Supporting Information

for

Recent Advances in Polarizable Force Fields for Macromolecules: Microsecond Simulations of Proteins Using the Classical Drude Oscillator Model

Jing Huang,[1] Pedro E.M. Lopes,[1] Benoît Roux[2] and Alexander D. MacKerell, Jr.[1,]*

[1] Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD 21201, USA

[2] Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637, USA

* 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

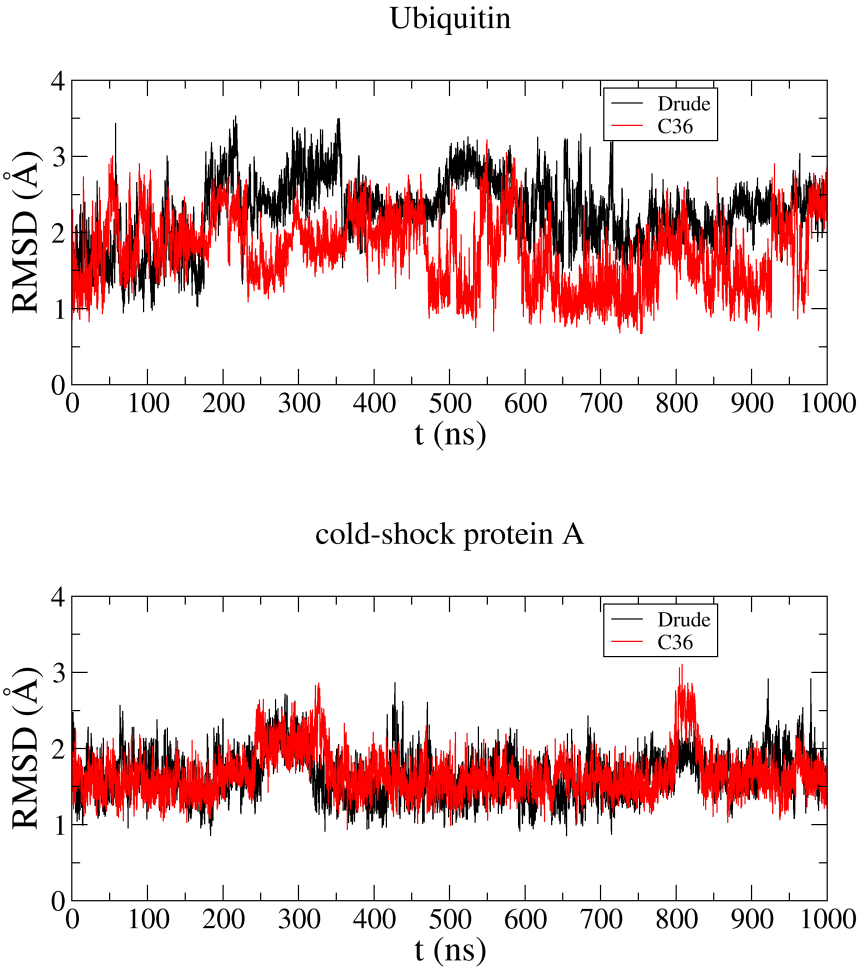
Simulation details.

Protein systems studied in this work are summarized in Table S1. Results from the first 100 ns of the simulations were presented previously.[1] MD simulations were performed in the NPT ensemble at 300 K and 1 atm using NAMD.[2] Periodic boundary conditions were applied and Lennard-Jones interactions were truncated at 12 Å with a smoothing function (force switch smoothing for the CHARMM36 simulations and switch smoothing for the Drude) from 10 to 12 Å. The nonbonded interaction lists were generated with a distance cutoff of 16 Å and updated heuristically. Electrostatic interactions were calculated using the particle mesh Ewald method[3] with a real space cutoff of 12 Å on an approximately 1 Å grid with a sixth-order spline. All covalent bonds involving hydrogen as well as the intramolecular geometries of water were constrained using SETTLE. The integration time step equals 1 fs for the Drude and 2 fs for the CHARMM36 simulations, and coordinates were saved every 0.1 ns.

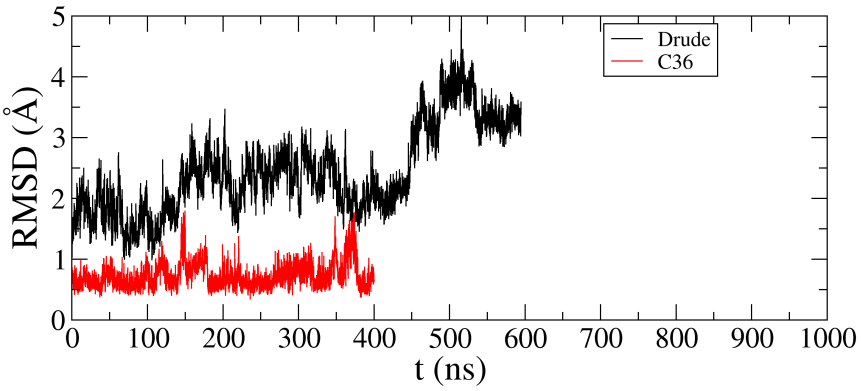
Table S1. Details of the MD simulations presented in this study. Given in the parenthesis is the trajectory length included in the protein dielectric analysis.

Protein	PDB ID	Box size (Å)	ions	Simulation length
Cold-shock protein A, 69 aa		52.4	Na^+	1000 ns with Drude FF (1000 ns)
	1mjc			1000 ns with C36 FF (1000 ns)
Ubiquitin, 76 aa		58.4	--	1000 ns with Drude FF (1000 ns)
	1ubq			1000 ns with C36 FF (1000 ns)
Crambin, 46 aa		52.0	--	1000 ns with Drude FF (1000 ns)
	1ejg			250 ns with C36 FF (120 ns)
Protein GB1 domain, 56 aa		56.7	Na^+	600 ns with Drude FF (500 ns)
	1p7e			400 ns with Drude FF (120 ns)
Circular permutant of ribosomal proteion S6, 74 aa				950 ns with Drude FF (500 ns)
	3zzp	61.0	Na^+	180 ns with C36 FF (120 ns)
DNA methyltransferase associated protein (DMAP1), 75 aa				1000 ns with Drude FF (500 ns)
	4iej	59.0	--	200 ns with C36 FF (120 ns)
PDZ domain from tight junction regulatory protein, 85 aa				1000 ns with Drude FF (500 ns)
	3vqf	58.0	Na^+	200 ns with C36 FF (120 ns)

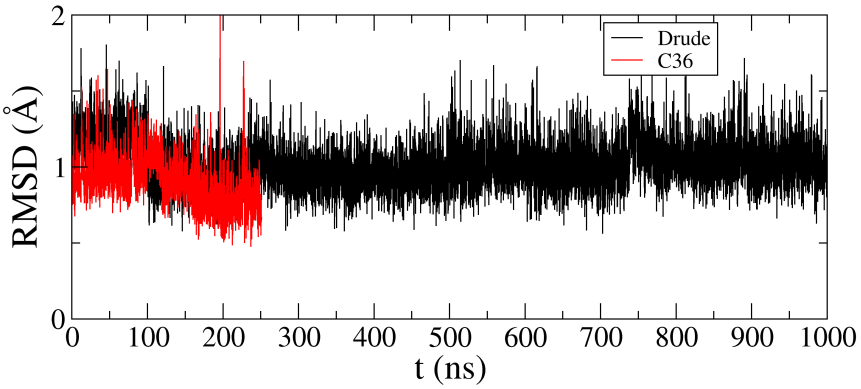
Figure S1. Protein Ca RMSD plots from MD simulations with the Drude and C36 simulations. All residues are included in RMSD calculations. The MD simulations of DNA methyltransferase associated protein (4iej) were run without Calcium ions that are important to its stability, which may contribute to its unfolding after hundred of nanosecond simulations with the Drude force field.



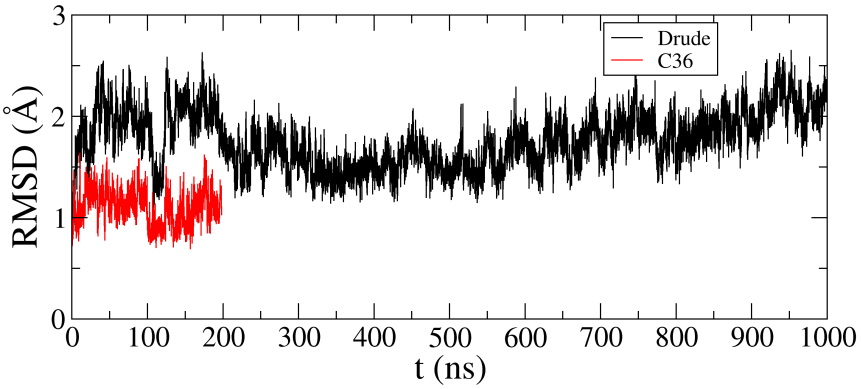
protein G B1 domain



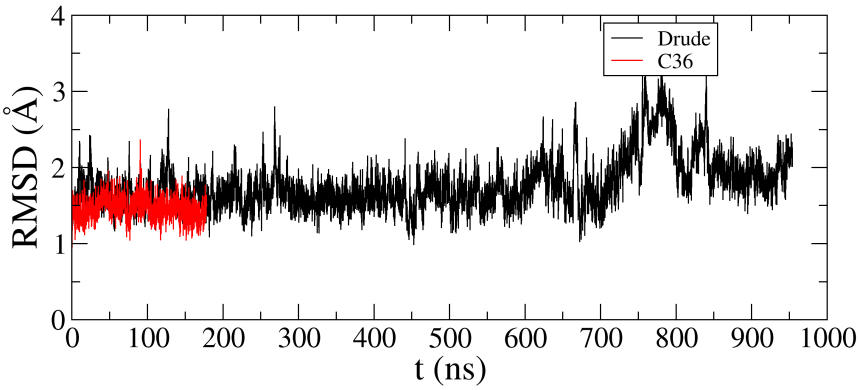
crambin



PDZ domain from tight junction regulatory protein



Circular permutant of ribosomal protein S6



DNA methyltransferase associated protein

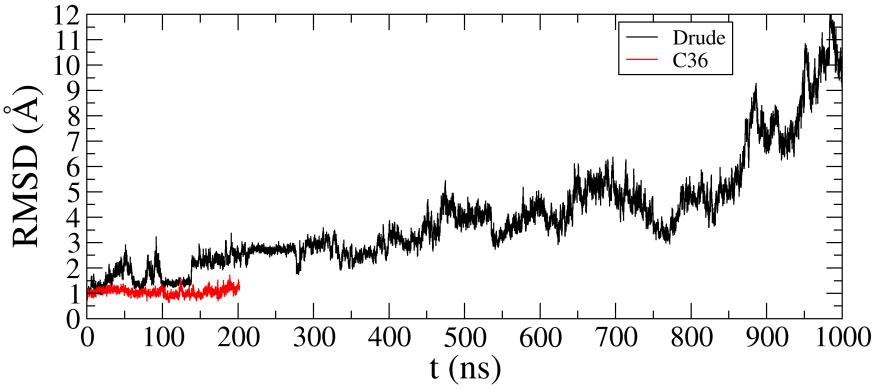
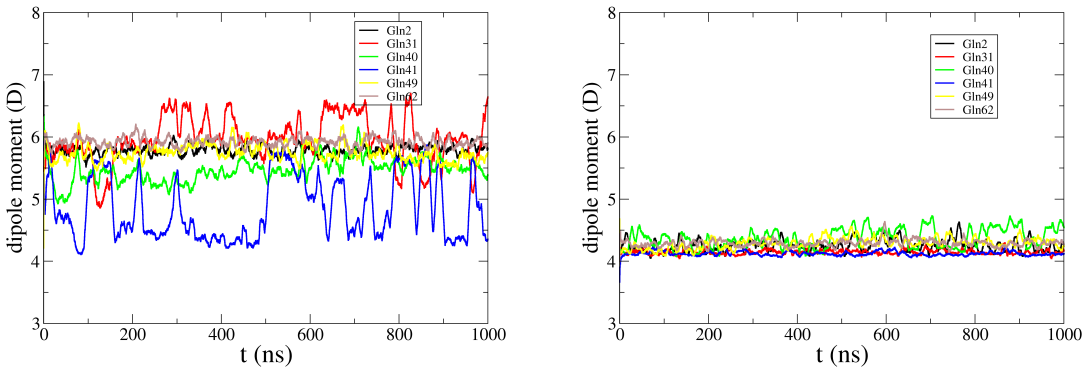
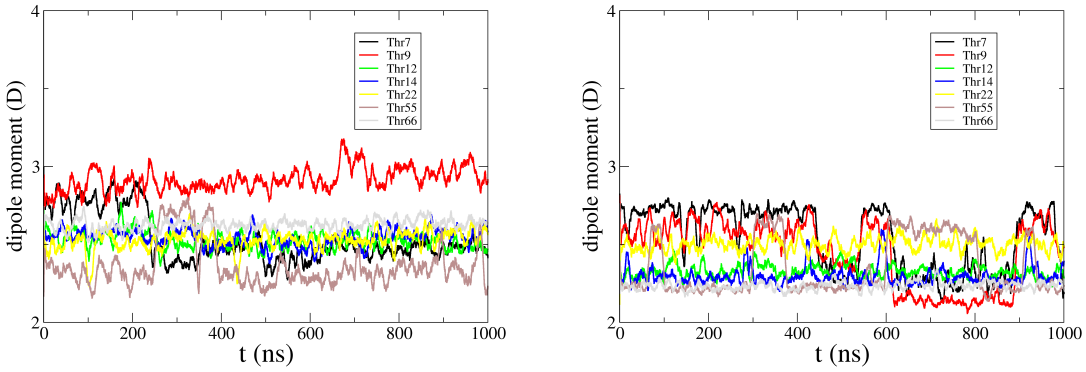


Figure S2. Side chain dipole moments from the 1 μs MD simulations of ubiquitin with the polarizable Drude force field (left panels) and the additive CHARMM36 force field (right panels). For charged amino acid the dipole moment is computed using the center of mass as the origin.

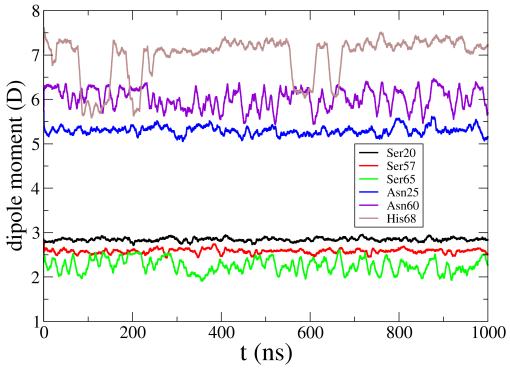
A) Glutamine

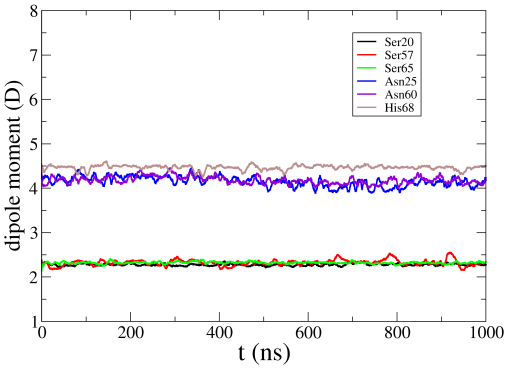


B) Threonine

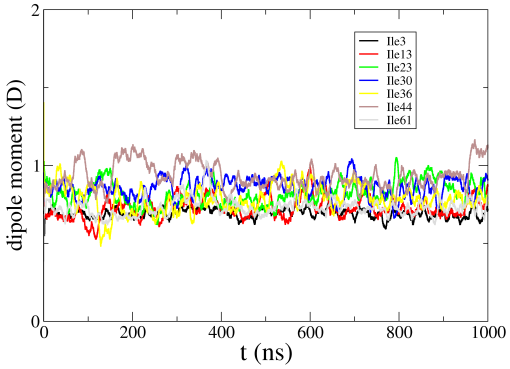


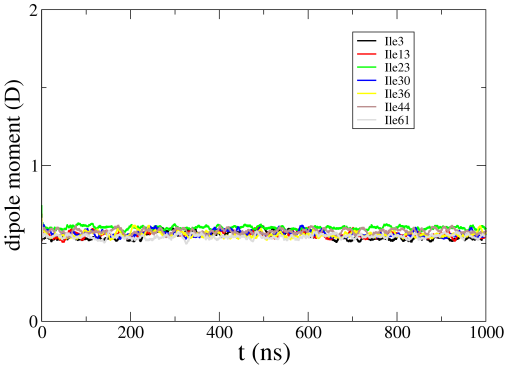
C) Serine, Asparagine and Histidine



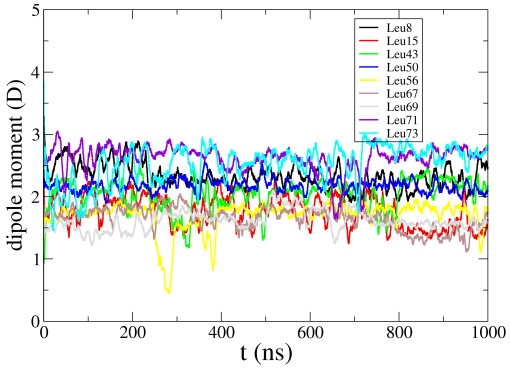


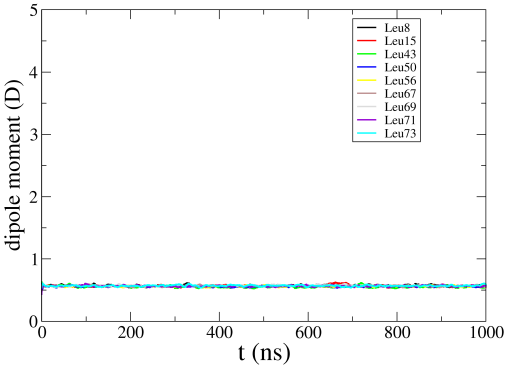
D) Isoleucine



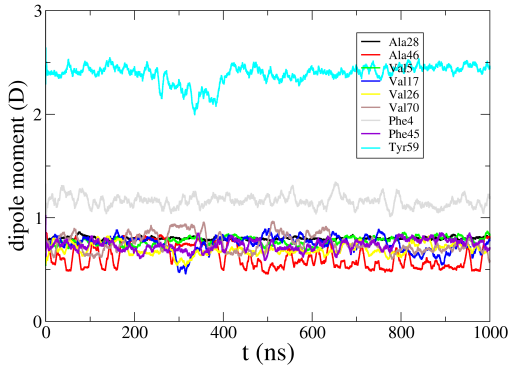


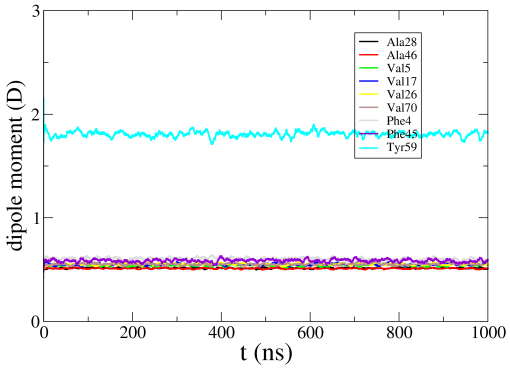
E) Leucine



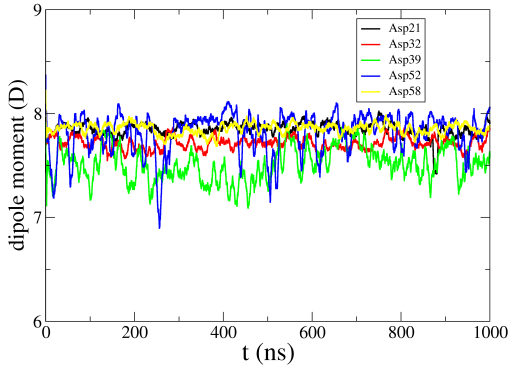


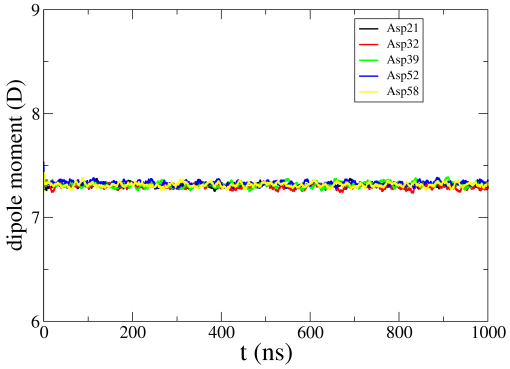
F) Alanine, Valine, Phenylalanine, and Tyrosine



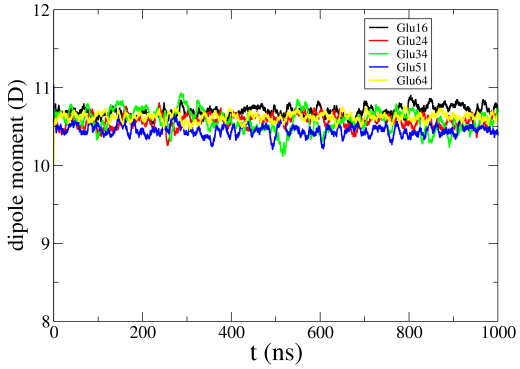


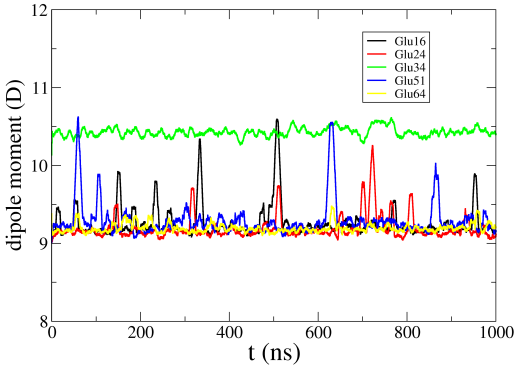
G) Aspartate



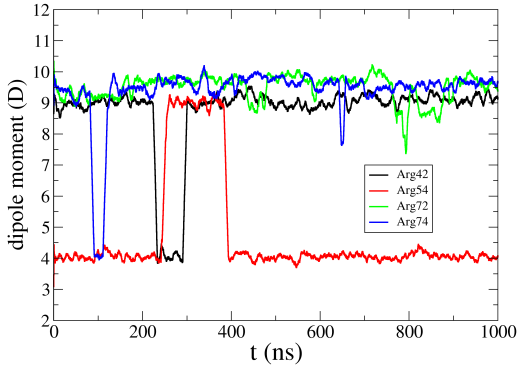


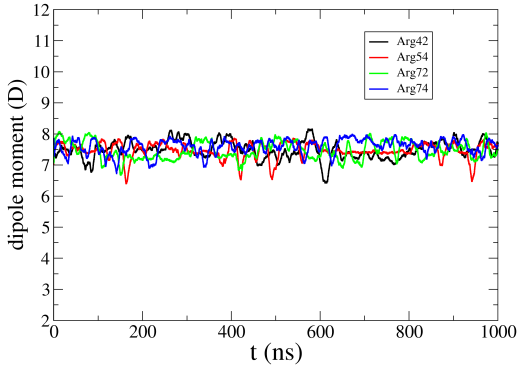
H) Glutamate



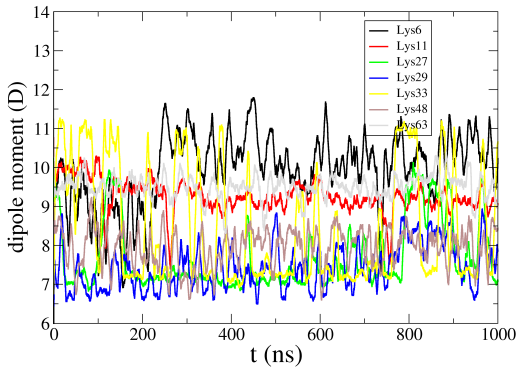


I) Arginine


dipole moment (D)
t (ns)
Arg42
Arg54
Arg72
Arg74


dipole moment (D)
t (ns)
Arg42
Arg54
Arg72
Arg74

J) Lysine


dipole moment (D)
t (ns)
Lys6
Lys11
Lys27
Lys29
Lys33
Lys48
Lys63

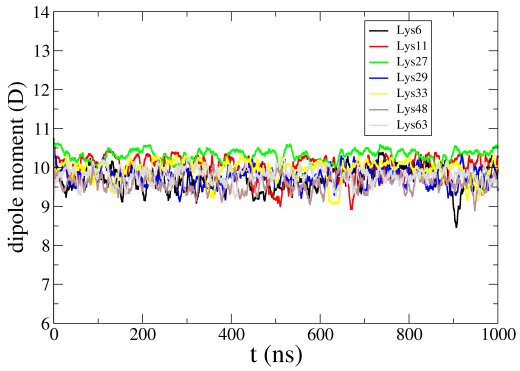
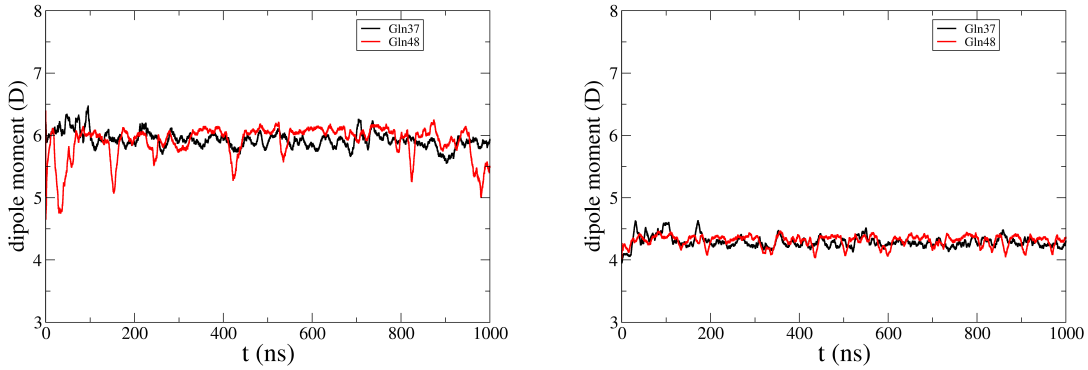
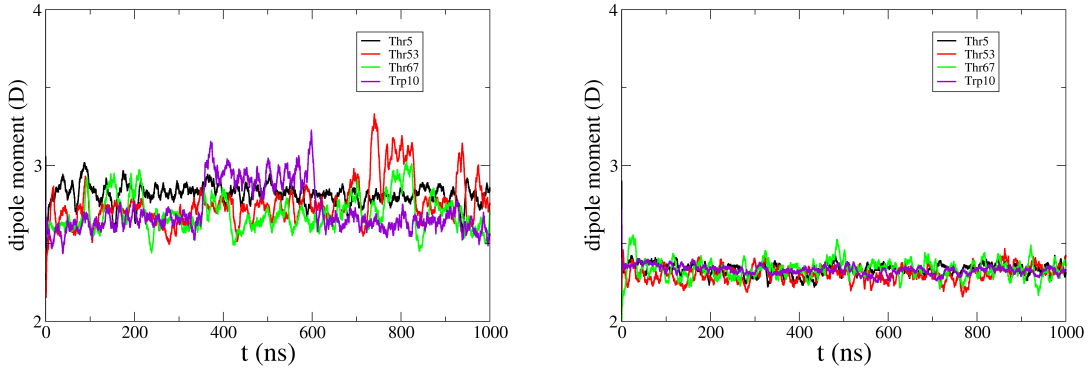

dipole moment (D)
t (ns)
Lys6
Lys11
Lys27
Lys29
Lys33
Lys48
Lys63

Fig. S3 Side chain dipole moments from the 1 μs MD simulations of cold-shock protein A with the polarizable Drude force field (left panels) and the additive CHARMM36 force field (right panels). For charged amino acid the dipole moment is computed using the center of mass as the origin.

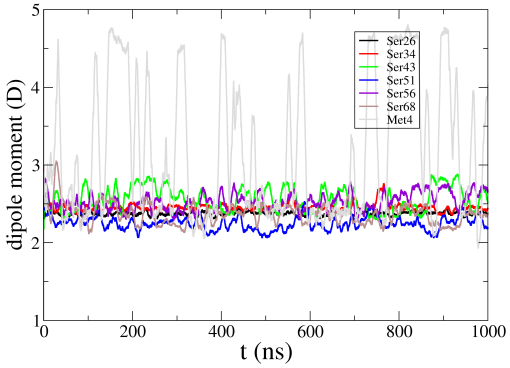
A) Glutamine

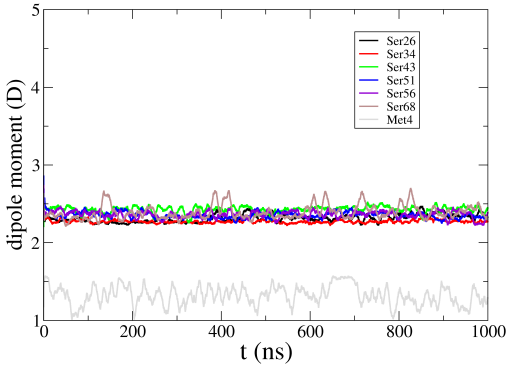


B) Threonine and Tryptophan

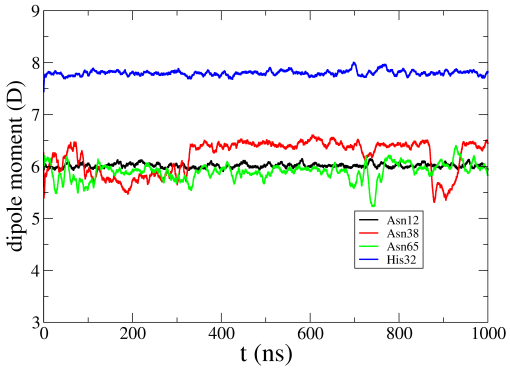


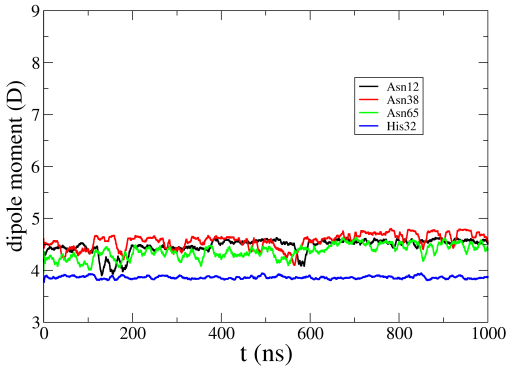
C) Serine and Methionine



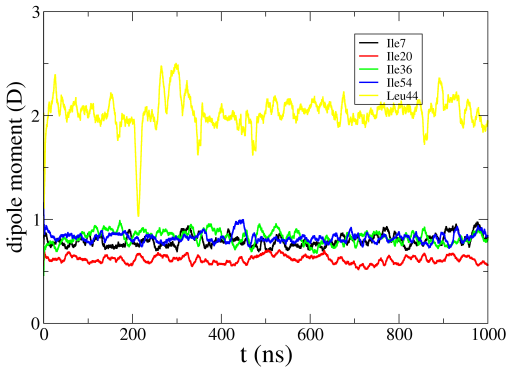


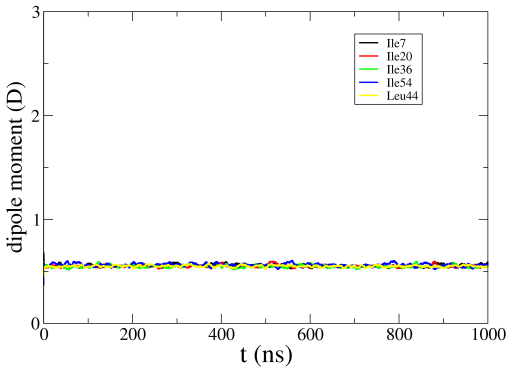
D) Asparagine and Histidine



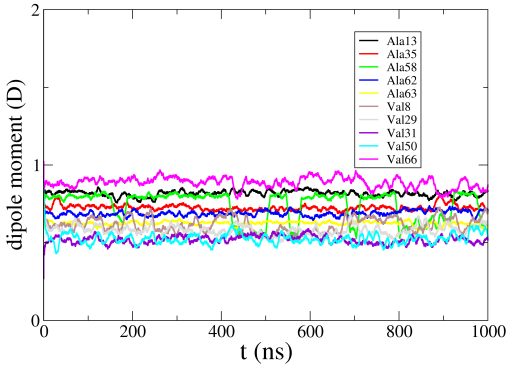


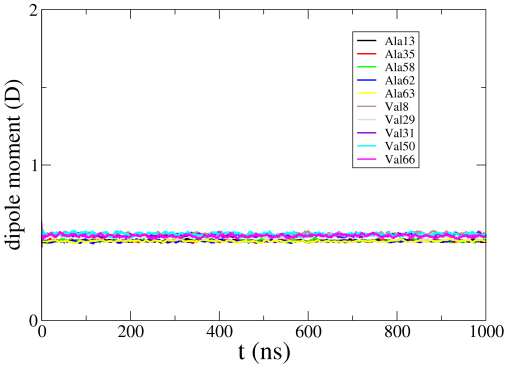
E) Isoleucine and Leucine



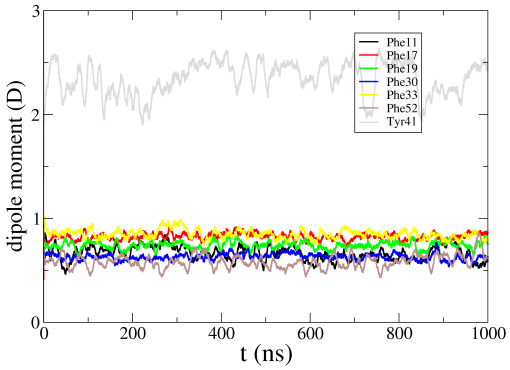


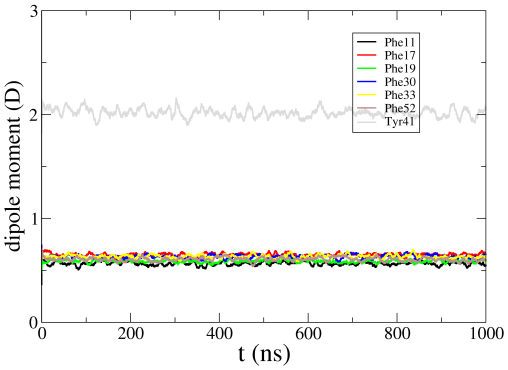
F) Alanine and Valine



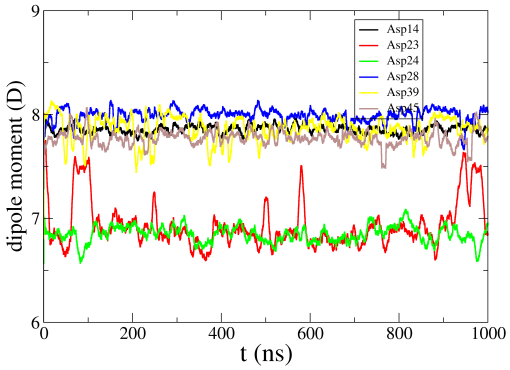


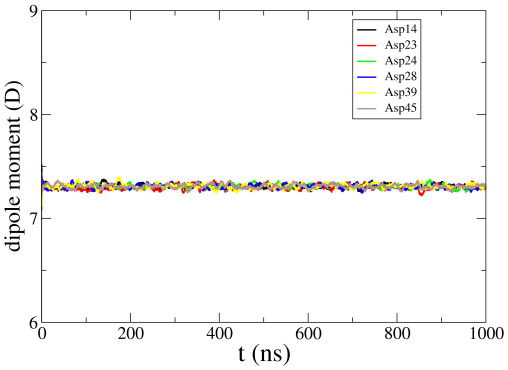
G) Phenylalanine and Tyrosine



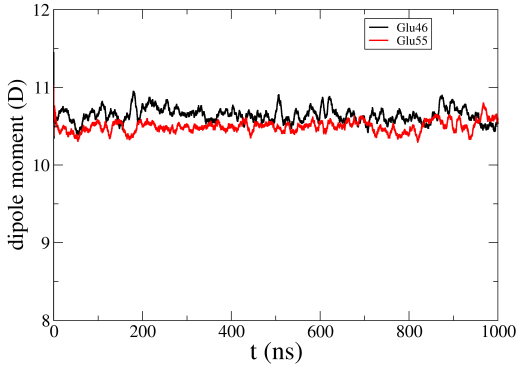


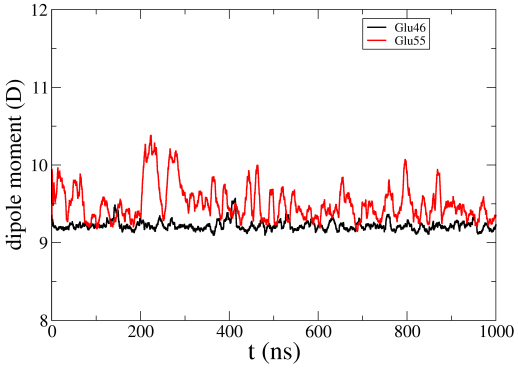
H) Aspartate



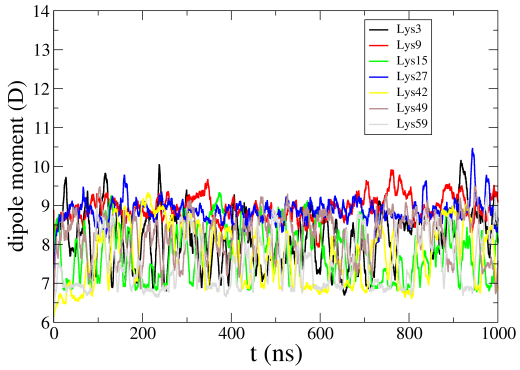


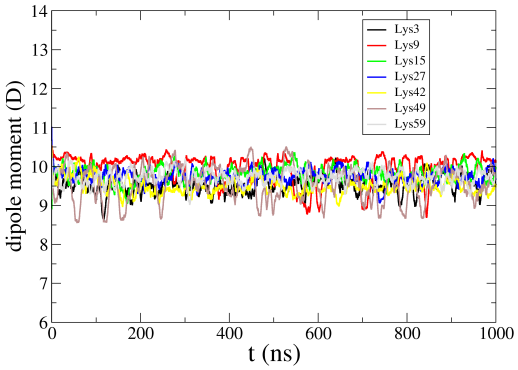
I) Glutamate


Glu46
Glu55
dipole moment (D)
t (ns)


Glu46
Glu55
dipole moment (D)
t (ns)

J) Lysine


Lys3
Lys9
Lys15
Lys27
Lys42
Lys49
Lys59
dipole moment (D)
t (ns)


Lys3
Lys9
Lys15
Lys27
Lys42
Lys49
Lys59
dipole moment (D)
t (ns)

Calculation of protein dielectric properties.

The optical dielectric constant, $\varepsilon^{\inf}$,is computed with

$$\frac{\varepsilon^{\inf}-1}{\varepsilon^{\inf}+2}=\frac{\langle A\rangle}{r^3}$$

, where <A> is the average molecular polarizability and $r=\sqrt{5/3}R_g$ where R_g is the radius of gyration. The static dielectric constant for the entire protein, ε_p, is given by

$$\frac{\langle \Delta M_p^2\rangle}{kTr^3}=\frac{(2\varepsilon_w+1)(\varepsilon_p-1)}{2\varepsilon_w+\varepsilon_p}$$

, where $\langle \Delta M_p^2\rangle$ is the square of average protein dipole moment fluctuation and ε_w is the dielectric constant of water. To compute dielectric constant for the protein interior, ε_{in}, the protein is modeled as a sphere with dielectric constant ε_{in} and radius r_{in}, and an outer spherical shell with dielectric constant ε_{out}, which is related by:

$$\frac{\langle \Delta M_{in}^2\rangle}{kTr^3}=\frac{(\varepsilon_{in}-1)[(1+2\varepsilon_{out})(2\varepsilon_w+\varepsilon_{out})-2\left(\frac{r_{in}}{r}\right)^3(\varepsilon_w-\varepsilon_{out})(1-\varepsilon_{out})]}{(\varepsilon_{in}+2\varepsilon_{out})(2\varepsilon_w+\varepsilon_{out})-2\left(\frac{r_{in}}{r}\right)^3(\varepsilon_w-\varepsilon_{out})(\varepsilon_{in}-\varepsilon_{out})}$$

and

$$\left(\frac{r_{in}}{r}\right)^3\varepsilon_{in}+\left[1-\left(\frac{r_{in}}{r}\right)^3\right]\varepsilon_{out}=\varepsilon_p$$

, where $\langle \Delta M_{in}^2\rangle$ is the square of average dipole moment fluctuation of the protein interior. The dielectric constant for the hydrophobic core is computed in the same way. The statistical uncertainty is estimated with block averages by dividing the MD trajectories into 5 blocks.

Figure S4. Trace of the molecular polarizability tensors from the 1 μs Drude simulations for ubiquitin and CspA.

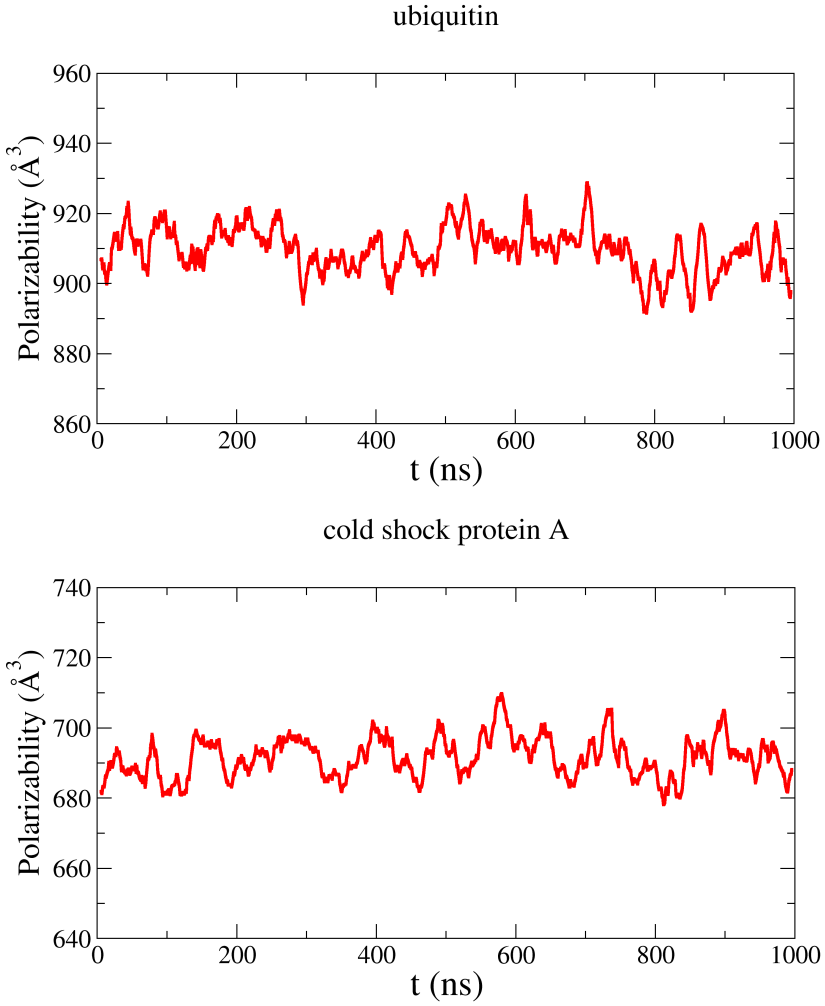


Table S2. Dielectric properties of proteins with the fully polarizable Drude model. The total charge, Q; radius of gyration, R_g; ensemble average molecular polarizabilities, A; average dipole moment, M; and average dipole moment fluctuation, ΔM^2, are listed. For non-neutral molecules the magnitude of dipole moment is origin-dependent, with the center of mass chosen as the origin.

	Q (e)	R_g (Å)	$\langle A \rangle$ (Å^3)	$\langle M \rangle$ (D)	$\langle \Delta M^2 \rangle$ (D^2)
1ubq	0	11.63	909.7	246.2	1341.1
1mjc	-1	10.87	692.7	235.6	1068.8
1ejg	0	9.88	418.9	85.6	137.9
3vqf	-1	11.97	990.5	204.8	1040.9
3zzp	-2	12.23	919.3	299.5	1343.1
1p7e	-2	10.81	630.9	252.2	906.5

Reference

(1) Lopes, P. E. M.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell, A. D., Jr., *Journal of Chemical Theory and Computation* **2013**, *9*, 5430.

(2) Jiang, W.; Hardy, D. J.; Phillips, J. C.; MacKerell, A. D., Jr.; Schulten, K.; Roux, B. *The journal of physical chemistry letters* **2010**, *2*, 87.

(3) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.