A grid-based backbone correction to the ff12SB protein force field for implicit-solvent simulations

*Alberto Perez[1], Justin L. MacCallum[2], Emiliano Brini[1], Carlos Simmerling[1,3] and Ken A. Dill,[1,4]**

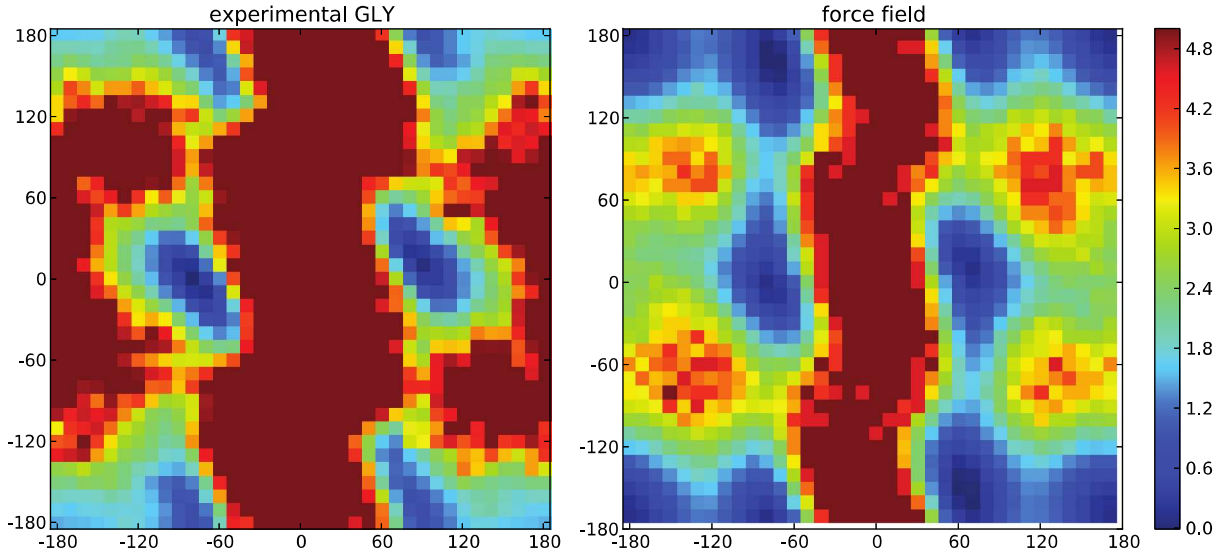
[1]Laufer Center for Physical and Quantitative Biology

[2]Department of Chemistry, University of Calgary

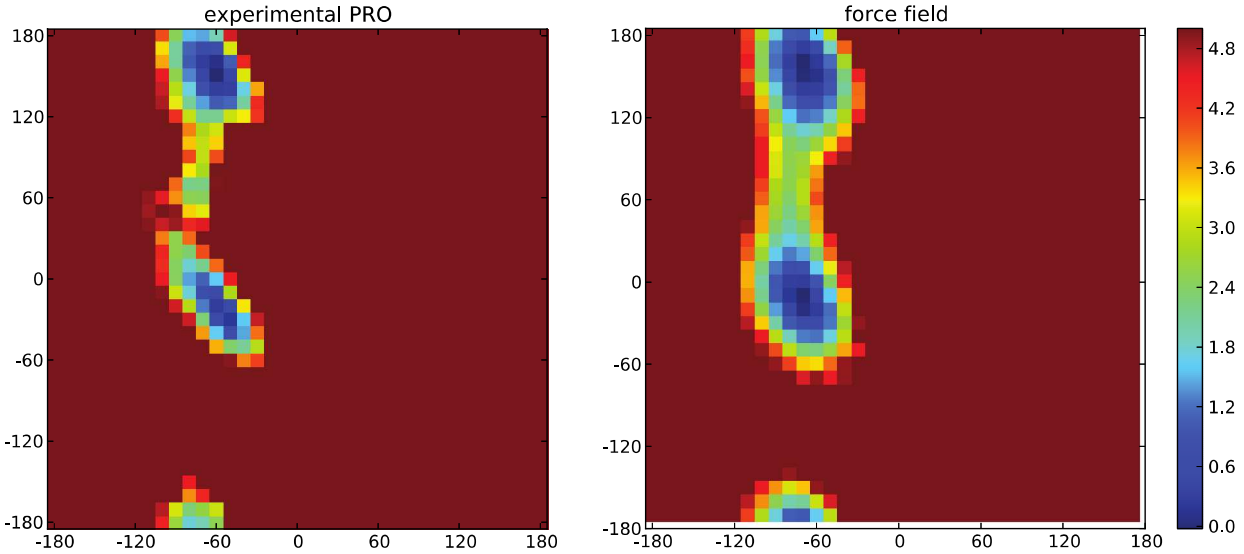
[3]Department of Chemistry, Stony Brook University

[4]Department of Physics and Astronomy, Stony Brook University

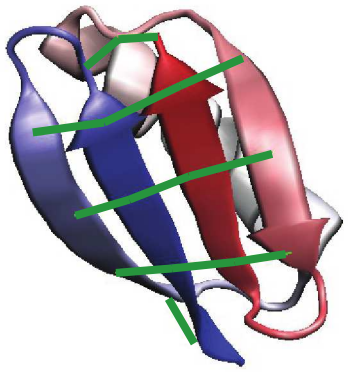
Supplementary Figures



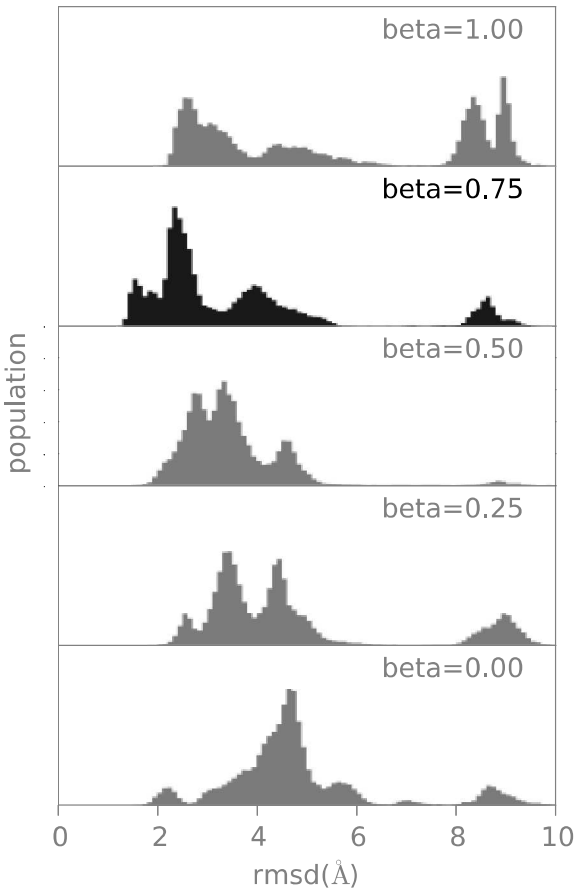
SI Figure S1. Ramachandran plot for Glycine. The experimental plot (left) was derived from data from the top500 dataset. The calculated plot with the original force field (right) was derived from trialanine implicit solvent simulations. Energies above 5 kcal/mol are shown in the plot as a value of 5kcal/mol. Phi angle values on the x axis and psi on the y axis.



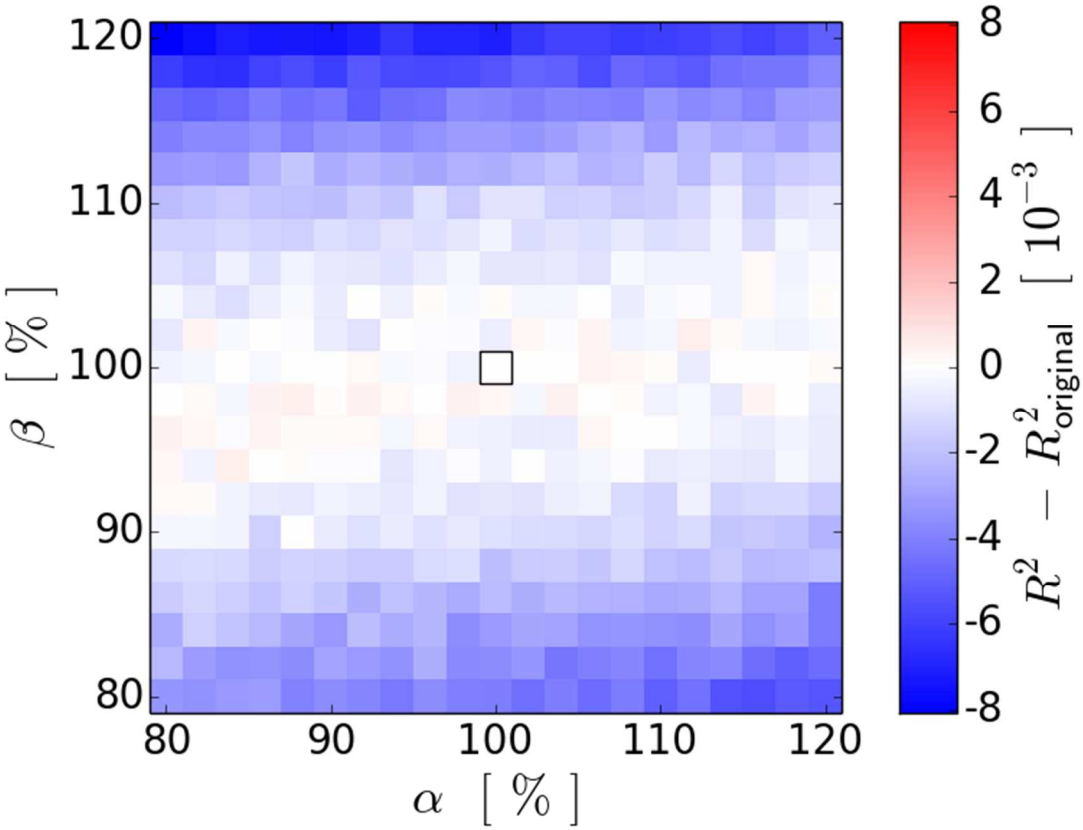
SI Figure S2. Ramachandran plot for Proline. The experimental plot (left) was derived from data from the top500 dataset. The calculated plot (right) was derived from trialanine implicit solvent simulations. Energies above 5 kcal/mol are shown in the plot as a value of 5kcal/mol. Phi angle values on the x axis and psi on the y axis.



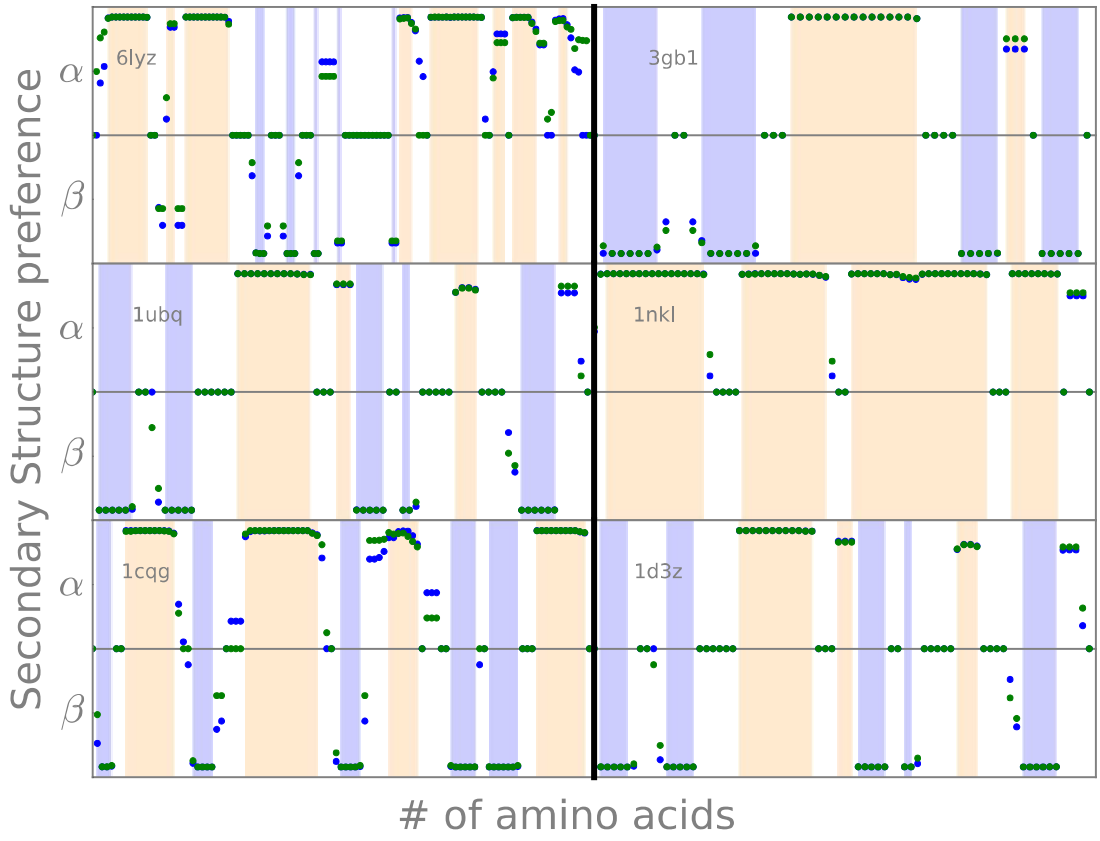
SI Figure S3. Folding of protein G. Restraints used (green) in the folding simulation of protein G.



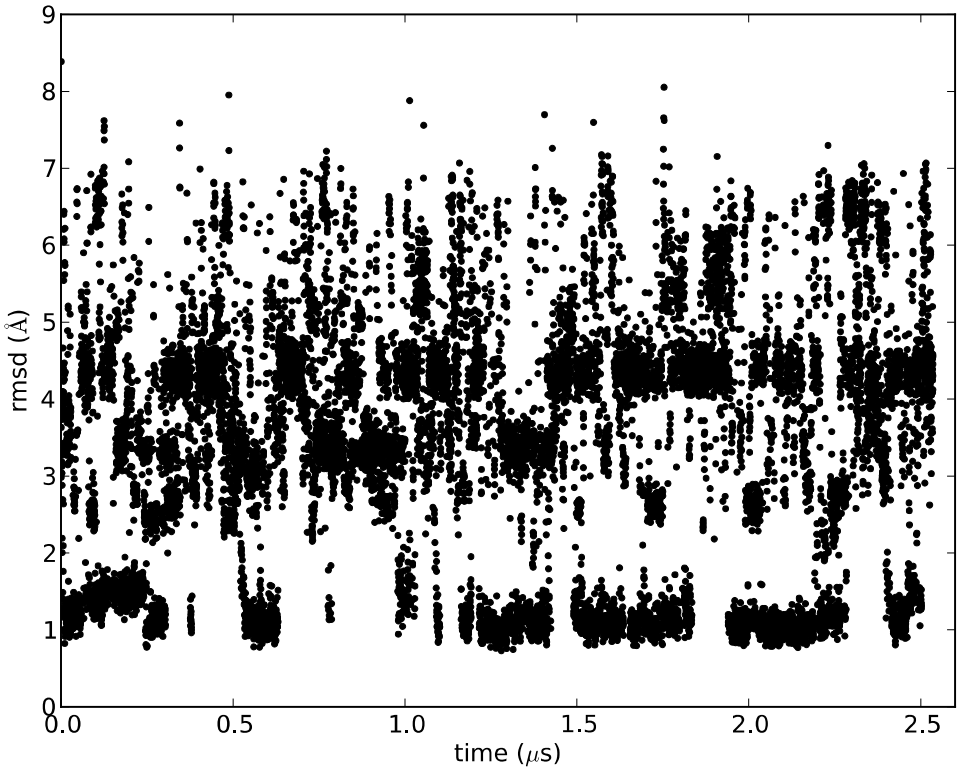
SI Figure S4. RMSD distribution from native in protein G folding simulations with different fixes for the beta region. Folding trajectories with AMAPbeta=0.75 is the one that has the most native like distribution.



SI figure S5. Sensitivity of the helical propensity to the scaling of the AMAP parameters. The x and y axes represent the scaling of the AMAP corrections for the alpha and beta region respectively. The square at the center of the graph (100% ; 100%) identifies the AMAP correction set used in the paper. The color map shows the difference between the coefficient of determination of the scaled set of parameters and the the original AMAP correction. Red tones identify an improvement of the results while blue tones identify a deterioration.



SI Figure S6. Secondary structure along sequence. Native like secondary structure is colored orange (α-helix) or blue (β-sheet) boxes. The preference for either α or β in the force fields is shown as a log($(N_\alpha+0.01)/(N_\beta+0.01)$), where N_i denotes the population percentages observed in simulations and 0.01 is added to avoid undetermined logs. In the plot values > 0 correspond to helix preference, < 0 to beta strand preference and = 0 means there is no preference. Green denotes ff13SB and blue denotes ff13SB+AMAP.



SI figure S7. Rmsd of Trpzip2 lowest temperature replica to the native structure (1HRX).

		α	3_{10}	π	β	Bridge	Coil	Turn
ff12SBff + AMAP	EK	53	5	0	0	0	33	10
	MAT	24	1	0	35	0	23	17
	NRF2	1	2	0	44	1	25	28
	PROTG	4	2	0	43	0	23	28
	RIBO	45	2	0	1	0	40	12
	TC5B	35	13	0	0	0	45	7
	TRPZIP	1	12	0	19	0	41	27
ff12SB	EK	54	4	0	0	0	33	9
	MAT	64	2	0	0	0	28	6
	NRF2	27	1	0	4	10	50	7
	PROTG	29	9	0	0	0	27	35
	RIBO	55	2	0	0	0	38	6
	TC5B	37	13	0	0	0	43	7
	TRPZIP	23	25	0	0	0	33	19

Table S1. Secondary structure percentages found in folding simulations for our test case of peptides using the original or modified force field. Notice that the original force field produces almost no beta secondary structure and in contrast has very large helical populations.