Supporting Information

For

Induction of peptide bond dipoles drives cooperative helix formation in the $(AAQAA)_3$ peptide

Jing Huang and Alexander D. MacKerell Jr.*

Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD 21201

* Corresponding author

Mailing Address: 20 Penn Street, Room 629, Baltimore, MD 21201

Email: alex@outerbanks.umaryland.edu

Phone: (410) 706-7442

Fax: (410) 706-5017

Backbone peptide dipole analysis

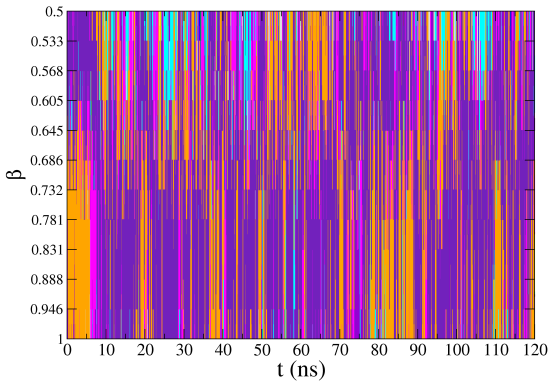
Peptide backbone dipoles are computed as the ensemble average from the simulations for the different conformational ranges. The peptide bond for the *i*-th residue is defined as the CO group in the *i-1* and N, H, Cα, Hα atoms in the *i*-th residues, yielding a charge neutral group in both force fields. For the first residue the carbonyl group in the N-terminal acetyl group is used.

Intrinsic dipoles moments for backbone peptide bond groups are obtained for the alanine tripeptide in the gas phase. These calculations were performed by obtaining all the ϕ, ψ values from the full $(AAQAA)_3$ HREMD simulation at 300K (ie. 90,000 x 15 ϕ, ψ sets), then restraining all the ϕ, ψ dihedral angles in the alanine tripeptide to these particular values using a harmonic restraining potential of 10000 kcal/mol/rad^2, performing 2000 steps of Adopted Basis Newton-Raphson (ABNR) minimization followed by 2000 steps of Steepest Descent (SD) minimization with a stepsize of 0.00001 to relax all the remaining degrees of freedom, and obtaining the dipole moment for the second peptide bond. Averages were then obtained over the different conformational regions in a manner identical to that performed for $(AAQAA)_3$. For the gas-phase alanine tripeptide the average dipole moments are reported in the Table 2 in the main text as "Intrinsic," with the difference between the total and intrinsic dipoles yielding the "Enhancement" of the peptide dipoles due to the full environment comprised of the remainder of the peptide and the surrounding solvent. The contributions to the dipoles was further separated into the induction effects of peptide-water and intrapeptide interactions by computing the "intramolecular enhancement" of dipole moments. For each frame from the full solvated $(AAQAA)_3$ HREMD simulation, we computed the peptide dipole moments after removing all water molecules and then relaxing the Drude particles in a SCF manner. The computed dipoles

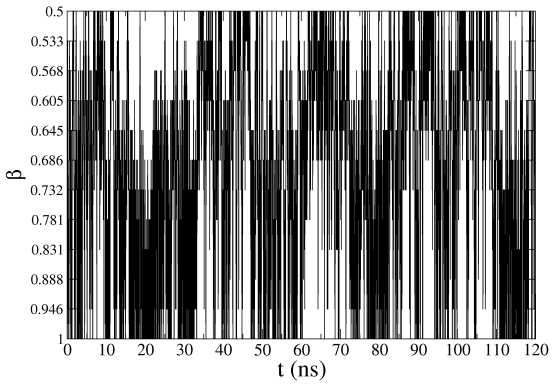
are ensemble averaged for the different conformational ranges, yielding the "intramolecular enhancement" values reported in the Table 2 in the main text.

Figure S1. Time series of the β values adopted in the H-REMD simulation with the Drude force field at 300K for A) all replicas; B) replica starting with β=1 (the 0th replica); C) replica starting with β=0.732 (the 5th replica) and D) replica starting with β=0.5 (the 11th replica).

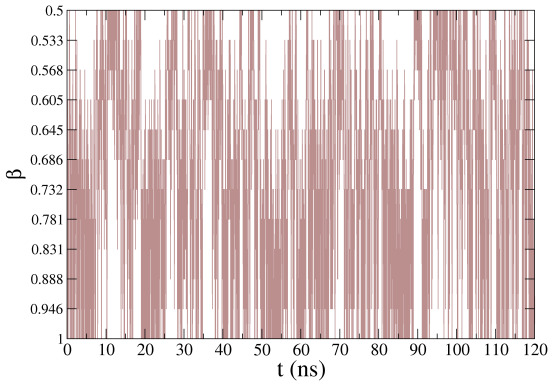
A



B



C



D

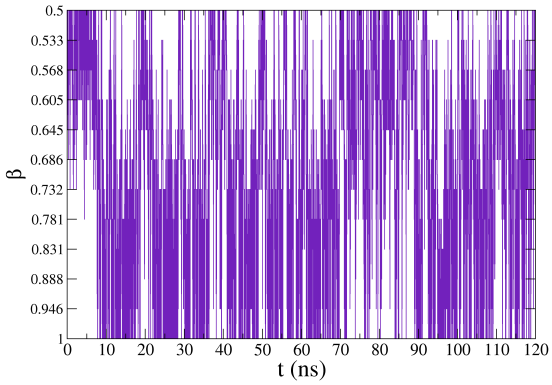
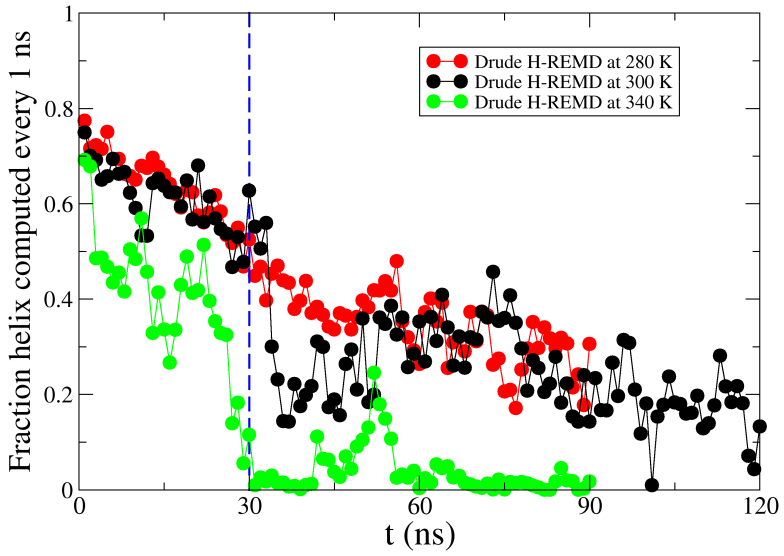
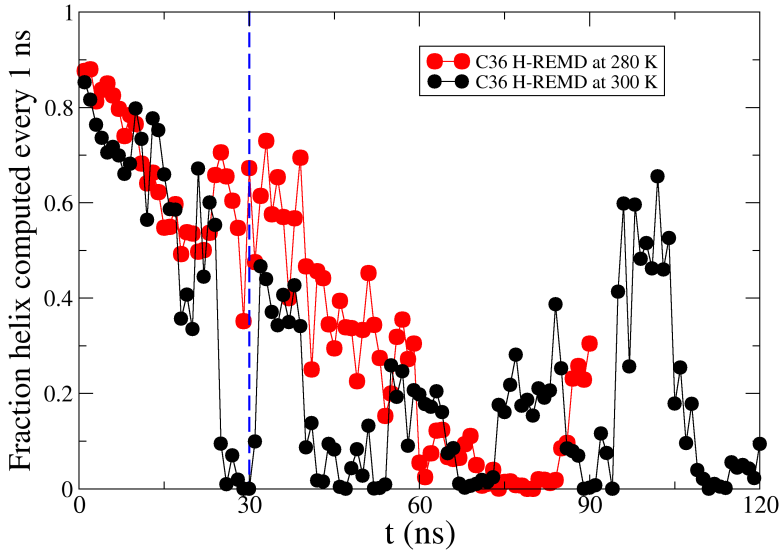


Figure S2. Average fraction helix computed over 1 ns blocks for the Drude and CHARMM36 H-REMD simulation at 280, 300 and 340 K. The blue broken lines indicate the simulation time before which is considered equilibrium and not included in the analysis. For the C36 simulation at 340 K, the initial H-REMD was carried out using 4 replicas (with biasing parameter β=1, 0.919, 0.842 and 0.773) for 60 ns, and then switched to H-REMD simulations using 12 replicas with β scaled exponentially from 1 to 0.5, the same replica exchange setup as used in the other simulations. H-REMD with 12 replicas was run for another 60 ns, resulting in the same trajectory length for analysis as the Drude simulation at 340K.




C36 H-REMD at 280 K
C36 H-REMD at 300 K
Fraction helix computed every 1 ns
t (ns)
0
0.2
0.4
0.6
0.8
1
0
30
60
90
120

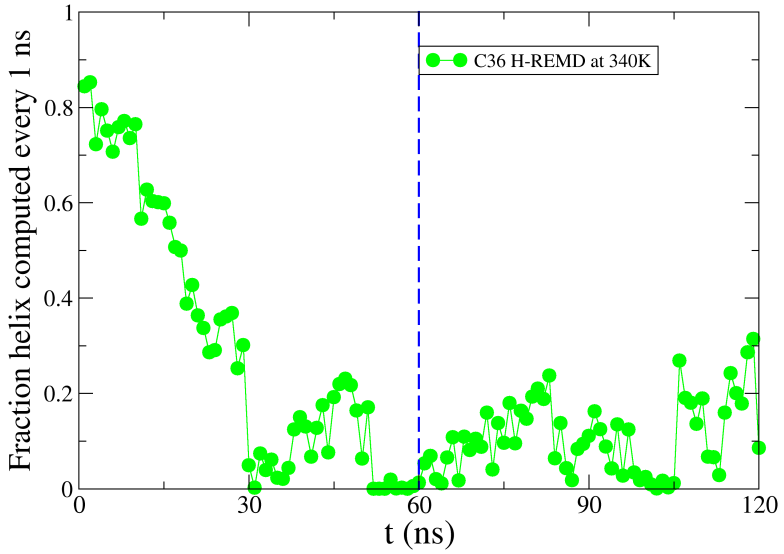

C36 H-REMD at 340K
Fraction helix computed every 1 ns
t (ns)
0
0.2
0.4
0.6
0.8
1
0
30
60
90
120

Figure S3. Definition of different conformations on the Ramachandran map, overlapped with the ϕ,ψ distribution of the top500 pdb database from Lovell *et al*. (1) The conformations are defined as follows: α region: -100° < ϕ < -30° and -67° < ψ < -7°; α_+ region: -160° < ϕ < -120° and -120° < ψ < -50°; β region: -180° < ϕ < -90° and 50° < ψ < 180°, or -180° < ϕ < -90° and -180° < ψ < -120°, or 160° < ϕ < 180° and 110° < ψ < 180°; ppII: -90° < ϕ < -20° and 50° < ψ < 180°, or -90° < ϕ < -20° and -180° < ψ < -120°.

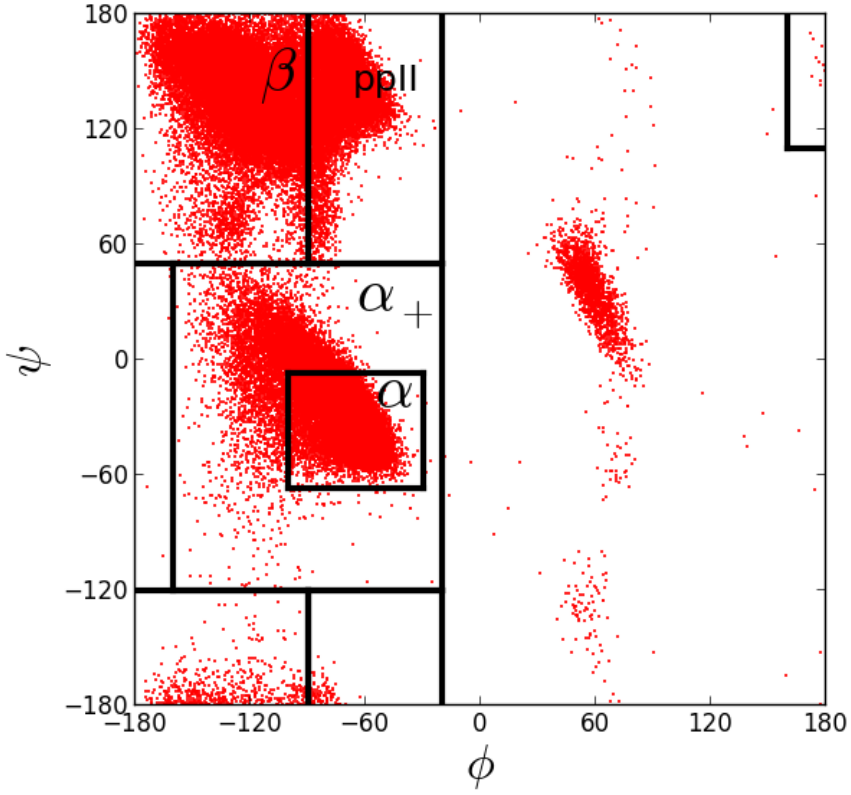


Figure S4. Probability density distribution of radius of gyration of the $(AAQAA)_3$ peptide computed for conformations from the 300K simulations with the Drude and C36 force fields.

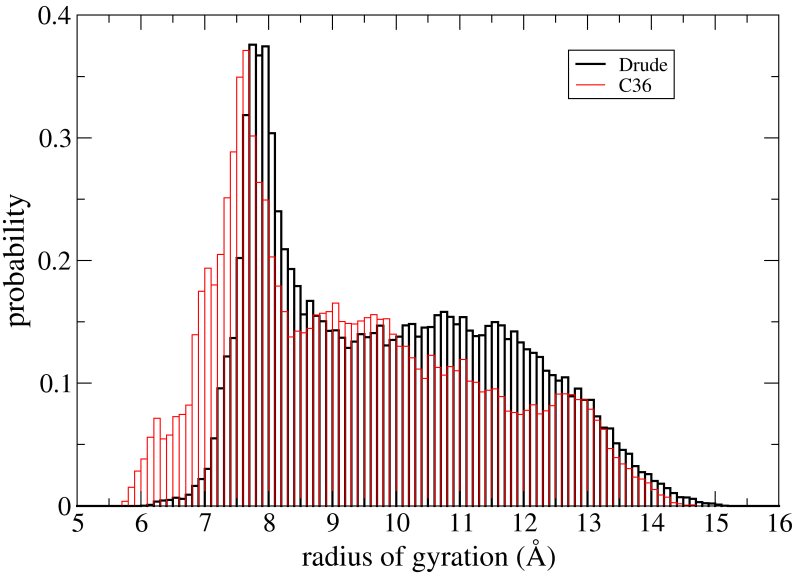


Table S1. Population of selected conformational regions with the Drude and CHARMM36 force fields at 280K, 300K and 340K.

280K	**%α_+**	**%β**	**%PPII**	**%α**	**%α-helix**	**α/α_+**	**α-helix/α**
Drude	43.5	38.6	16.0	40.7	34.6	0.94	0.85
CHARMM36	42.6	14.5	37.5	34.4	25.6	0.81	0.74

300K	**%α_+**	**%β**	**%PPII**	**%α**	**%α-helix**	**α/α_+**	**α-helix/α**
Drude	36.3	44.9	17.0	33.0	25.1	0.91	0.76
CHARMM36	40.2	20.7	31.5	30.0	19.8	0.75	0.66

340K	**%α_+**	**%β**	**%PPII**	**%α**	**%α-helix**	**α/α_+**	**α-helix/α**
Drude	18.9	57.8	20.5	14.1	3.8	0.75	0.27
CHARMM36	35.0	25.0	30.5	22.8	10.9	0.65	0.48

Supporting References

1. Lovell, S. C., I. W. Davis, W. B. Arendall, P. I. W. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, and D. C. Richardson. 2003. Structure validation by C α geometry: φ, ψ and C β deviation. Proteins: Struct., Funct., Bioinf. 50:437-450.