simulations at 300 K. The plot is normalised by scaling the population of the oligomers with the number of monomers from which the oligomers are formed.