Supporting Information for:

Further along the Road Less Traveled: AMBER ff15ipq, an Original Protein Force Field Built on a Self-Consistent Physical Model

Karl T. Debiec[123], David S. Cerutti[4], Lewis R. Baker[3], Angela M. Gronenborn[2], David A. Case[5], and Lillian T. Chong[3*].

[1]Molecular Biophysics and Structural Biology Graduate Program, University of Pittsburgh and Carnegie Mellon University

[2]Department of Structural Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15260

[3]Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260

[4]Department of Chemistry, Michigan State University, East Lansing, MI 48824

[5]Department of Chemistry and Chemical Biology, Rutgers University, New Brunswick, NJ 08854

*To whom correspondence should be addressed. Tel: +1 412 624-6026; Fax: +1 412 624-8301; Email: ltchong@pitt.edu

Supporting Materials and Methods

S1. Simulations of Salt-Bridge Formation using Polarizable Force Fields

CHARMM Drude-2013 simulations. Systems were initially prepared with the CHARMM36 force field using psfgen,[1,2] equilibrated using NAMD 2.10.0,[3] and uploaded to the CHARMM-GUI server to obtain Drude-2013 topologies.[4,5] The Drude-2013 systems were then subjected to another round of energy minimization and equilibration prior to production simulation. The equilibration process for these systems involved three stages of unrestrained simulation in which a 1-ps simulation was first performed at constant pressure using a 0.1-fs timestep, followed by a 20-ps simulation at constant pressure using a 0.5-fs timestep, and finally a 1-ns simulation at constant pressure using a 1-fs timestep. Production simulations were then carried out for 100 ns at constant pressure using a 1-fs timestep. Throughout equilibration and production temperature was maintained at 298 K using a dual Langevin thermostat (frictional constant of 1 ps^{-1}) and 1 K for Drude particles (frictional constant of 20-ps^{-1}),[6] while pressure was maintained at 1 atm using a Langevin piston barostat (piston period and decay time of 200 fs and 100 fs, respectively).[7] Van der Waals interactions were smoothly switched off between 8 and 10 Å, while short-range electrostatic interactions were truncated at 10 Å and long-range electrostatic interactions were calculated using the particle mesh Ewald method.[8] To enable a 1-fs time step, a hard-wall barrier was applied to restrict the maximum distance between Drude particles and their hosts to 0.2 Å,[9] and bonds to hydrogen were constrained to their equilibrium lengths using the M-SHAKE algorithm.[10]

AMOEBA simulations. Systems were built using TINKER and converted to AMBER format using the AMBERTools 15 software package.[11,12] Each system was subjected to energy minimization followed by a 20-ps simulation at constant temperature and a 1-ns simulation at constant pressure. Production simulations were then carried out for 50 ns at constant pressure. Throughout equilibration and production temperature was maintained at 298 K using a Langevin thermostat (frictional constant of 1 ps^{-1}), while pressure was maintained at 1 atm using a Berendsen barostat (time constant of 1-ps).[13] Van der Waals interactions were truncated at 12 Å, while short-range electrostatic interactions were truncated at 7 Å and long-range electrostatic interactions were calculated using the PME method.[8] Dynamics were integrated using a 1-fs timestep.

S2. Derivation of Amino Acid Side-Chain Analogue Parameters with AMOEBA

Parameters that are consistent with the AMOEBA polarizable force field were derived for imidazolium, guanidinium, and acetate using the software Poltype.[14] The structures of these side-chain analogues were optimized at the HF/6-31G* level of theory, followed by density calculation at the MP2/6-311G** level of theory using the Gaussian 09 software package.[15] Multipoles were fit using Stone's distributed multipole analysis as implemented in the GDMA program,[16,17] and refined via fitting of the electrostatic potential using a convergence criterion of 0.1 kcal/mol•$electron^2$. Van der Waals radii and well-depths were assigned based on both the element and valence orbitals of each atom, while atomic polarizabilities were assigned based solely on the element. Parameters for bonded interactions were selected based on the analogues' chemical connectivity from a database of small-molecule parameters.[14]

S3. Simulations of GB3, Ubiquitin, and Binase with AMBER ff99SB-ILDN and CHARMM22*

Systems for the simulations reported in Figure S1 were prepared and equilibrated using the Desmond 3.0.1.0 software package.[18] Each system was subjected to energy minimization followed by a 20-ps equilibration at constant temperature, and a 1-ns equilibration at constant pressure. Temperature was maintained at 298 K and pressure at 1 atm using the Martyna-Tobias-Klein thermostat and barostat

(time constants of 1 ps and 2 ps, respectively).[19] To enable a 2-fs time step, bonds to hydrogen were constrained to their equilibrium values using the M-SHAKE algorithm.[10] A short-range nonbonded cutoff of 10 Å was used, and long-range electrostatics were calculated using the particle mesh Ewald (PME) method.[8]

Production simulations were carried out for 1 μs at constant pressure using a 64-node Anton special-purpose supercomputer and the Multigrator integrator.[20,21] Temperature was maintained at 298 K using the Nosé-Hoover thermostat and pressure at 1 atm using the Martyna-Tobias-Klein barostat (time constants of 1 ps). To enable a 2.5-fs time step, bonds to hydrogen were constrained to their equilibrium values using the M-SHAKE algorithm.[10] Van der Waals and short-range electrostatic interactions were truncated at 10 Å; long-range electrostatic interactions were calculated using the Gaussian split Ewald method,[22] and were updated every third timestep.

Supporting Figures

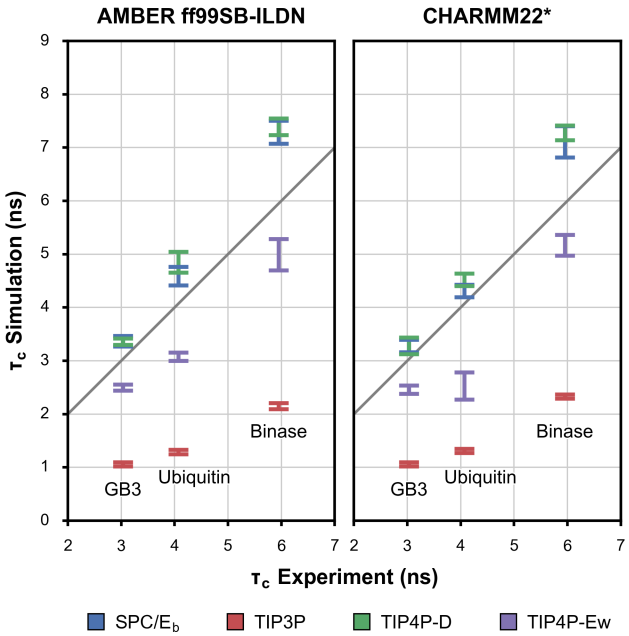

Figure S1. Comparison of simulated vs. experimental rotational correlation times τ_c of GB3, ubiquitin, and binase. Simulated τ_c values were obtained using two different force fields (AMBER ff99SB-ILDN and CHARMM22*)[23,24] and four different water models (SPC/E_b, TIP3P, TIP4P-D, and TIP4P-Ew).[25–28] Experimental τ_c values were measured using NMR relaxation and corrected for differences in temperature and D_2O content between simulation and experiment.[29–32] Error bars represent 95% confidence intervals; standard errors were obtained by dividing each 1-μs simulation into five 200-ns blocks and calculating the standard deviation.

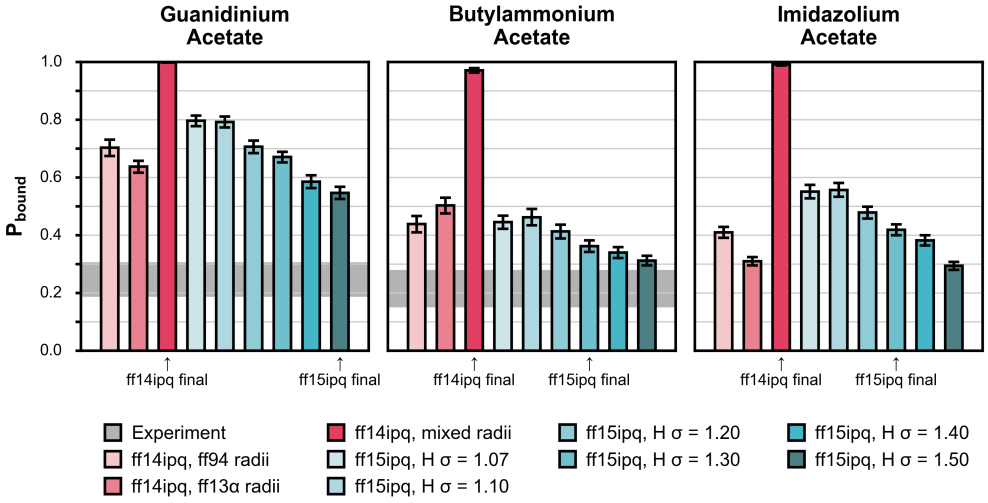

Figure S2. Probability of binding (P_{bound}) between acetate and one or more molecules of three cationic side-chain analogues using the ff14ipq and ff15ipq charge sets paired with different Lennard Jones radii. The ff13α radii tested with ff14ipq included an increase in the σ of the carboxylate oxygen (type 'O') and a decrease in the ammonium nitrogen (type 'N') relative to ff94;[33] the final version used 'mixed radii' in which the ff13α σ were applied only to interactions with water.[34] The ff14ipq force field was simulated with the TIP4P-Ew water model,[28] while ff15ipq was simulated with SPC/E_b.[25] The P_{bound} values corresponding to the experimentally-determined K_A values of guanidinium acetate and butylammonium acetate are depicted as horizontal gray bars;[35,36] no experimental value is available for the imidazolium acetate system. Error bars represent 95% confidence intervals calculated using a block averaging method.[37]

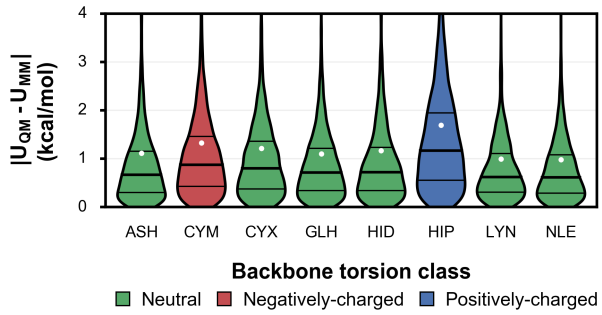

Figure S3. Distributions of errors in molecular mechanical energies U_{MM} relative to their quantum mechanical targets U_{QM} for amino acid dipeptides representing alternative protonation states (Ash, Cym, Glh, Hid, Hip, Lyn), the disulphide form of cysteine (Cyx), and the noncanonical amino acid norleucine (Nle). 25th, 50th, and 75th percentiles are represented by horizontal lines, and root mean square errors are represented by white circles. Each dataset is colored by its corresponding backbone torsion class.

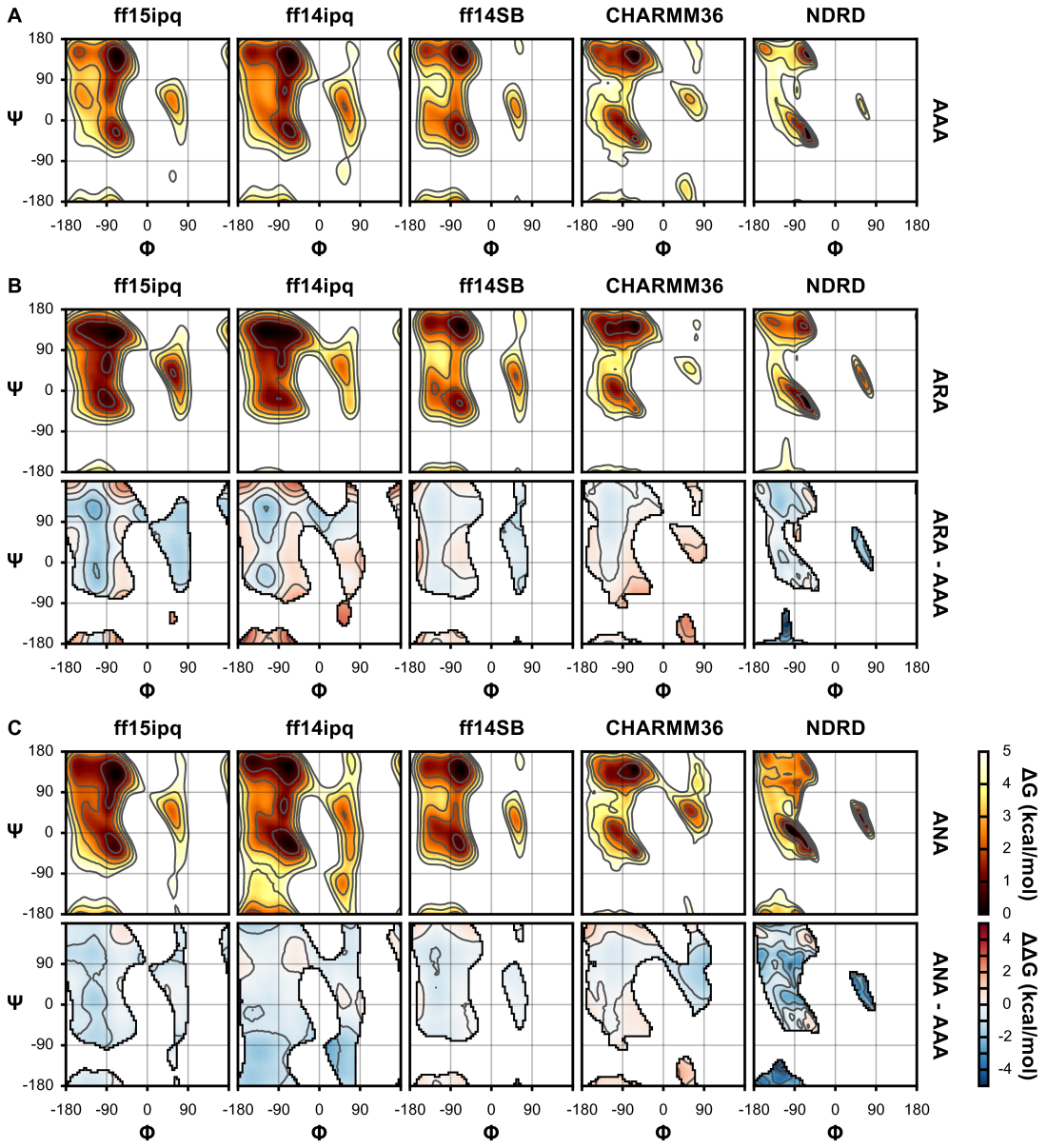

Figure S4. Comparison of residue-specific conformational preferences of the central residue of Ace-Ala-Xaa-Ala-Nme tetrapeptides observed in umbrella sampling simulations with distributions of Ala-Xaa-Ala obtained from the Neighbor-Dependent Ramachandran Distribution (NDRD) dataset consisting of loop conformations observed in experimentally solved structures.[38] Simulations were performed using four different force fields (AMBER ff15ipq, ff14ipq, ff14SB, and CHARMM36),[1,34,39] which were each paired with either the water model with which they were derived or that with which they are most-commonly used. For each system, both the absolute free energy as well as the difference in free energy relative to Ala-Ala-Ala is shown. Regions in which the free energies of both Ala-Xaa-Ala and Ala-Ala-Ala are greater than 5 kcal/mol are omitted for clarity.

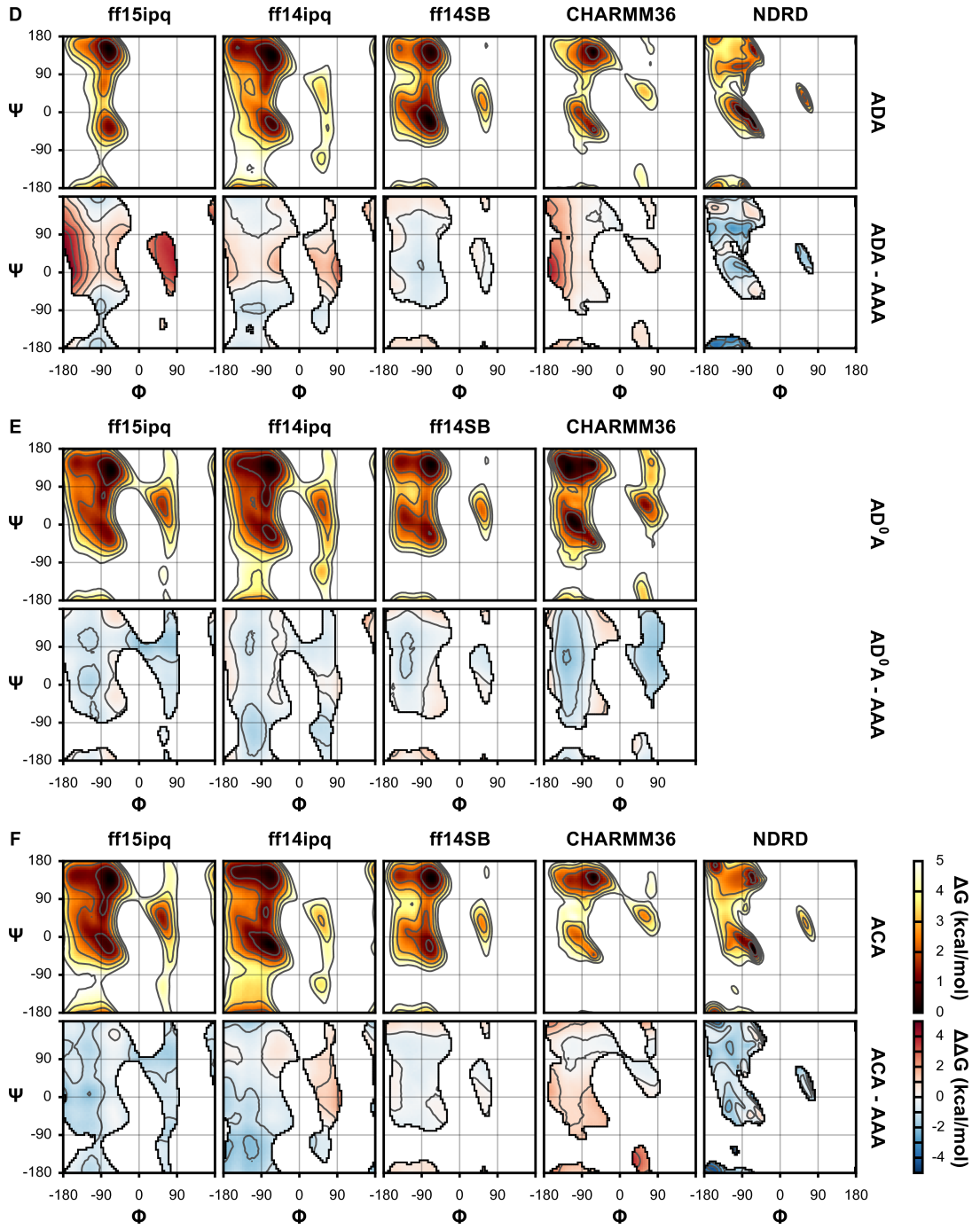

Figure S4 (Continued).

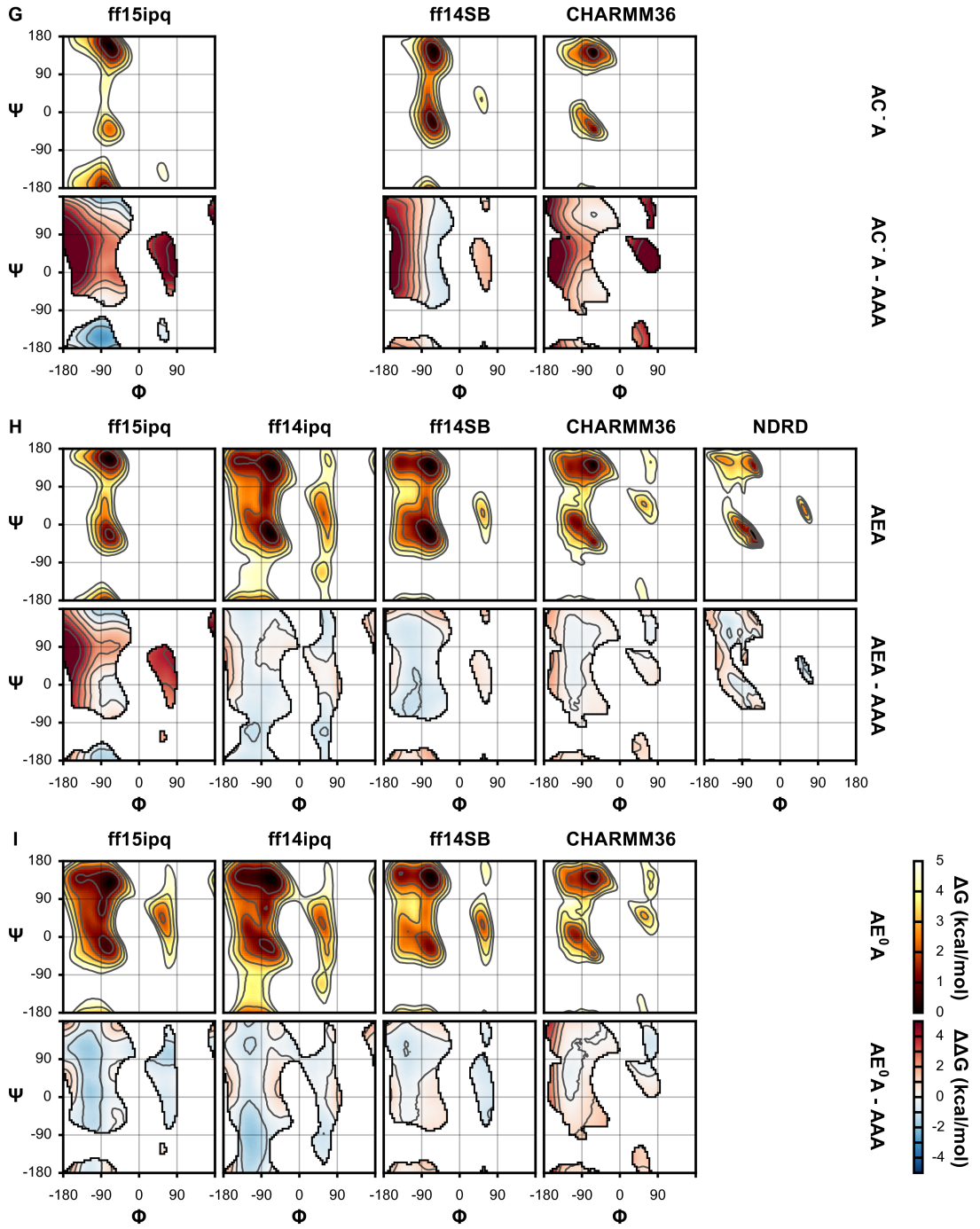

Figure S4 (Continued).

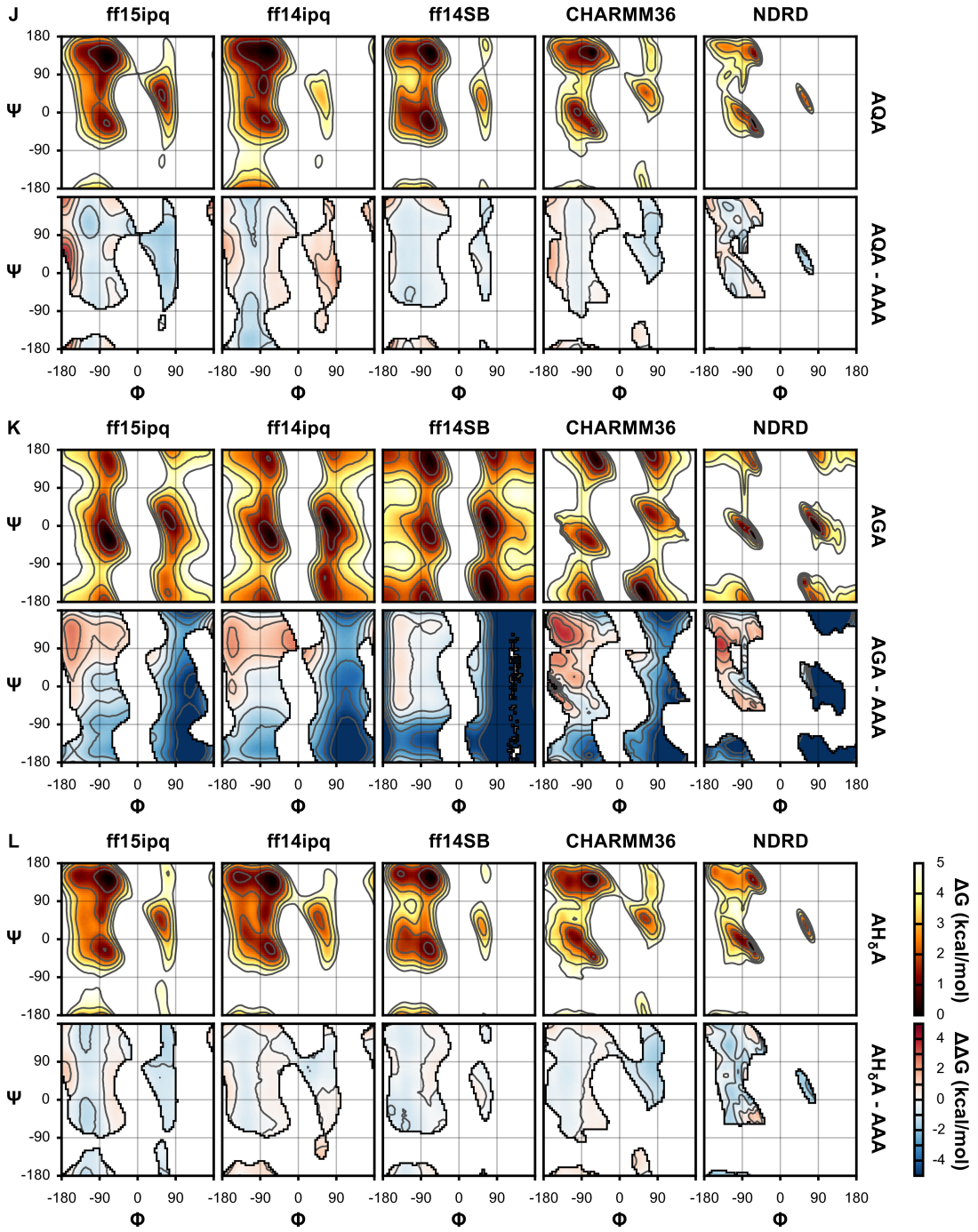

Figure S4 (Continued).

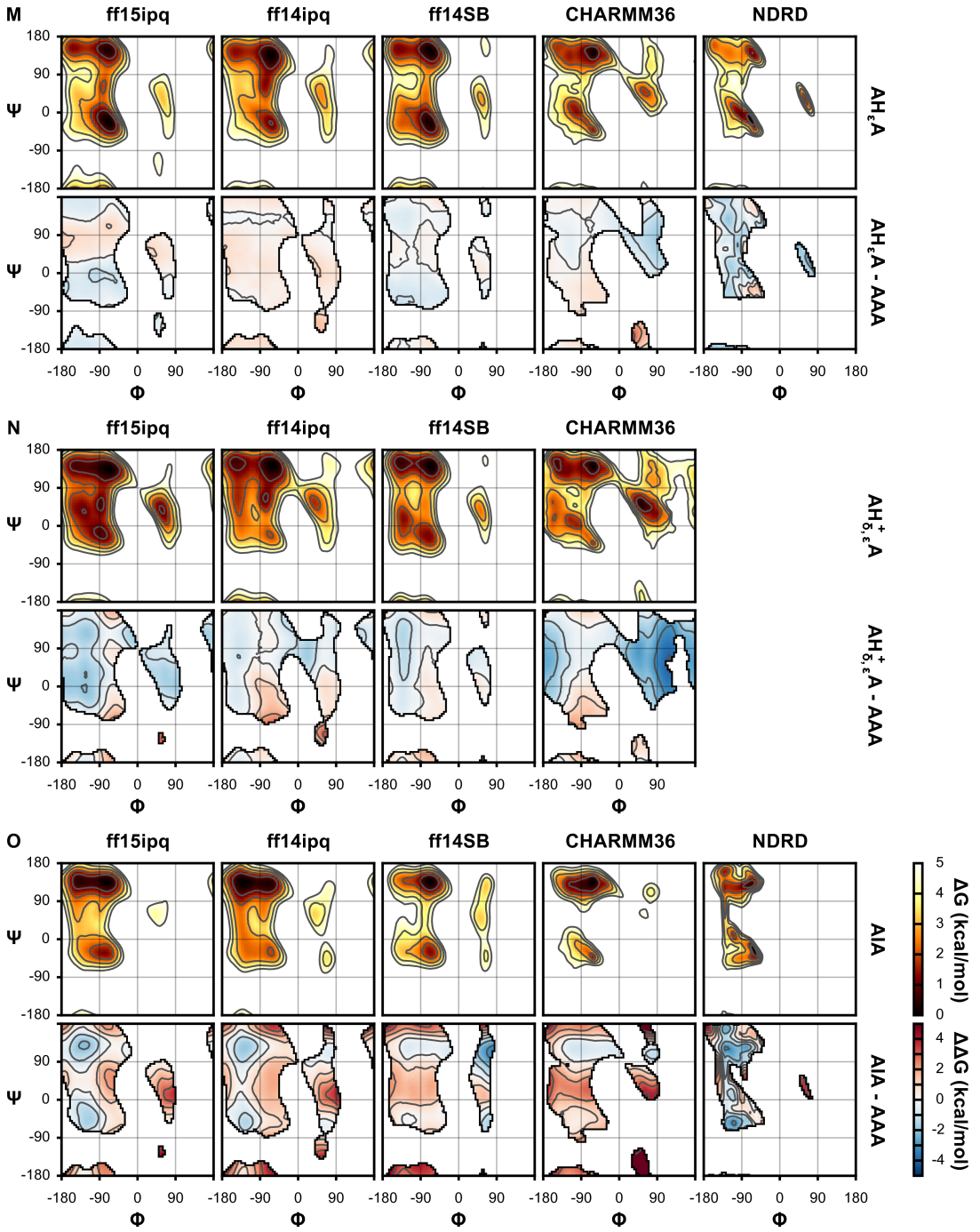

Figure S4 (Continued).

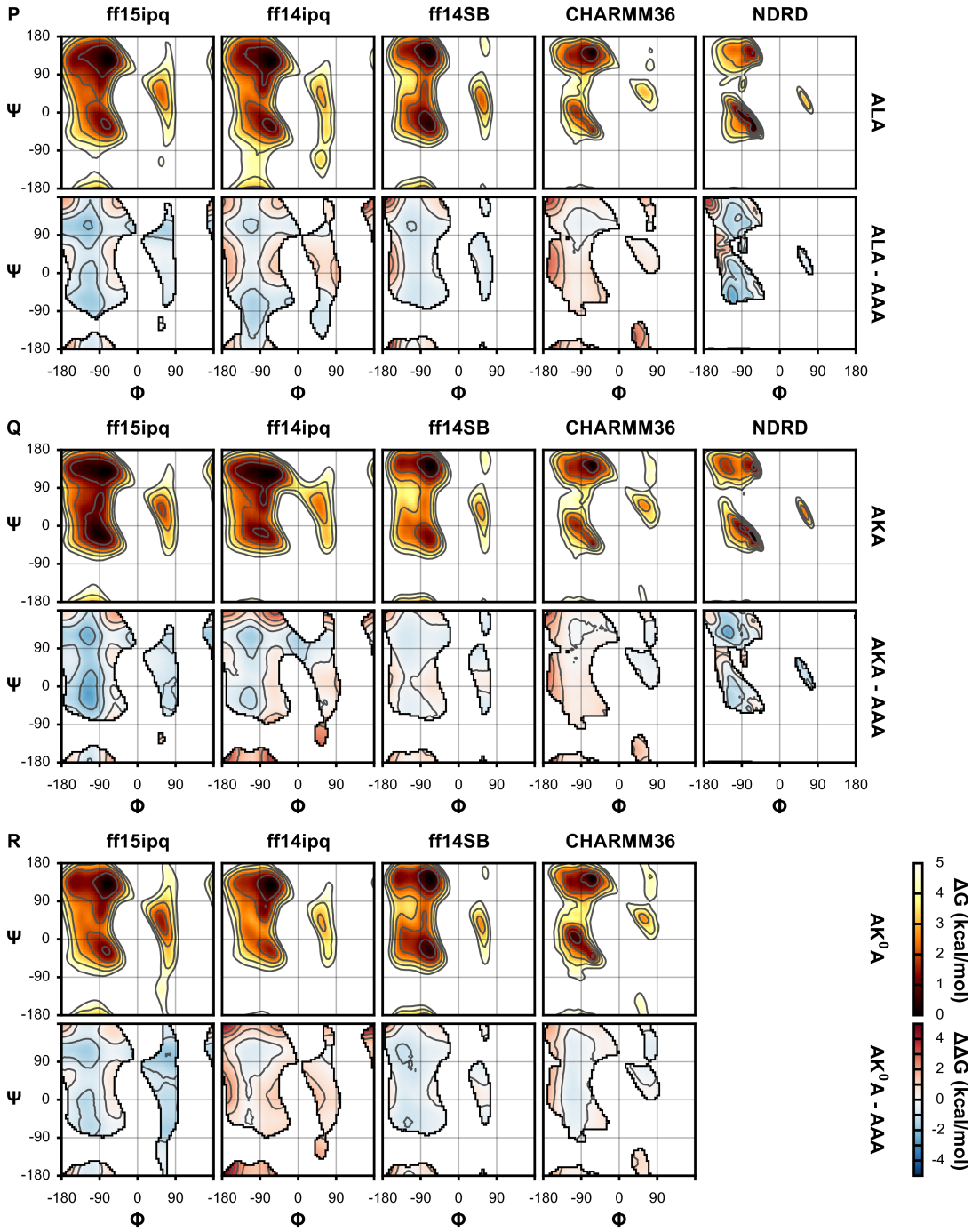

Figure S4 (Continued).

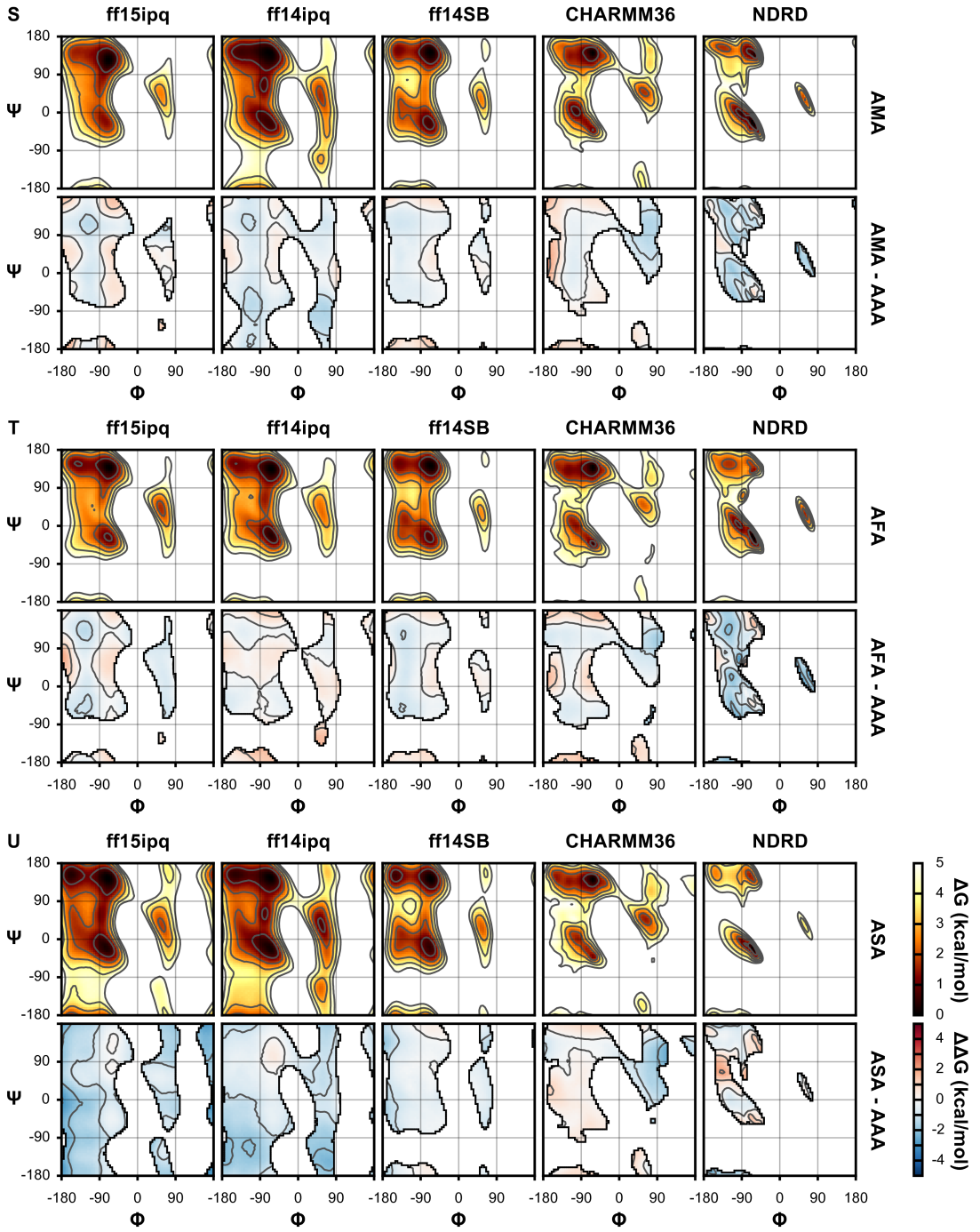

Figure S4 (Continued).

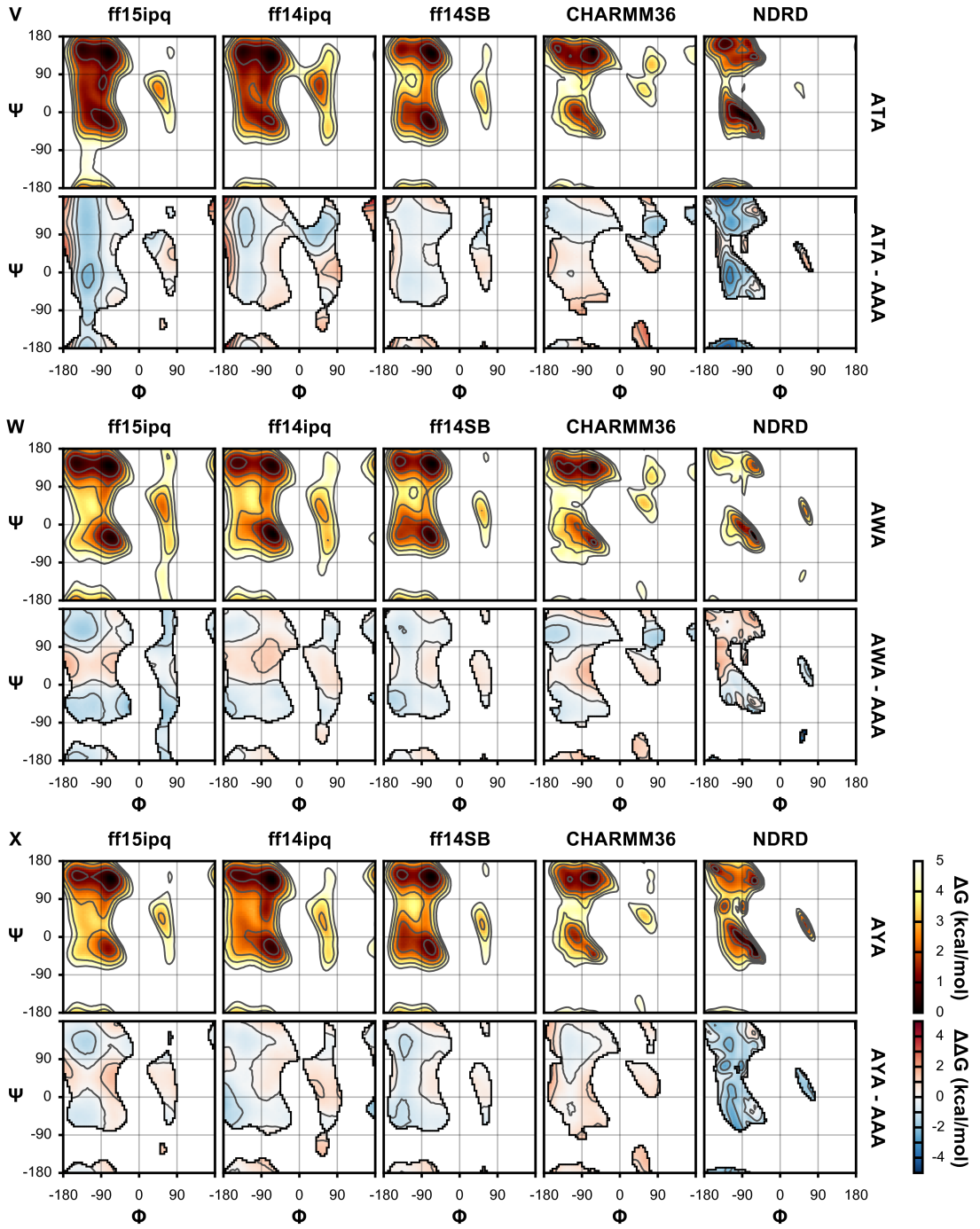

Figure S4 (Continued).

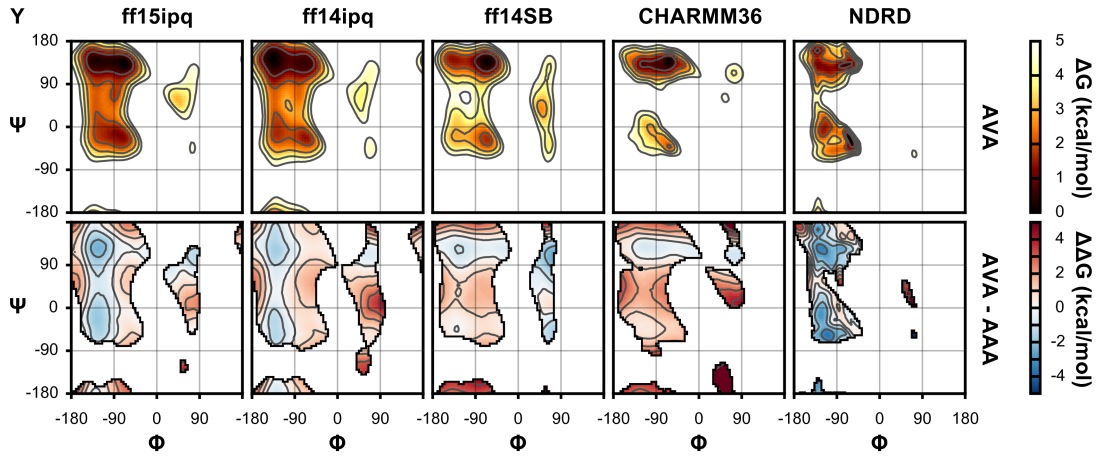

Figure S4 (Continued).

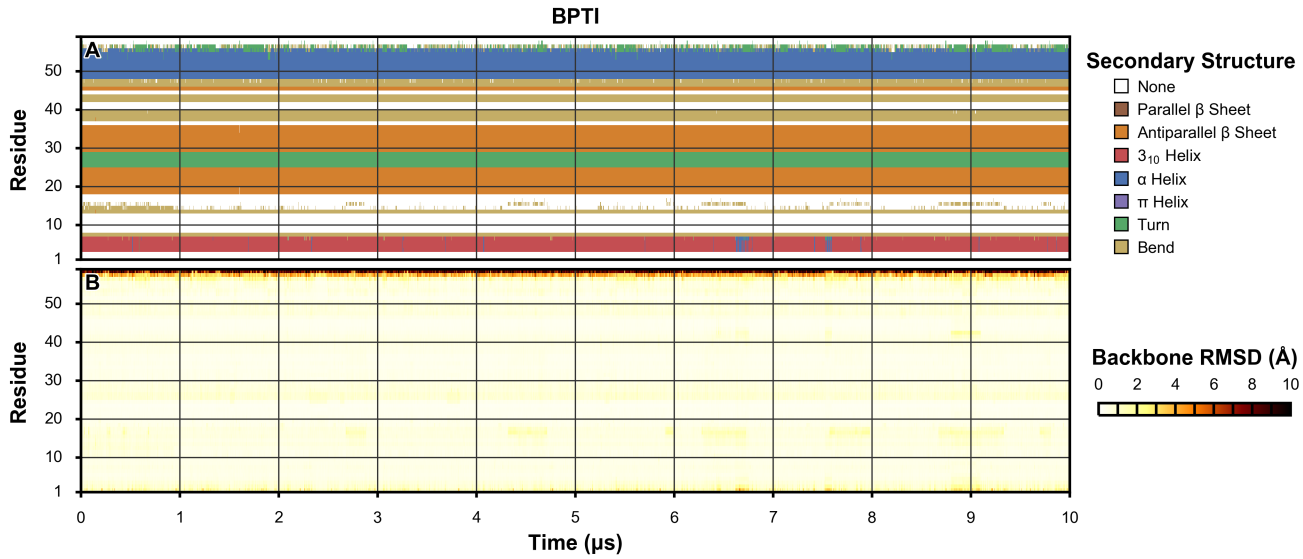

Figure S5. Secondary structure (**A**) and per-residue backbone RMSD relative to the crystal structure (PDB code 5PTI)[40] (**B**) of BPTI observed over the course of a 10-μs simulation.

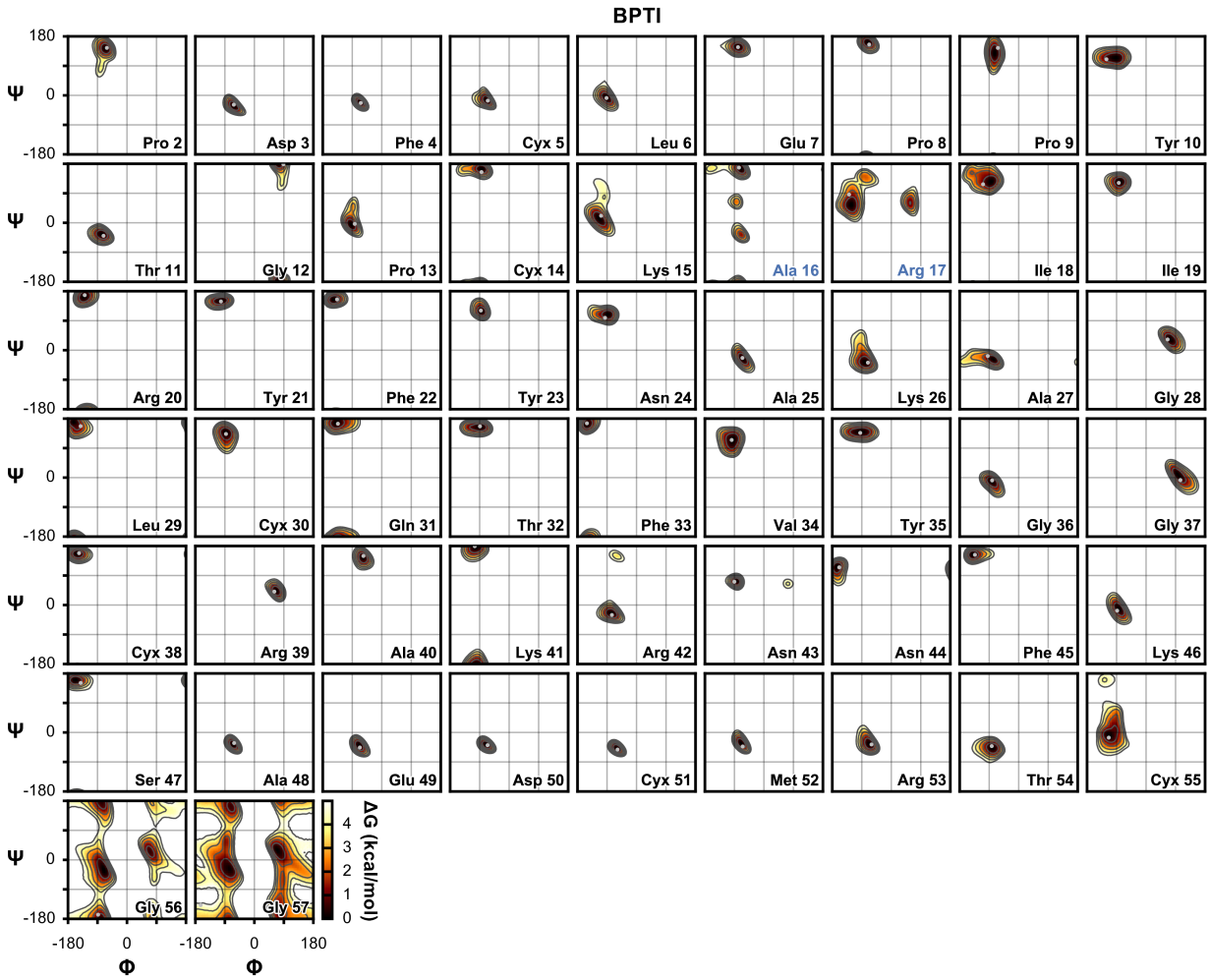

Figure S6. Backbone conformational sampling of BPTI observed in a 10-μs simulation. The Φ/Ψ angles observed in the crystal structure (PDB code 5PTI)[40] are shown as gray points. Overall retention of the crystal conformation of most residues is good; regions of deviation mentioned in the main text are highlighted in blue.

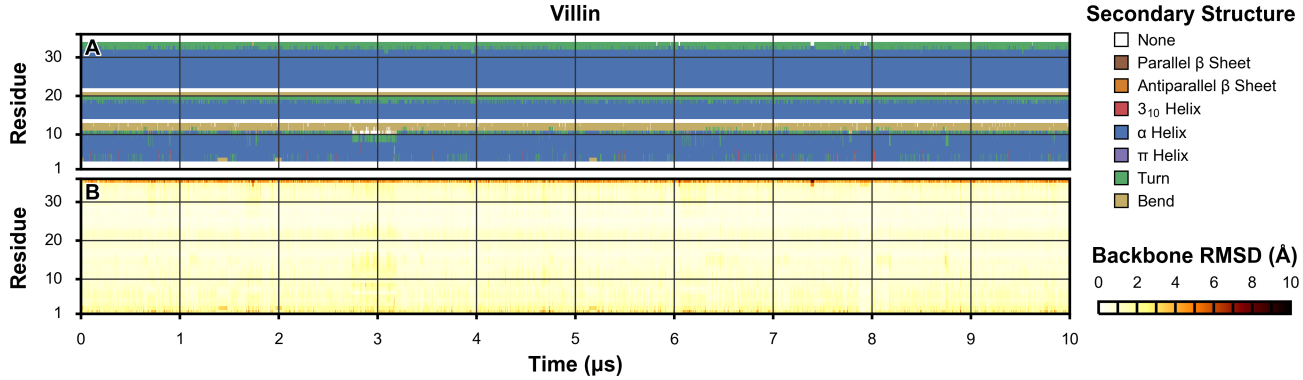

Figure S7. Secondary structure (**A**) and per-residue backbone RMSD relative to the crystal structure (PDB code 2F4K)[41] (**B**) of the double-norleucine mutant of the villin headpiece subdomain observed over the course of a 10-μs simulation.

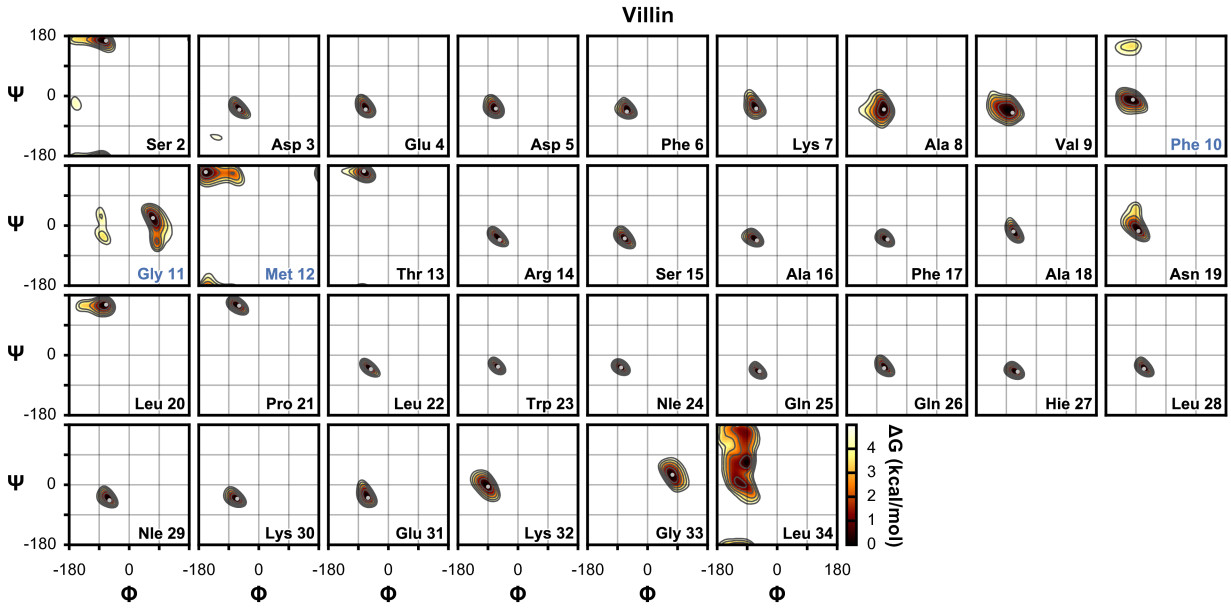

Figure S8. Backbone conformational sampling of the double-norleucine mutant of the villin headpiece subdomain observed in a 10-μs simulation. The Φ/Ψ angles observed in the crystal structure (PDB code 2F4K)[41] are shown as gray points. Overall retention of the crystal conformation of most residues is good; regions of deviation mentioned in the main text are highlighted in blue.

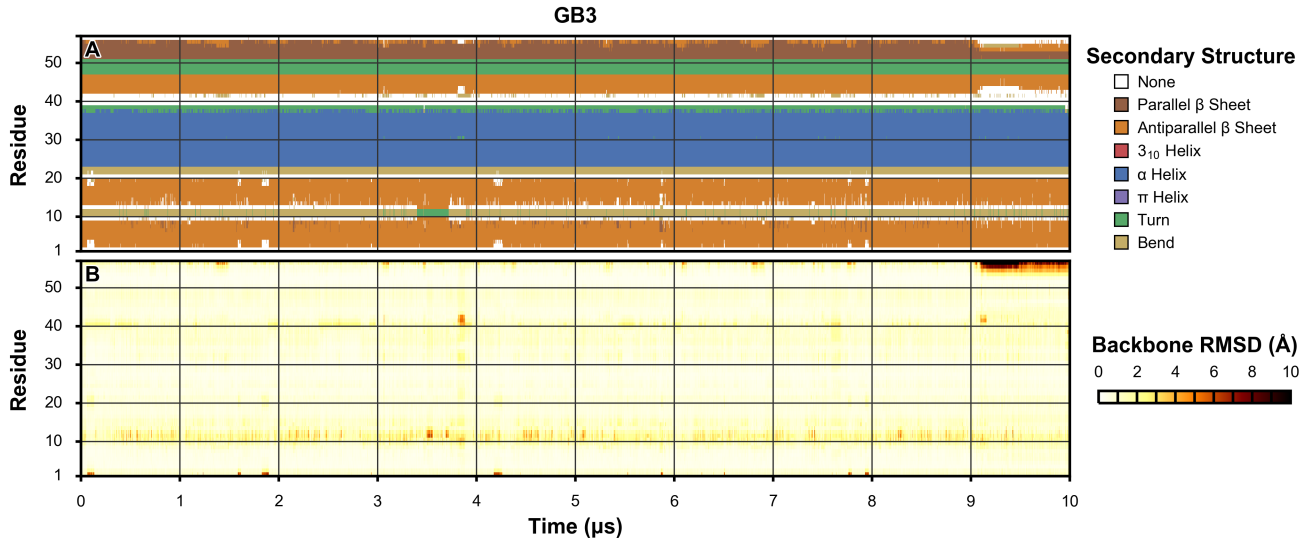

Figure S9. Secondary structure (**A**) and per-residue backbone RMSD relative to the NMR structure (PDB code 1P7E)[42] (**B**) of GB3 observed over the course of a 10-μs simulation.

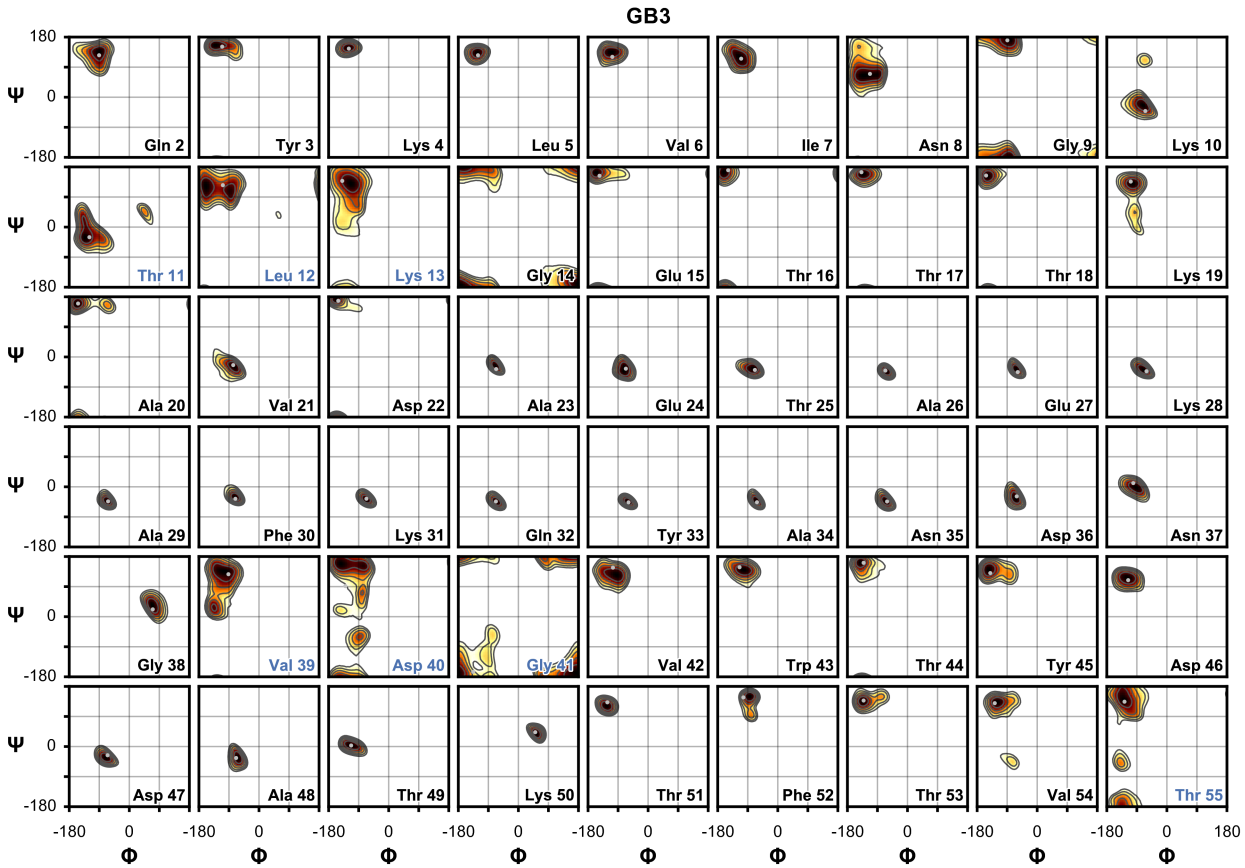

Figure S10. Backbone conformational sampling of GB3 observed in a 10-μs simulation. The Φ/Ψ angles observed in the NMR structure (PDB code 1P7E)[42] are shown as gray points. Overall retention of the NMR conformation of most residues is good; regions of deviation mentioned in the main text are highlighted in blue.

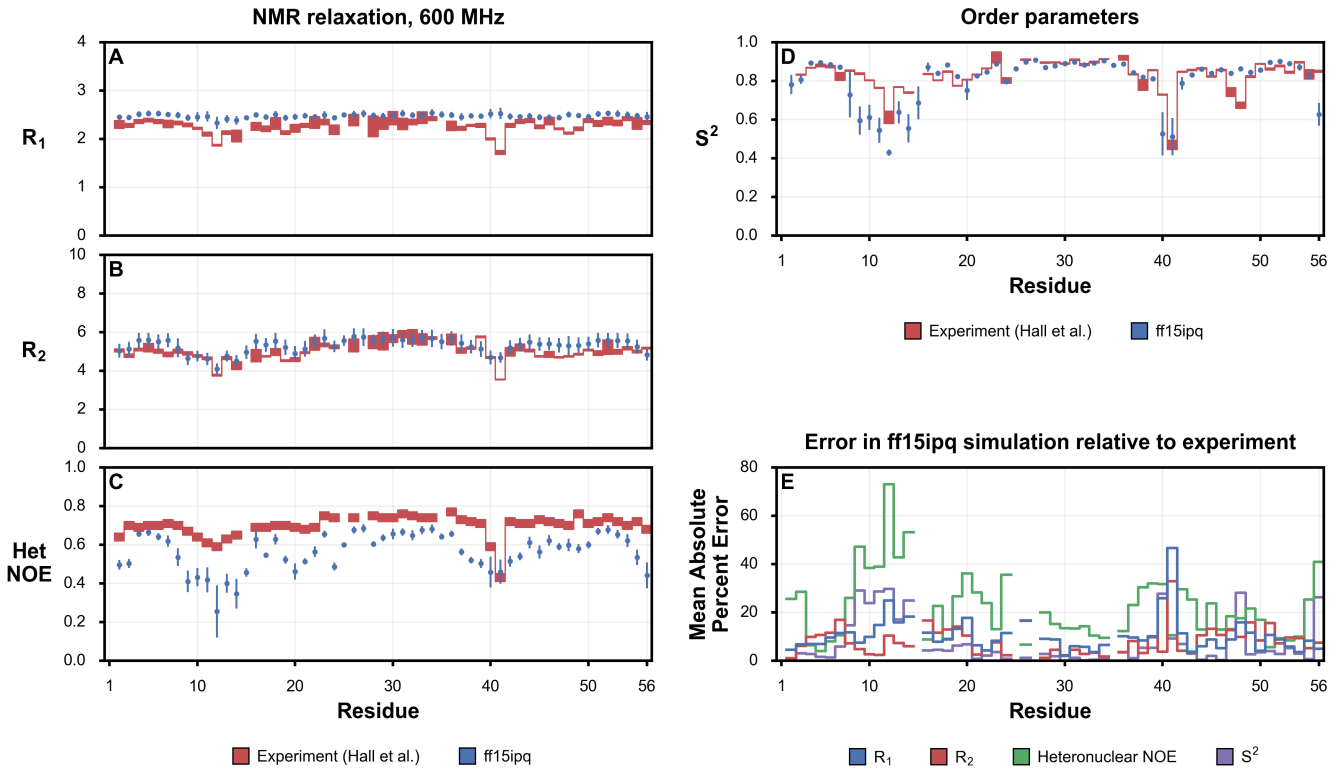

Figure S11 Comparison of NMR relaxation and order parameters for the GB3 protein from simulation and experiment. All parameters from simulation were calculated by applying the iRED method[43] to a set of ten 200-μs simulations with the ff15ipq force field, considering the autocorrelation function of each backbone N-H vector to 17 ns, which is five times the overall rotational correlation time (τ_c) of the protein.[44] Experimental R_1, R_2, and heteronuclear NOE data were available at five magnetic field strengths (400, 500, 600, 700, and 800 MHz), and experimental order parameters (S^2) were based on the combined data set from all five magnetic field stengths.[29] At a magnetic field strength of 600 MHz (**A**, **B**, **C)**, ff15ipq yields good agreement in R_2, reasonable agreement in R_1 with a consistent offset, and somewhat poorer agreement in heteronuclear NOE; similar agreement is achieved at other magnetic field strengths. Error bars represent 95% confidence intervals; standard errors were obtained by calculating the standard deviation across the ten simulations. Comparison of simulated and experimental S^2 values (**D**) shows acceptable agreement, with deviations in many of the same regions as those observed for the NMR relaxation parameters. The mean absolute percent error (MAPE) in simulated R_1, R_2, and heteronuclear NOE (**E**) was calculated by averaging the proportional errors in each parameter across the five magnetic field strengths, while the error in S^2 was calculated relative to the single experimental dataset. Residues with above-average MAPE in R_1 and R_2, including Leu 12, Asp 40, and Gly 41, are discussed in the main text.

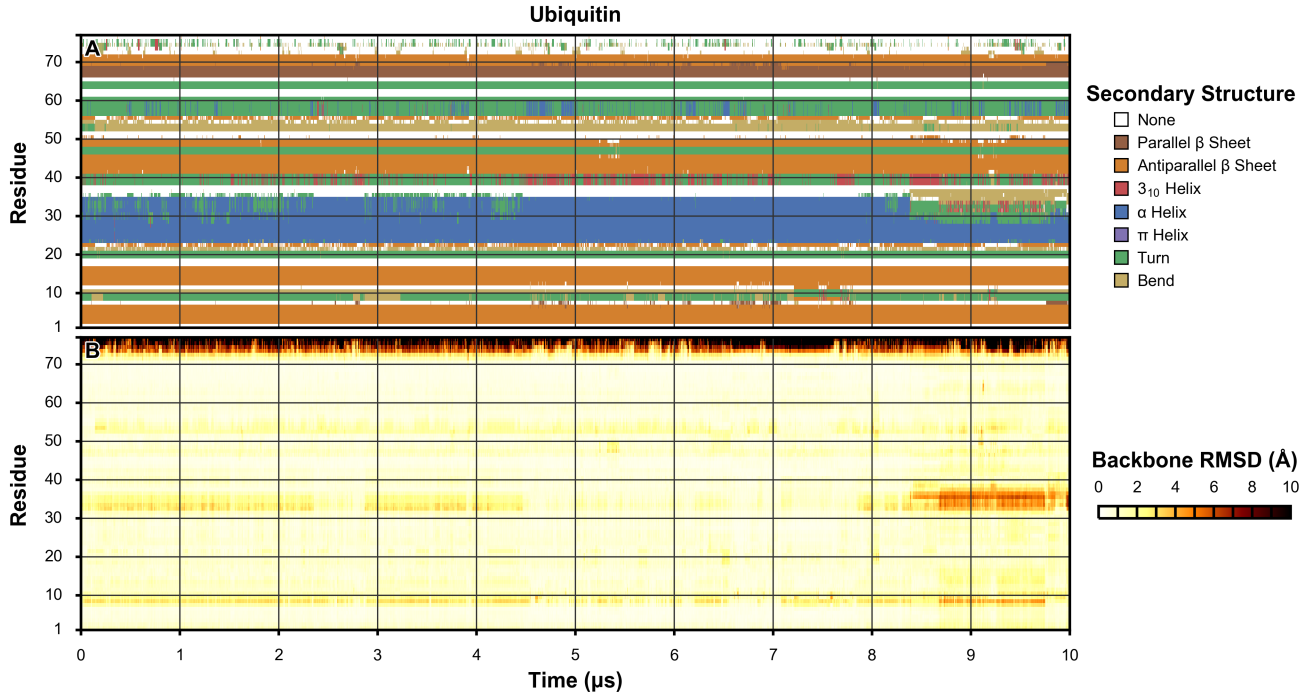

Figure S12. Secondary structure (**A**) and per-residue backbone RMSD relative to the crystal structure (PDB code 1UBQ)[45] (**B**) of ubiquitin observed over the course of a 10-μs simulation.

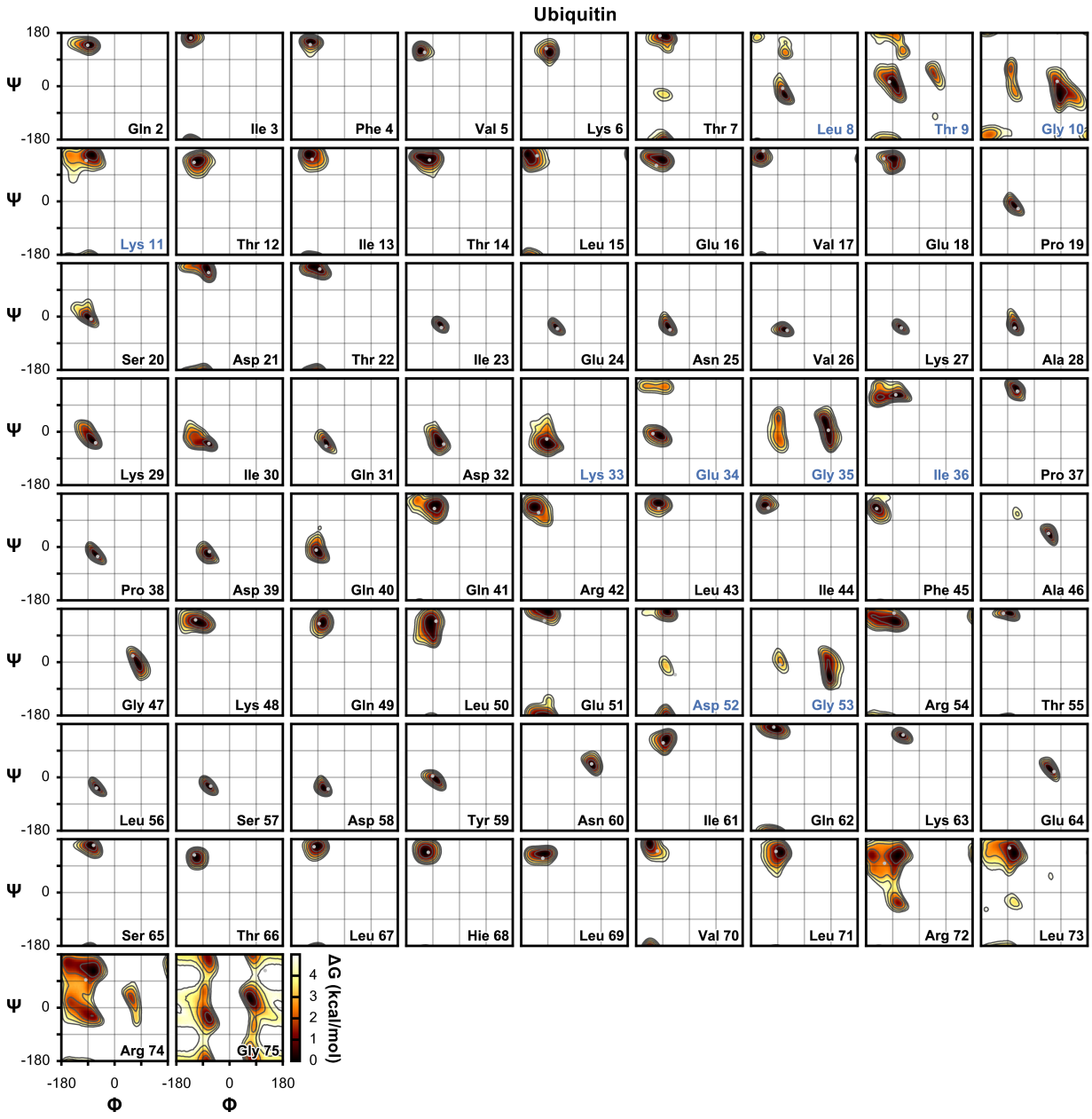

Figure S13. Backbone conformational sampling of ubiquitin observed in a 10-μs simulation. The Φ/Ψ angles observed in the crystal structure (PDB code 1UBQ)[45] are shown as gray points. Overall retention of the crystal conformation of most residues is good; regions of deviation mentioned in the main text are highlighted in blue.

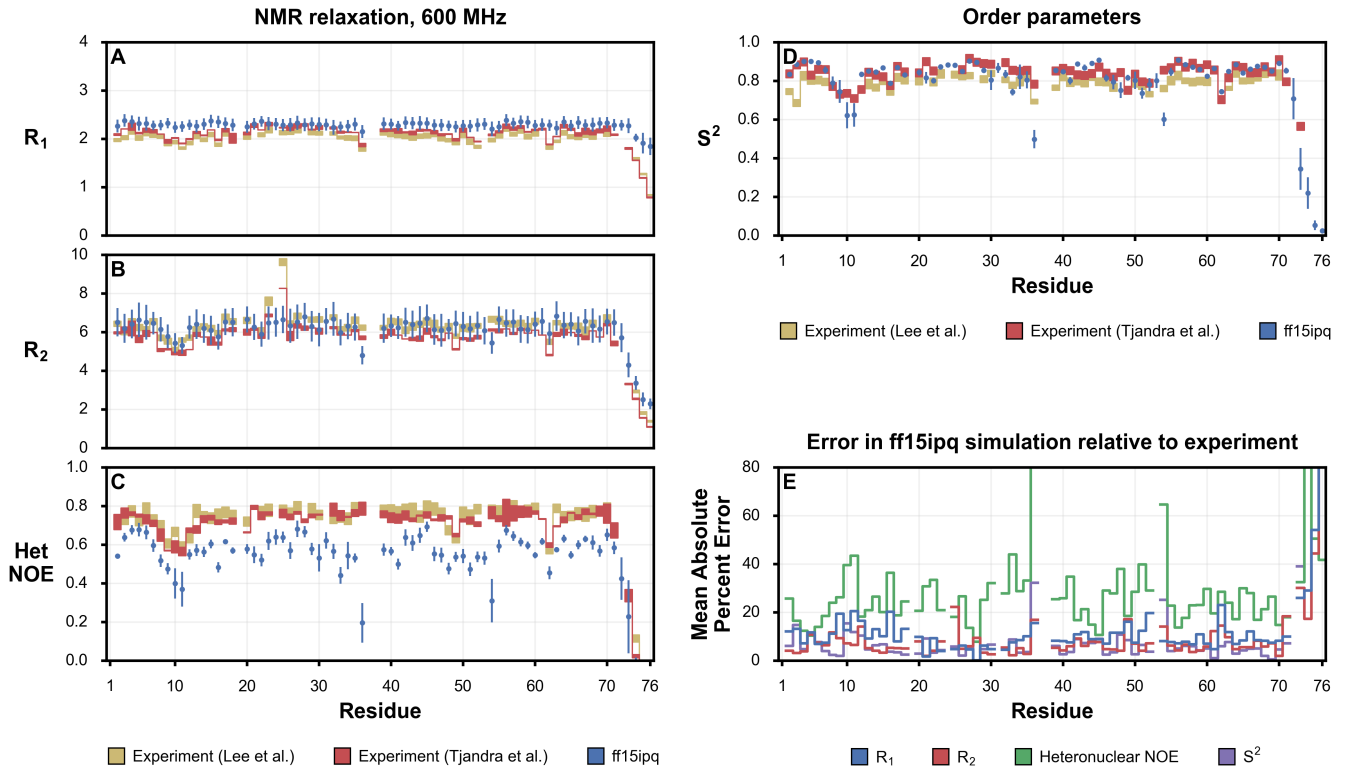

Figure S14 Comparison of NMR relaxation and order parameters for the ubiquitin protein from simulation and experiment. All parameters from simulation were calculated by applying the iRED method[43] to a set of ten 200-ns simulations with the ff15ipq force field, considering the autocorrelation function of each backbone N-H vector to 23 ns, which is five times the overall rotational correlation time (τ_c) of the protein.[44] Experimental R_1, R_2, and heteronuclear NOE data were available from Lee *et al.* at four magnetic field strengths (400, 500, 600, and 750 MHz), and from Tjandra *et al.* at 600 MHz, and experimental order parameters (S^2) were available from both groups.[30,46] At a magnetic field strength of 600 MHz (**A**, **B**, **C**), ff15ipq obtains good agreement in R_2, reasonable agreement in R_1 with a consistent offset, and somewhat poorer agreement in heteronuclear NOE; similar agreement is achieved at other magnetic field strengths. Error bars represent 95% confidence intervals; standard errors were obtained by calculating the standard deviation across the ten simulations. Comparison of simulated and experimental order parameters (**D**) shows acceptable agreement, with deviations in many of the same regions as observed for the NMR relaxation parameters. The mean absolute percent error (MAPE) in simulated R_1, R_2, and heteronuclear NOE (**E**) was calculated by averaging the proportional errors in each parameter across the four magnetic field strengths, while the error in S^2 was calculated relative to the two experimental datasets. Residues with above-average MAPE in R_1 and R_2, including Lys 11, Asn 25, and Asp 52, are discussed in the main text.

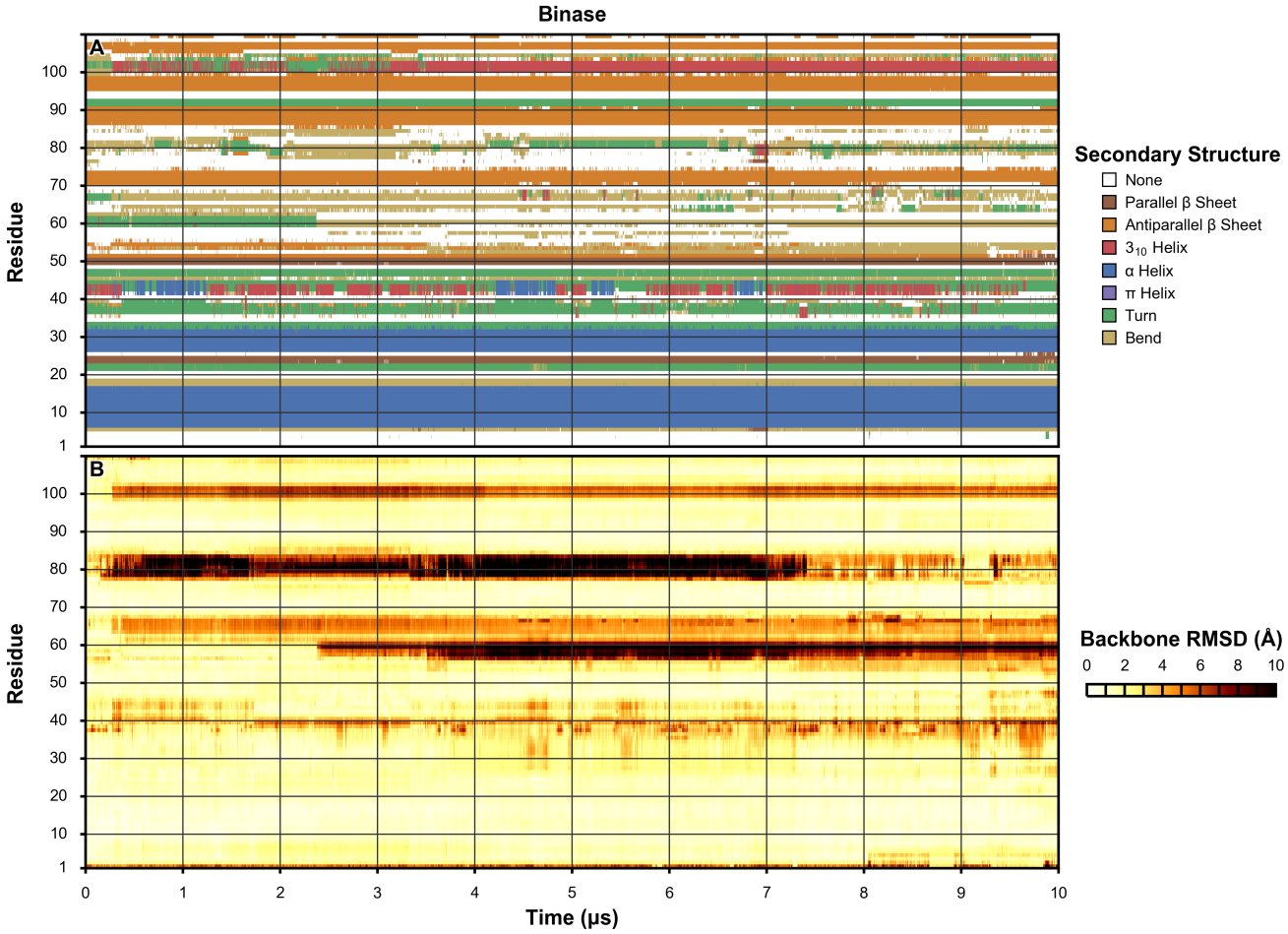

Figure S15. Secondary structure (**A**) and average per-residue backbone RMSD relative to the twenty structures of the NMR ensemble (PDB code 1BUJ)[47] (**B**) of binase observed over the course of a 10-μs simulation.

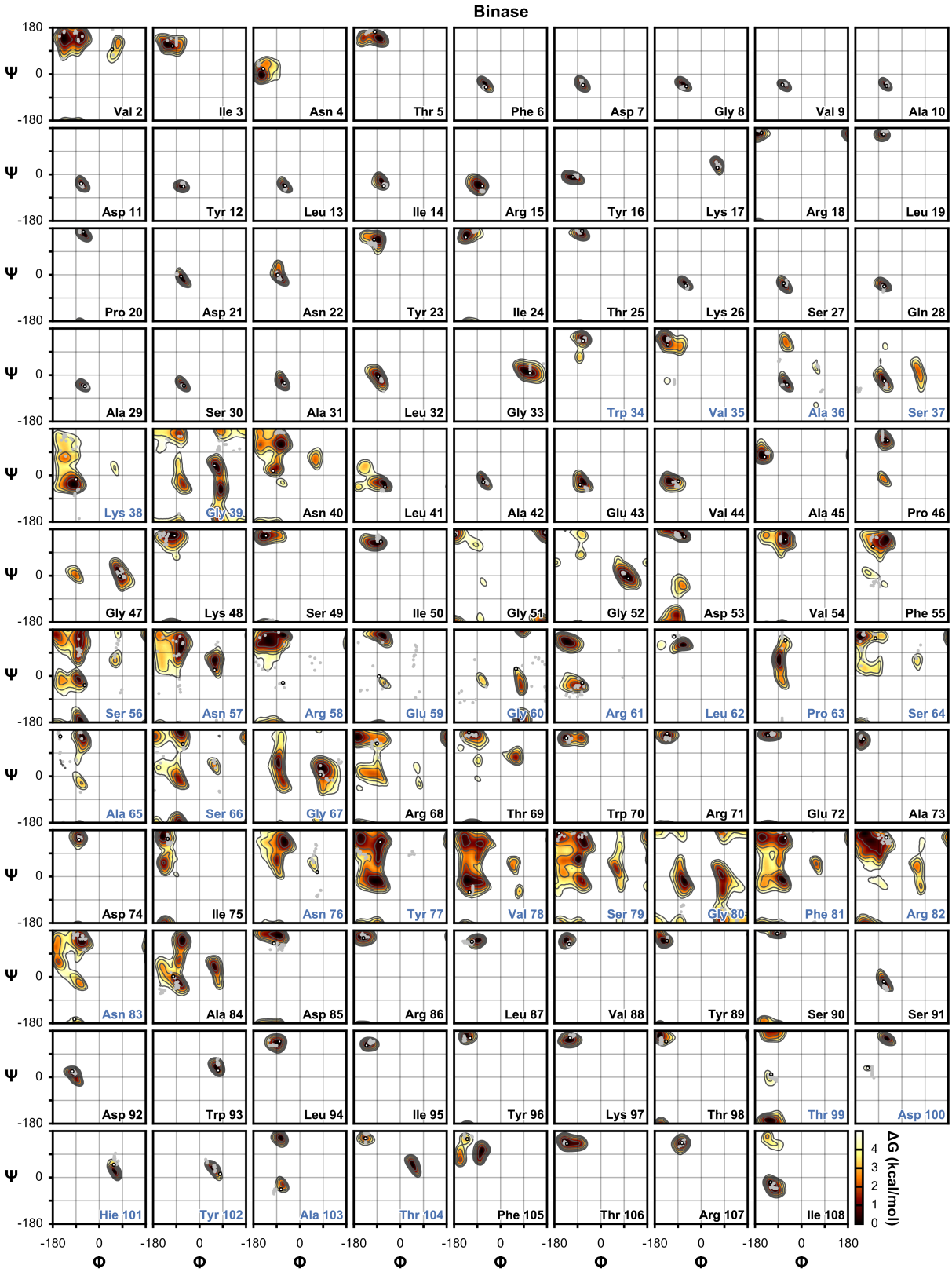

Figure S16. Backbone conformational sampling of binase observed in a 10-μs simulation. The Φ/Ψ angles observed in the twenty structures of the NMR ensemble (PDB code 1BUJ)[47] are shown as gray points, while those observed in the crystal structure (PDB code 1GOU)[48] are shown as white points. Overall retention of the crystal conformation of most residues is good; regions of deviation mentioned in the main text are highlighted in blue.

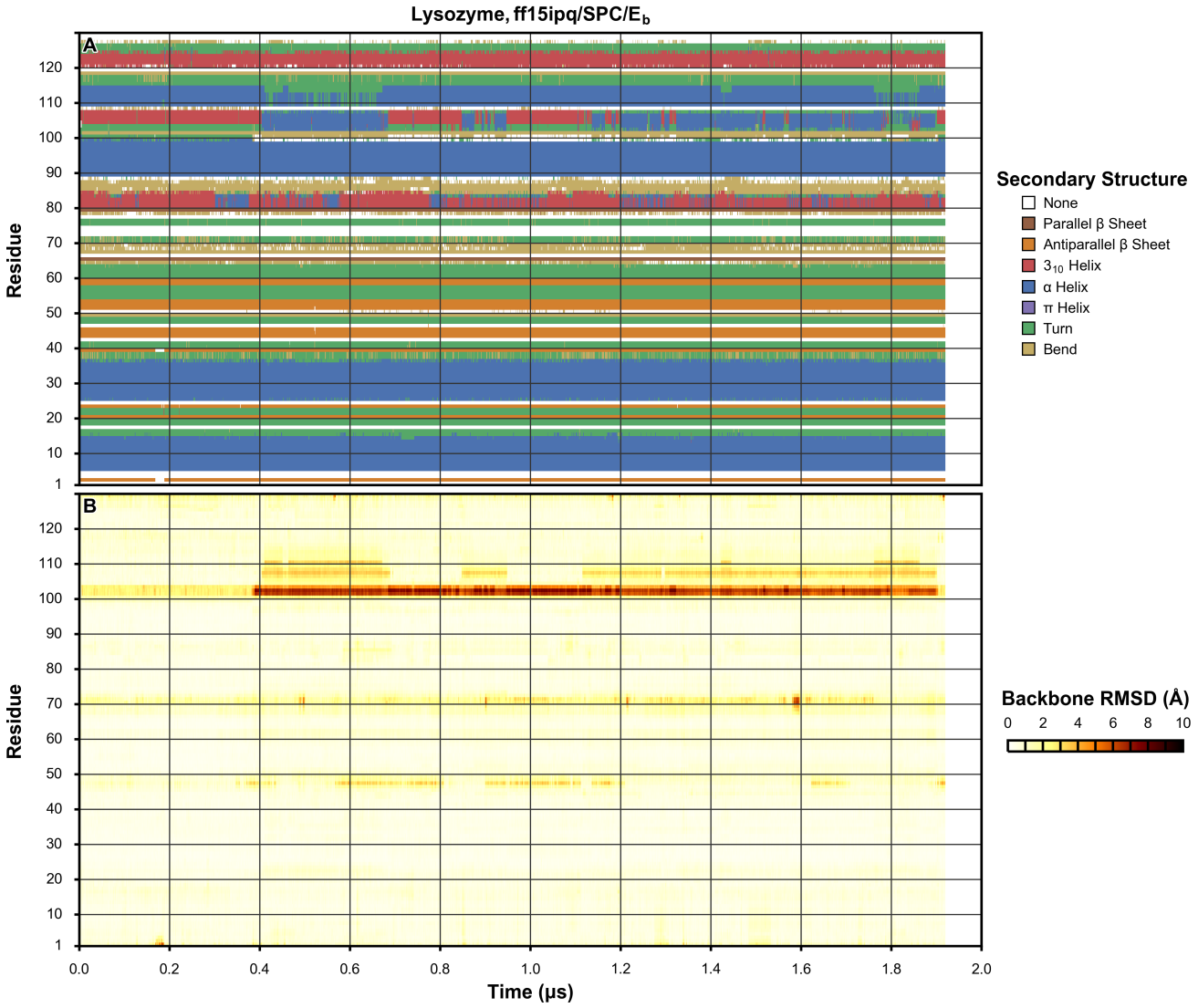

Figure S17. Secondary structure (**A**) and per-residue backbone RMSD relative to the crystal structure (PDB code 4LZT)[49] (**B**) of lysozyme observed over the course of a 2-μs simulation.

Figure S18. Backbone conformational sampling of lysozyme observed in a 10-μs simulation. The Φ/Ψ angles observed in the crystal structure (PDB code 4LZT)[49] are shown as gray points. Overall retention of the crystal conformation of most residues is good; regions of deviation mentioned in the main text are highlighted in blue.

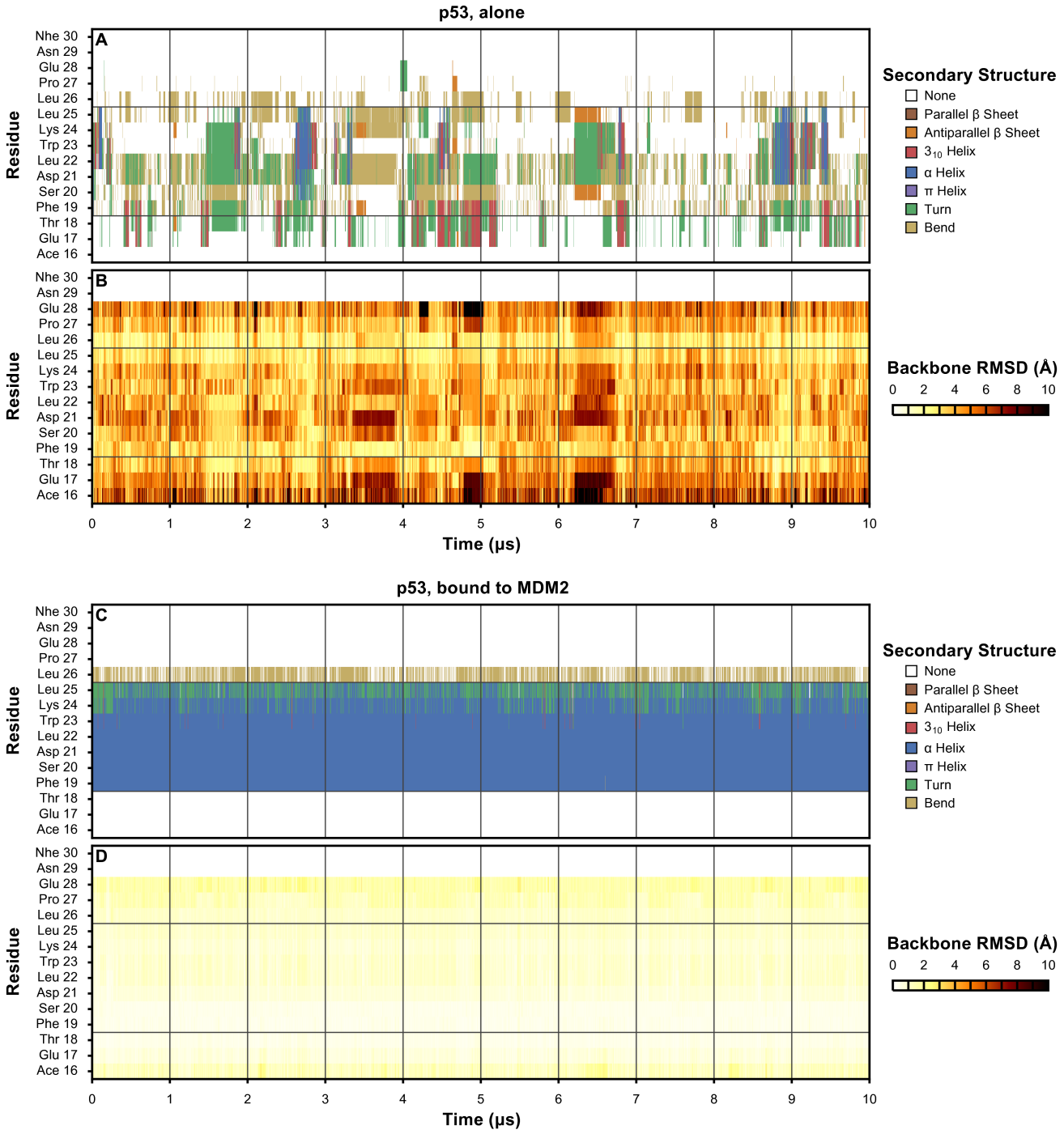

Figure S19. Secondary structure (**A, C**) and per-residue backbone RMSD relative to the crystal structure (PDB code 1YCR)[50] (**B, D**) of p53 observed over the course of 10-μs simulations alone (**A**, **B**) and in complex with MDM2 (**C**, **D**). Horizontal gridlines indicate the portion of p53 that forms an α-helix in the experimental structure of the p53/MDM2 complex.

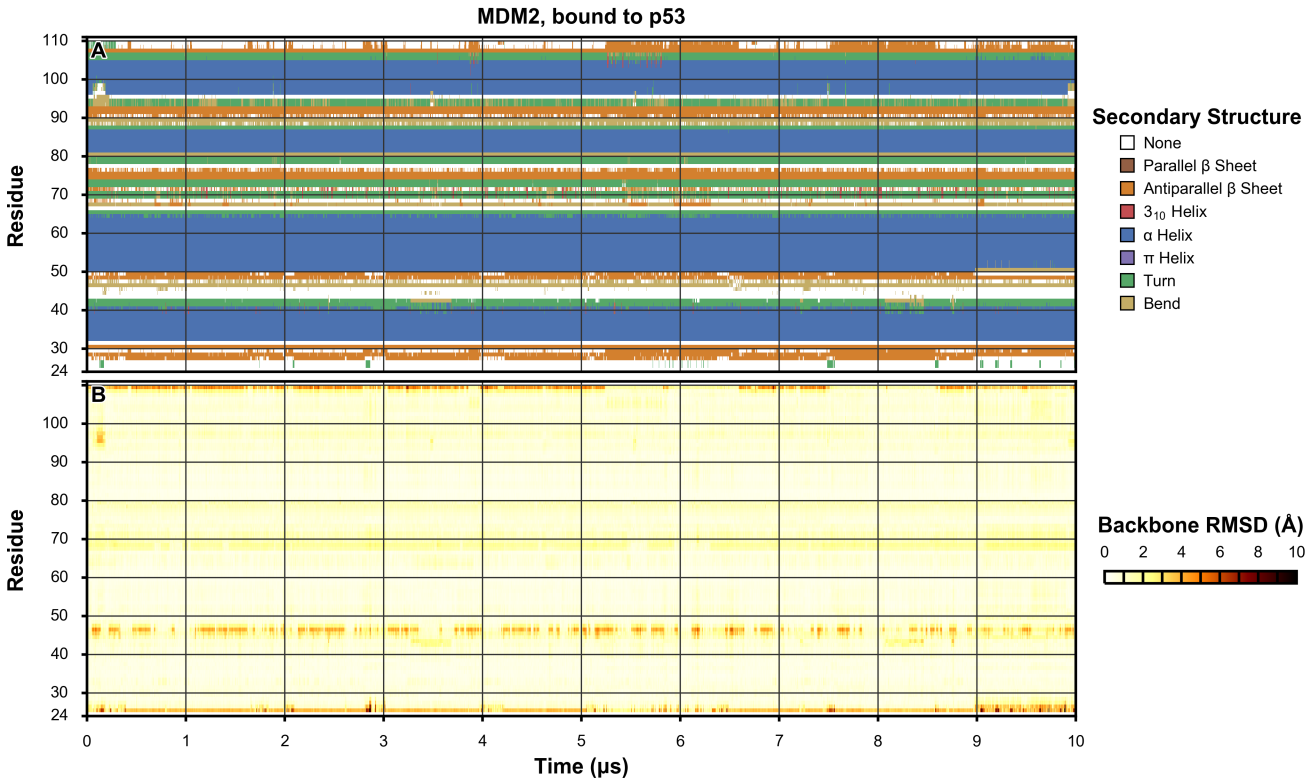

Figure S20. Secondary structure (**A**) and per-residue backbone RMSD relative to the crystal structure (PDB code 1YCR)[50] (**B**) of MDM2 observed over the course of a 10-μs simulation in complex with p53.

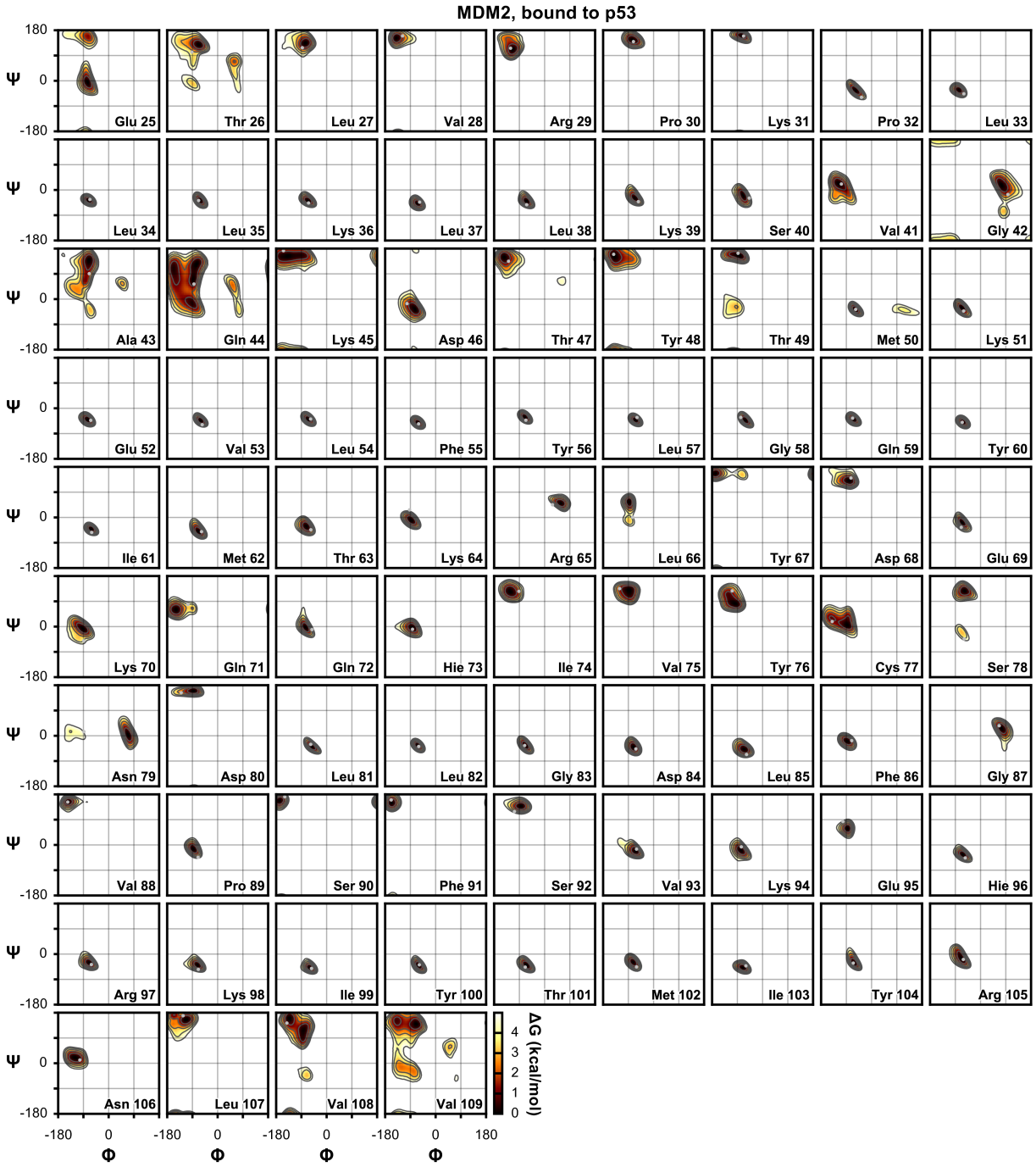

Figure S21. Backbone conformational sampling of MDM2 observed in a 10-μs simulation in complex with p53. The Φ/Ψ angles observed in the crystal structure (PDB code: 1YCR)[50] are shown as gray points.

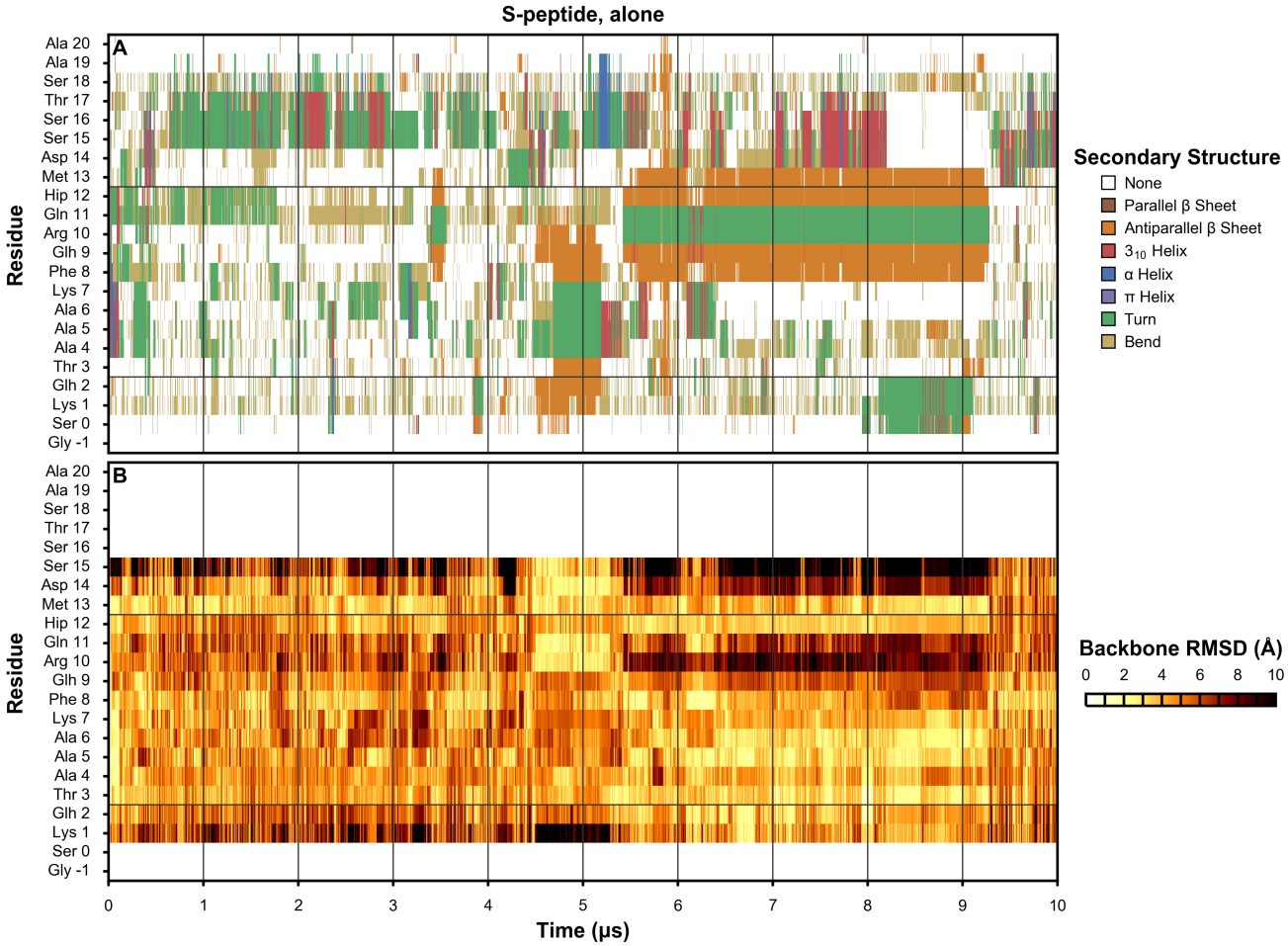

Figure S22. Secondary structure (**A**) and per-residue backbone RMSD relative to the crystal structure (PDB code 1RNU)[51] (**B**) of S-peptide observed over the course of a 10-μs simulation alone. Horizontal gridlines indicate the portion of S-peptide that forms an α-helix in the experimental structure of the S-peptide/S-protein complex.

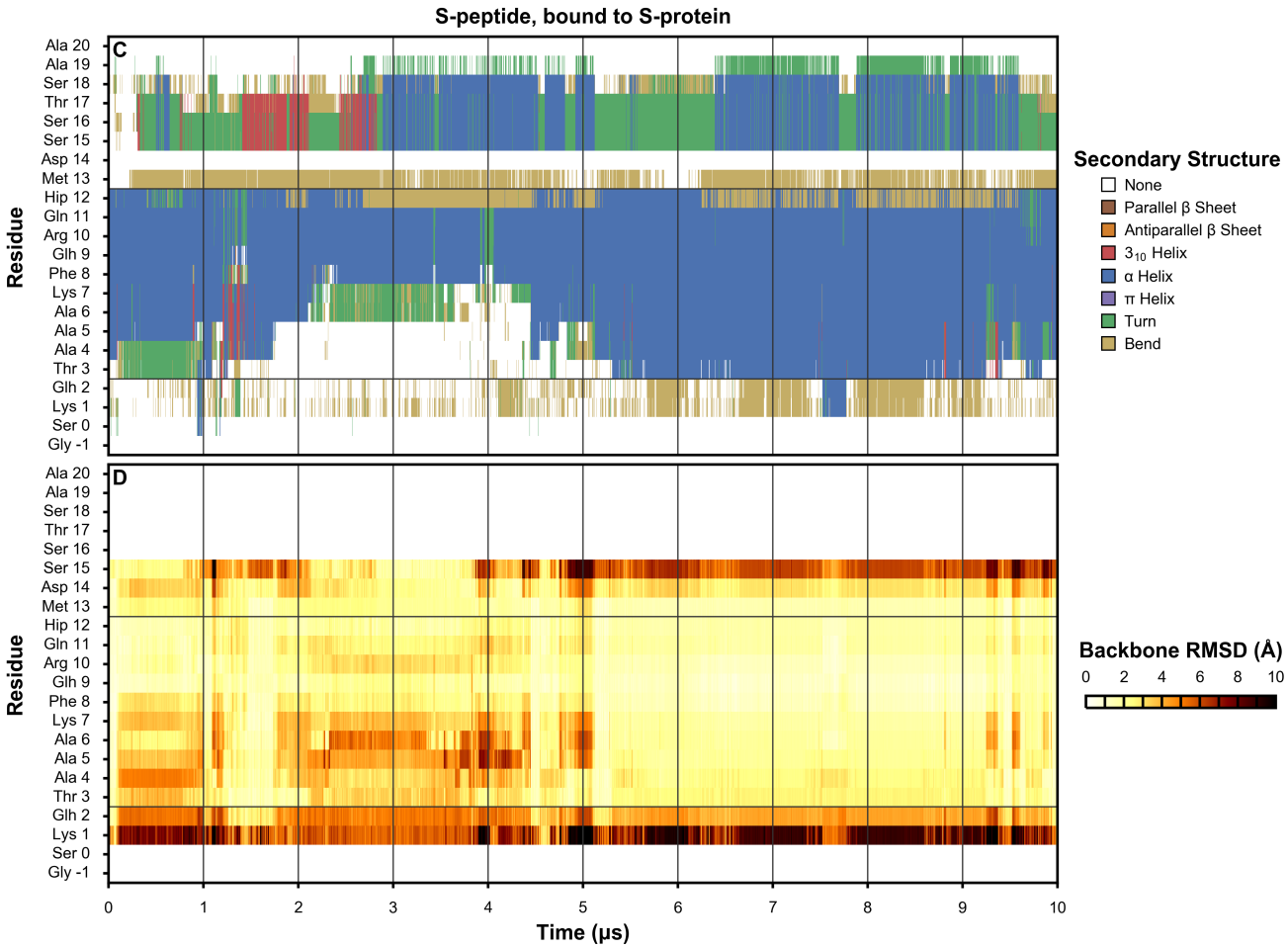

Figure S22 (Continued). Secondary structure (**C**) and per-residue backbone RMSD relative to the crystal structure (PDB code 1RNU)[51] (**D**) of S-peptide observed over the course of a 10-μs simulation in complex with S-protein. Horizontal gridlines indicate the portion of S-peptide that forms an α-helix in the experimental structure of the S-peptide/S-protein complex.

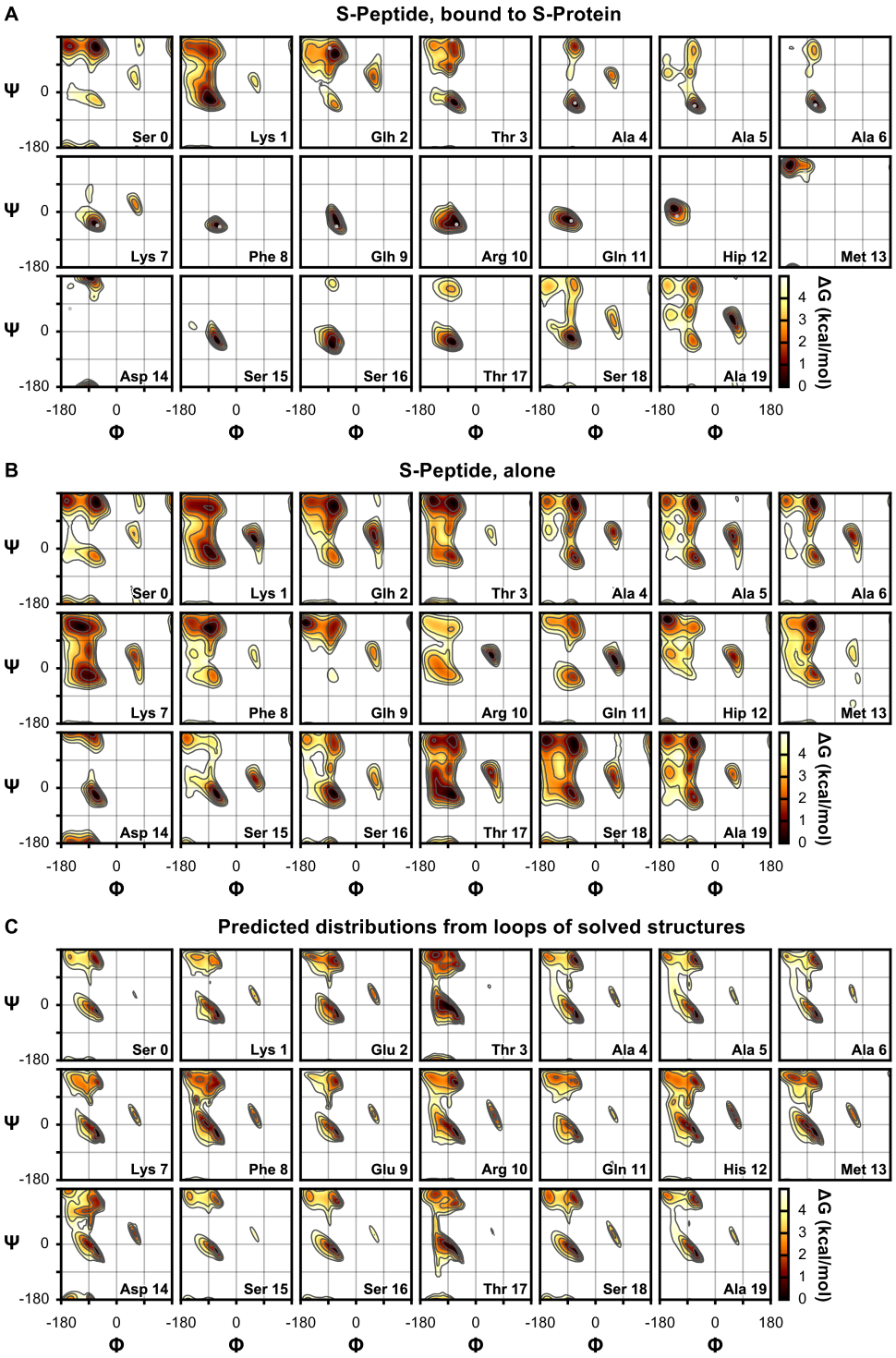

Figure S23. Backbone conformational sampling of S-Peptide observed in 10-μs simulations in complex with S-protein (**A**) and alone (**B**). The Φ/Ψ angles of residues present and resolved in the crystal structure of the S-peptide/S-protein complex (PDB code 1RNU)[51] are shown as gray points. For comparison are shown distributions for the S-peptide sequence obtained from the Neighbor-Dependent Ramachandran Distribution (NDRD) dataset, derived from conformations observed in the loops of solved structures (**C**).[38]

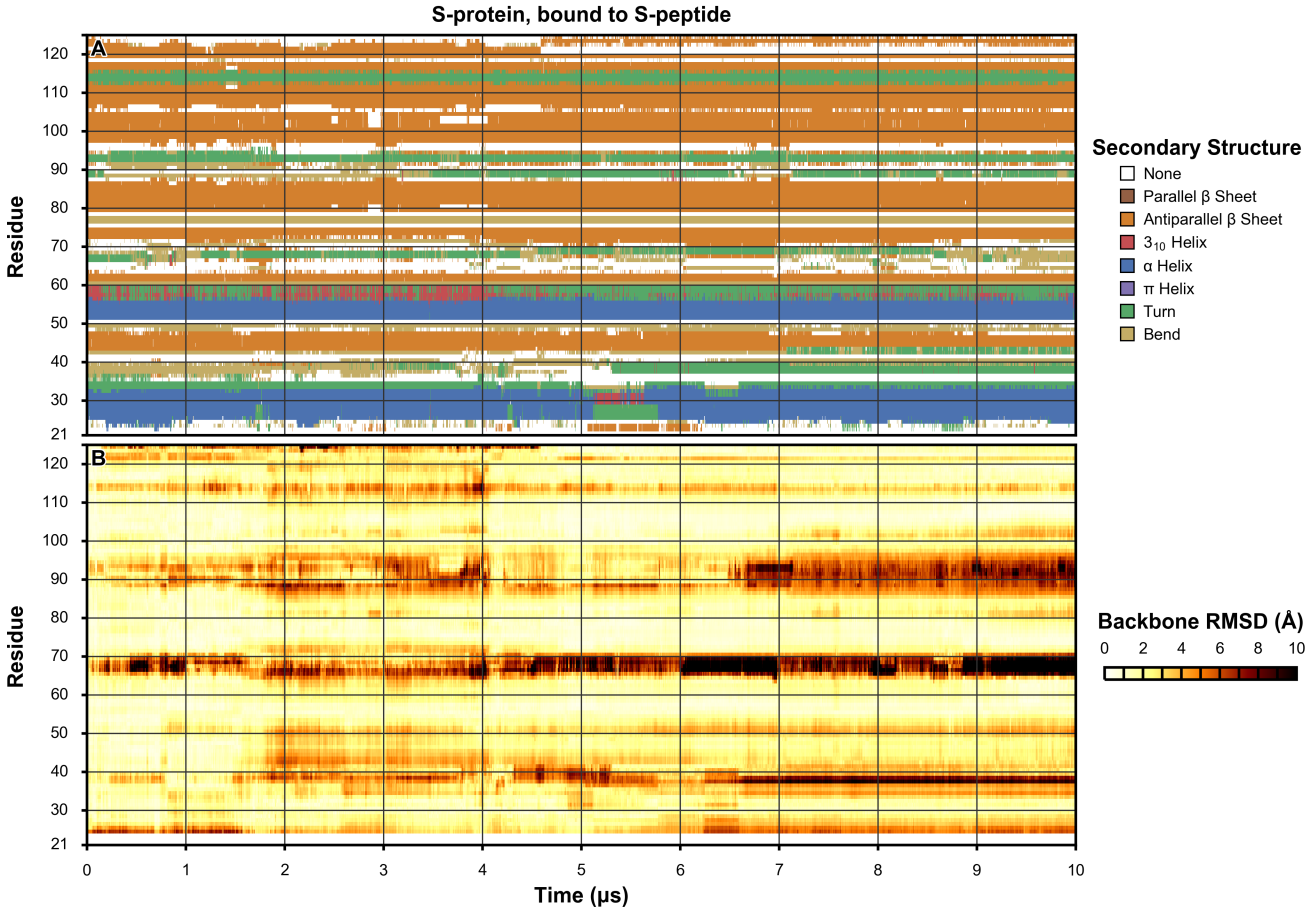

Figure S24. Secondary structure (**A**) and per-residue backbone RMSD relative to the crystal structure (PDB code 1RNU)[51] (**B**) of S-protein observed over the course of a 10-μs simulations in complex with S-peptide.

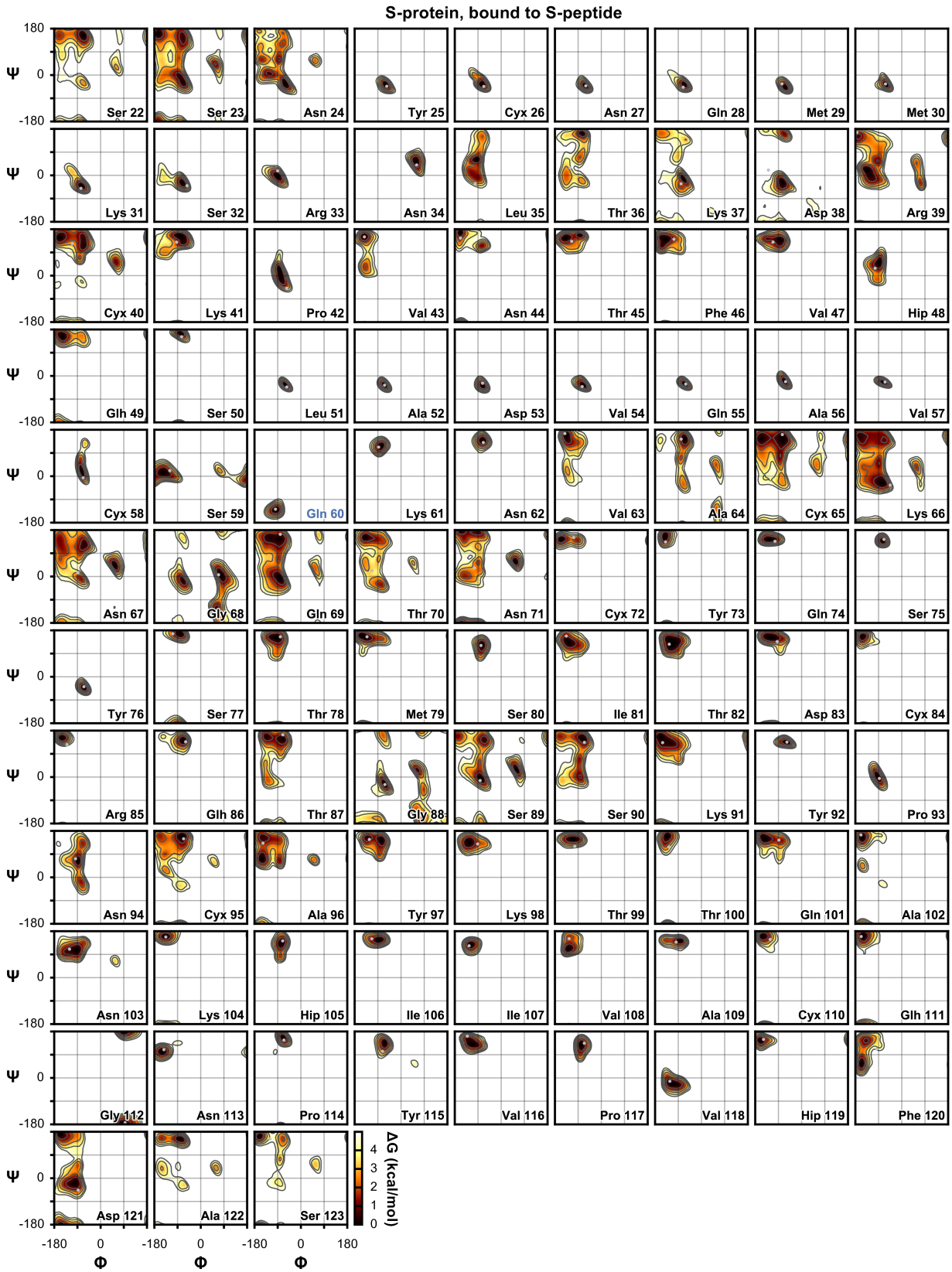

Figure S25. Backbone conformational sampling of S-protein observed in a 10-μs simulation in complex with S-peptide. The Φ/Ψ angles observed in the crystal structure (PDB code: 1RNU)[51] are shown as gray points. Overall retention of the crystal conformation of most residues is good; Gln 60, which retains its uncommon 'plateau' conformation throughout the simulation, is highlighted in blue.

Supporting Tables

Table S1. Backbone torsion classes and terms applied to heavy atoms for 28 nonterminal residue forms.

Residue	Φ/Ψ Class	Φ'/Ψ' Subclass	ϕ	ψ	ϕ'	ψ'
Ala	Neutral	Alanine	C-N-CX-C	N-CX-C-N	C-N-CX-CT	CT-CX-C-N
Ash	Neutral	Two-branched	C-N-CX-C	N-CX-C-N	C-N-CX-2C	2C-CX-C-N
Asn	Neutral	Two-branched	C-N-CX-C	N-CX-C-N	C-N-CX-2C	2C-CX-C-N
Cys	Neutral	Two-branched	C-N-CX-C	N-CX-C-N	C-N-CX-2C	2C-CX-C-N
Cyx	Neutral	Two-branched	C-N-CX-C	N-CX-C-N	C-N-CX-2C	2C-CX-C-N
Glh	Neutral	Two-branched	C-N-CX-C	N-CX-C-N	C-N-CX-2C	2C-CX-C-N
Gln	Neutral	Two-branched	C-N-CX-C	N-CX-C-N	C-N-CX-2C	2C-CX-C-N
Leu	Neutral	Two-branched	C-N-CX-C	N-CX-C-N	C-N-CX-2C	2C-CX-C-N
Lyn	Neutral	Two-branched	C-N-CX-C	N-CX-C-N	C-N-CX-2C	2C-CX-C-N
Met	Neutral	Two-branched	C-N-CX-C	N-CX-C-N	C-N-CX-2C	2C-CX-C-N
Nle	Neutral	Two-branched	C-N-CX-C	N-CX-C-N	C-N-CX-2C	2C-CX-C-N
Ser	Neutral	Two-branched	C-N-CX-C	N-CX-C-N	C-N-CX-2C	2C-CX-C-N
Ile	Neutral	Three-branched	C-N-CX-C	N-CX-C-N	C-N-CX-3C	C-N-CX-3C
Thr	Neutral	Three-branched	C-N-CX-C	N-CX-C-N	C-N-CX-3C	C-N-CX-3C
Val	Neutral	Three-branched	C-N-CX-C	N-CX-C-N	C-N-CX-3C	C-N-CX-3C
Hid	Neutral	Aromatic	C-N-CX-C	N-CX-C-N	C-N-CX-TA	TA-CX-C-N
Hie	Neutral	Aromatic	C-N-CX-C	N-CX-C-N	C-N-CX-TA	TA-CX-C-N
Phe	Neutral	Aromatic	C-N-CX-C	N-CX-C-N	C-N-CX-TA	TA-CX-C-N
Tyr	Neutral	Aromatic	C-N-CX-C	N-CX-C-N	C-N-CX-TA	TA-CX-C-N
Trp	Neutral	Aromatic	C-N-CX-C	N-CX-C-N	C-N-CX-TA	TA-CX-C-N
Asp	Negatively-Charged		C-N-TM-C	N-TM-C-N	C-N-TM-2C	2C-TM-C-N
Cym	Negatively-Charged		C-N-TM-C	N-TM-C-N	C-N-TM-2C	2C-TM-C-N
Glu	Negatively-Charged		C-N-TM-C	N-TM-C-N	C-N-TM-2C	2C-TM-C-N
Arg	Positively-Charged		C-N-TP-C	N-TP-C-N	C-N-TP-C8	C8-TP-C-N
Hip	Positively-Charged		C-N-TP-C	N-TP-C-N	C-N-TP-C8	C8-TP-C-N
Lys	Positively-Charged		C-N-TP-C	N-TP-C-N	C-N-TP-C8	C8-TP-C-N
Gly	Glycine		C-N-TG-C	N-TG-C-N		
Pro	Proline		C-TN-TJ-C	TN-TJ-C-N	C-TN-TJ-CT	CT-TJ-C-N

Supporting References

(1) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell Jr., A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone Φ, ψ and Side-Chain χ1 and χ2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.

(2) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.

(3) Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J. C.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: Greater Scalability for Parallel Molecular Dynamics. *J. Comput. Phys.* **1999**, *151*, 283–312.

(4) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29* (11), 1859–1865.

(5) Brooks, B. R.; Brooks, C. L.; MacKerell Jr., A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaeffer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30* (10), 1545–1614.

(6) Jiang, W.; Hardy, D. J.; Phillips, J. C.; MacKerell Jr., A. D.; Schulten, K.; Roux, B. High-Performance Scalable Molecular Dynamics Simulations of a Polarizable Force Field Based on Classical Drude Oscillators in NAMD. *J. Phys. Chem. Lett.* **2011**, *2*, 87–92.

(7) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J. Chem. Phys.* **1995**, *103* (11), 4613–4621.

(8) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (19), 8577–8593.

(9) Lopes, P. E. M.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell Jr., A. D. Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2013**, *9* (12), 5430–5449.

(10) Kräutler, V.; van Gunsteren, W. F.; Hünenberger, P. H. A Fast SHAKE Algorithm to Solve Distance Constraint Equations for Small Molecules in Molecular Dynamics Simulations. *J. Comput. Chem.* **2001**, *22* (5), 501–508.

(11) Ponder, J. W. TINKER 7.1.2. 2015.

(12) Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; Cheatham III, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Götz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P. A.; Kaus, J. W.; Kovalenko, A.; Lee, T.; Le Grand, S.; Li, P.; Luchko, T.; Luo, R.; Madej, B. D.; Merz, K. M.; Monard, G.; Needham, H.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A. E.; Salomon-Ferrer, R.; Simmerling, C.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; York, D. M.; Kollman, P. A. AMBER 2015. University of California, San Francisco 2015.

(13) Berendsen, H. J. C.; Postma, P. M.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81* (8), 3684–3690.

(14) Wu, J. C.; Chattree, G.; Ren, P. Automation of AMOEBA Polarizable Force Field Parameterization for Small Molecules. *Theor. Chem. Acc.* **2012**, *131*, 1138.

(15) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J.

E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V; Cioslowski, J.; Fox, D. J. Gaussian 09. Gaussian, Inc.: Wallingform CT, 2009.

(16) Stone, A. J.; Alderton, M. Distributed Multipole Analysis. *Mol. Phys.* **1985**, *56* (5), 1047–1064.

(17) Stone, A. J. Distributed Multipole Analysis: Stability for Large Basis Sets. *J. Chem. Theory Comput.* **2005**, *1* (6), 1128–1132.

(18) Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossváry, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In *Proceedings of the 2006 ACM/IEEE SCl06 Conference*; 2006; pp 1–13.

(19) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant Pressure Molecular Dynamics Algorithms. *J. Chem. Phys.* **1994**, *101* (5), 4177–4189.

(20) Shaw, D. E.; Deneroff, M. M.; Dror, R. O.; Kuskin, J. S.; Larson, R. H.; Salmon, J. K.; Young, C.; Batson, B.; Bowers, K. J.; Chao, J. C.; Eastwood, M. P.; Gagliardo, J.; Grossman, J. P.; Ho, H. R.; Ierardi, D. J.; Kolossváry, I.; Klepeis, J. L.; Layman, T.; McLeavey, C.; Moraes, M. A.; Mueller, R.; Priest, E. C.; Shan, Y.; Spengler, J.; Theobald, M.; Towles, B.; Wang, S. C. Anton, a Special-Purpose Machine for Molecular Dynamics Simulation. *Commun. ACM* **2008**, *51* (7), 91–97.

(21) Lippert, R. A.; Predescu, C.; Ierardi, D. J.; Mackenzie, K. M.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. Accurate and Efficient Integration for Molecular Dynamics Simulations at Constant Temperature and Pressure. *J. Chem. Phys.* **2013**, *139*, 1–11.

(22) Shan, Y.; Klepeis, J. L.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. Gaussian Split Ewald: A Fast Ewald Mesh Method for Molecular Simulation. *J. Chem. Phys.* **2005**, *122* (5), 1–13.

(23) 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* **2010**, *78* (8), 1950–1958.

(24) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. How Robust Are Protein Folding Simulations with Respect to Force Field Parameterization? *Biophys. J.* **2011**, *100* (9), L47–L49.

(25) Takemura, K.; Kitao, A. Water Model Tuning for Improved Reproduction of Rotational Diffusion and NMR Spectral Density. *J. Phys. Chem. B* **2012**, *116*, 6279–6287.

(26) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926–935.

(27) Piana, S.; Donchev, A. G.; Robustelli, P.; Shaw, D. E. Water Dispersion Interactions Strongly Influence Simulated Structural Properties of Disordered Protein States. *J. Phys. Chem. B* **2015**, *119*, 5113–5123.

(28) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120* (20), 9665–9678.

(29) Hall, J. B.; Fushman, D. Variability of the 15N Chemical Shielding Tensors in the B3 Domain of Protein G from 15N Relaxation Measurements at Several Fields. Implications for Backbone Order Parameters. *J. Am. Chem. Soc.* **2006**, *128* (24), 7855–7870.

(30) Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A. Rotational Diffusion Anisotropy of Human Ubiquitin from 15N NMR Relaxation. *J. Am. Chem. Soc.* **1995**, *117* (50), 12562–12566.

(31) Pang, Y.; Buck, M.; Zuiderweg, E. R. P. Backbone Dynamics of the Ribonuclease Binase Active Site Area Using Multinuclear (15N and 13CO) NMR Relaxation and Computational Molecular Dynamics. *Biochemistry* **2002**, *41* (8), 2655–2666.

(32) Wong, V.; Case, D. A. Evaluating Rotational Diffusion from Protein MD Simulations. *J. Phys. Chem. B* **2008**, *112*, 6013–6024.

(33) Cerutti, D. S.; Rice, J. E.; Swope, W. C.; Case, D. A. Derivation of Fixed Partial Charges for

Amino Acids Accommodating a Specific Water Model and Implicit Polarization. *J. Phys. Chem. B* **2013**, *117*, 2328–2338.
(34) Cerutti, D. S.; Swope, W. C.; Rice, J. E.; Case, D. A. ff14ipq: A Self-Consistent Force Field for Condensed-Phase Simulations of Proteins. *J. Chem. Theory Comput.* **2014**, *10*, 4515–4534.
(35) Tanford, C. The Association of Acetate with Ammonium and Guanidinium Ions. *J. Am. Chem. Soc.* **1954**, *76*, 945–946.
(36) Springs, B.; Haake, P. Equilibrium Constants for Association of Guanidinium and Ammonium Ions with Oxyanions. *Bioorg. Chem.* **1977**, *6* (2), 181–190.
(37) Flyvbjerg, H.; Petersen, H. G. Error Estimates on Averages of Correlated Data. *J. Chem. Phys.* **1989**, *91* (1), 461–466.
(38) Ting, D.; Wang, G.; Shapovalov, M. V; Mitra, R.; Jordan, M. I.; Dunbrack Jr., R. L. Neighbor-Dependent Ramachandran Probability Distributions of Amino Acids Developed from a Hierarchical Dirichlet Process Model. *PLoS Comput. Biol.* **2010**, *6* (4), e1000763.
(39) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
(40) Wlodawer, A.; Walter, J.; Huber, R.; Slölin, L. Structure of Bovine Pancreatic Trypsin Inhibitor. Results of Joint Neutron and X-Ray Refinement of Crystal Form II. *J. Mol. Biol.* **1984**, *180*, 301–329.
(41) Kubelka, J.; Chiu, T. K.; Davies, D. R.; Eaton, W. A.; Hofrichter, J. Sub-Microsecond Protein Folding. *J. Mol. Biol.* **2006**, *359* (3), 546–553.
(42) Ulmer, T. S.; Ramirez, B. E.; Delaglio, F.; Bax, A. Evaluation of Backbone Proton Positions and Dynamics in a Small Protein by Liquid Crystal NMR Spectroscopy. *J. Am. Chem. Soc.* **2003**, *125* (30), 9179–9191.
(43) Prompers, J. J.; Brüschweiler, R. General Framework for Studying the Dynamics of Folded and Nonfolded Proteins by NMR Relaxation Spectroscopy and MD Simulation. *J. Am. Chem. Soc.* **2002**, *124* (16), 4522–4534.
(44) Gu, Y.; Li, D.-W.; Brüschweiler, R. NMR Order Parameter Determination from Long Molecular Dynamics Trajectories for Objective Comparison with Experiment. *J. Chem. Theory Comput.* **2014**, *10*, 2599–2607.
(45) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. Structure of Ubiquitin Refined at 1.8 Å Resolution. *J. Mol. Biol.* **1987**, *194* (3), 531–544.
(46) Lee, A. L.; Wand, A. J. Assessing Potential Bias in the Determination of Rotational Correlation Times of Proteins by NMR Relaxation. *J. Biomol. NMR* **1999**, *13* (2), 101–112.
(47) Reibarkh, M. Y.; Nolde, D. E.; I, V. L.; Bocharov, A. A.; Shulga, A. A.; Kirpichnikov, M. P.; Arseniev, A. S. Three-Dimensional Structure of Binase in Solution. *FEBS Lett.* **1998**, *431*, 250–254.
(48) Polyakov, K. M.; Lebedev, A. A.; Okorokov, A. L.; Panov, K. I.; Schulga, A. A.; Pavlovsky, A. G.; Karpeisky, M. Y.; Dodson, G. G. The Structure of Substrate-Free Microbial Ribonuclease Binase and of Its Complexes with 3′GMP and Sulfate Ions. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2002**, *58* (5), 744–750.
(49) Walsh, M. A.; Schneider, T. R.; Sieker, L. C.; Dauter, Z.; Lamzin, S. Refinement of Triclinic Hen Egg-White Lysozyme at Atomic Resolution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1998**, *54*, 522–546.
(50) Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P. Structure of the MDM2 Oncoprotein Bound to the p53 Tumor Suppressor Transactivation Domain. *Science* **1996**, *274*, 948–953.
(51) Kim, E. E.; Varadarajan, R.; Wyckoff, H. W.; Richards, F. M. Refinement of the Crystal Structure of Ribonuclease S. Comparison with and between the Various Ribonuclease A Structures. *Biochemistry* **1992**, *31* (49), 12304–12314.