Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting Information: Improving IDP Theoretical Chemical Shift Accuracy and Efficiency Through a Combined MD/ADMA/DFT and Machine Learning Approach

Michael J. Bakker,[†] Arnošt Mládek,[†] Hugo Semrád,[†,‡] Vojtěch Zapletal[†],[†] and Jana Pavlíková Přecechtělová*[,†]

[†]*Faculty of Pharmacy in Hradec Králové, Charles University, Akademika Heyrovského 1203, 500 05 Hradec Králové, Czech Republic*

[‡]*Department of Chemistry, Faculty of Science, Masaryk University, 611 37 Brno, Czech Republic*

E-mail: precechj@faf.cuni.cz

Phone: +420 495 067 488

Table S1: Ensemble averaged CSs calculated for the CLUSTER ensemble with and without geometry optimization. ^{15}N CSs were referenced using ^{e}Nref1, ^{f}Nref2, and ^{g}Nref3). ^{31}P CSs were referenced using ^{h}Pref1, ^{i}Pref2, and ^{j}Pref3. ^{k} Experimental values of ^{1}H, ^{13}C, and ^{15}N CSs are taken from the Supporting Information of Ref.[S1], ^{31}P CSs are taken from Ref.[S2].

Residue	Atom	NON-OPTIMIZED	OPTIMIZED	Exp.^{k}
pS19	H	8.71 ± 0.27	8.22 ± 0.28	8.859
	HA	4.74 ± 0.06	4.75 ± 0.09	4.316
	HB1	4.27 ± 0.06	4.24 ± 0.06	3.905
	HB2	4.02 ± 0.06	4.07 ± 0.06	3.905
	C'	183.1 ± 0.61	183.21 ± 0.53	174.3
	CA	64.8 ± 0.48	64.34 ± 0.46	66.8
	CB	68.06 ± 0.37	68.29 ± 0.45	59.1
	N^{a}	152.2 ± 0.98	150.36 ± 1.3	119.7
	N^{b}	138.76 ± 0.98	136.92 ± 1.3	119.7
	N^{c}	124.92 ± 0.98	123.08 ± 1.3	119.7
	P^{d}	0.7 ± 0.6	0.81 ± 0.17	3.76
	P^{e}	31.85 ± 0.6	31.95 ± 0.17	3.76
	P^{f}	−5.21 ± 0.6	−5.11 ± 0.17	3.76
pS40	H	7.76 ± 0.21	7.96 ± 0.22	8.841
	HA	4.61 ± 0.07	4.74 ± 0.07	4.238
	HB1	4.23 ± 0.06	4.81 ± 0.8	3.904
	HB2	4.1 ± 0.08	4.6 ± 0.8	3.904
	C	181.15 ± 0.64	182.22 ± 0.53	174.4
	CA	63.21 ± 0.51	63.12 ± 0.49	66.3
	CB	67.44 ± 0.45	70.14 ± 3.36	59.8
	N^{a}	146.88 ± 1.47	151.07 ± 1.41	117.2
	N^{b}	133.44 ± 1.47	137.63 ± 1.41	117.2
	N^{c}	119.6 ± 1.47	123.8 ± 1.41	117.2
	P^{d}	0.77 ± 0.6	6.63 ± 2.14	4.18
	P^{e}	31.91 ± 0.6	37.77 ± 2.14	4.18
	P^{f}	−5.15 ± 0.6	0.71 ± 2.14	4.18

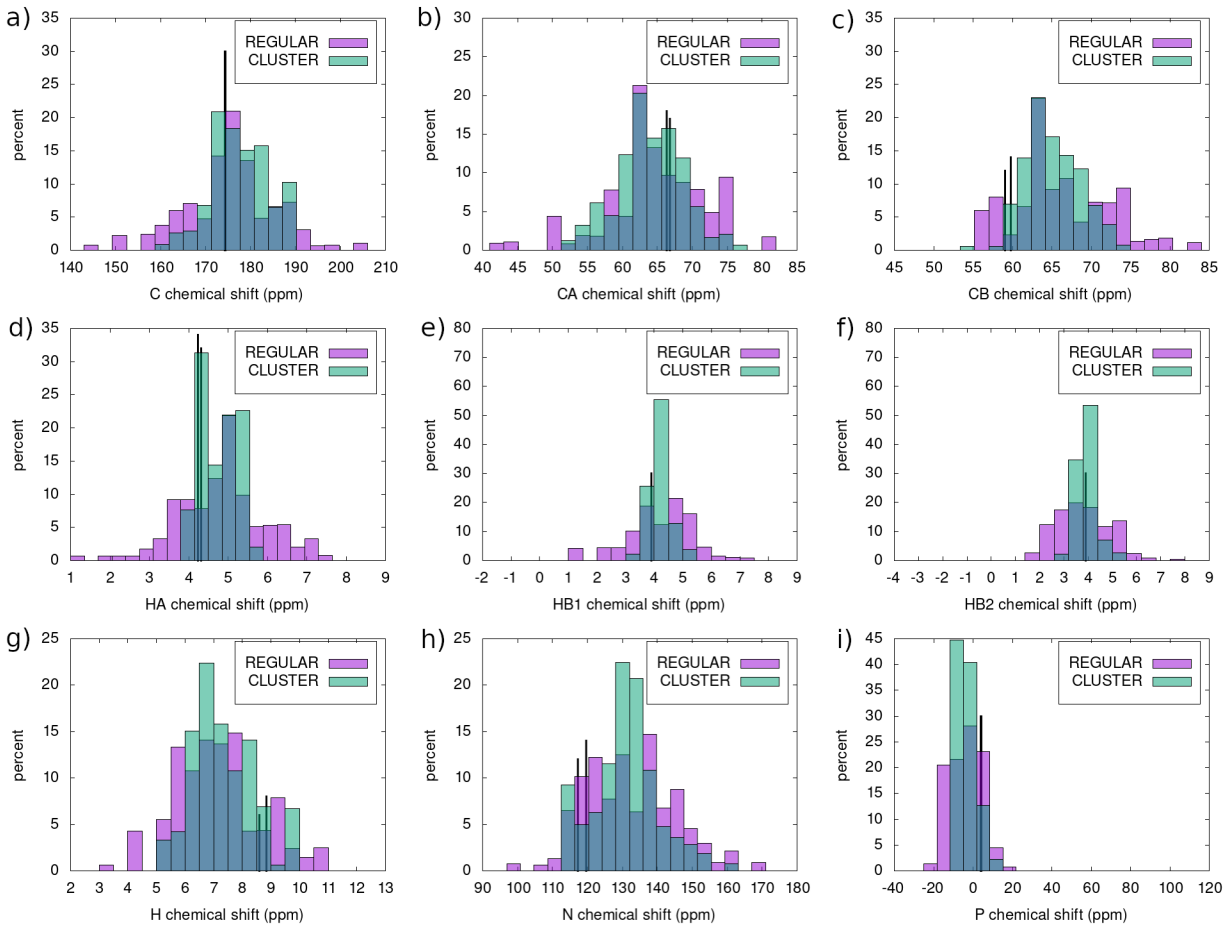

Figure S1: Histograms of CSs of the combined pS19 and pS40 obtained for C' (a), CA (b), CB (c), HA (d), HB1 (e), HB2 (f), H^N (g), N Ref1 (h), and P Ref1 (i) obtained using the CLUSTER and REGULAR ensembles. Black lines are included to represent experimentally obtained CS values for reference.

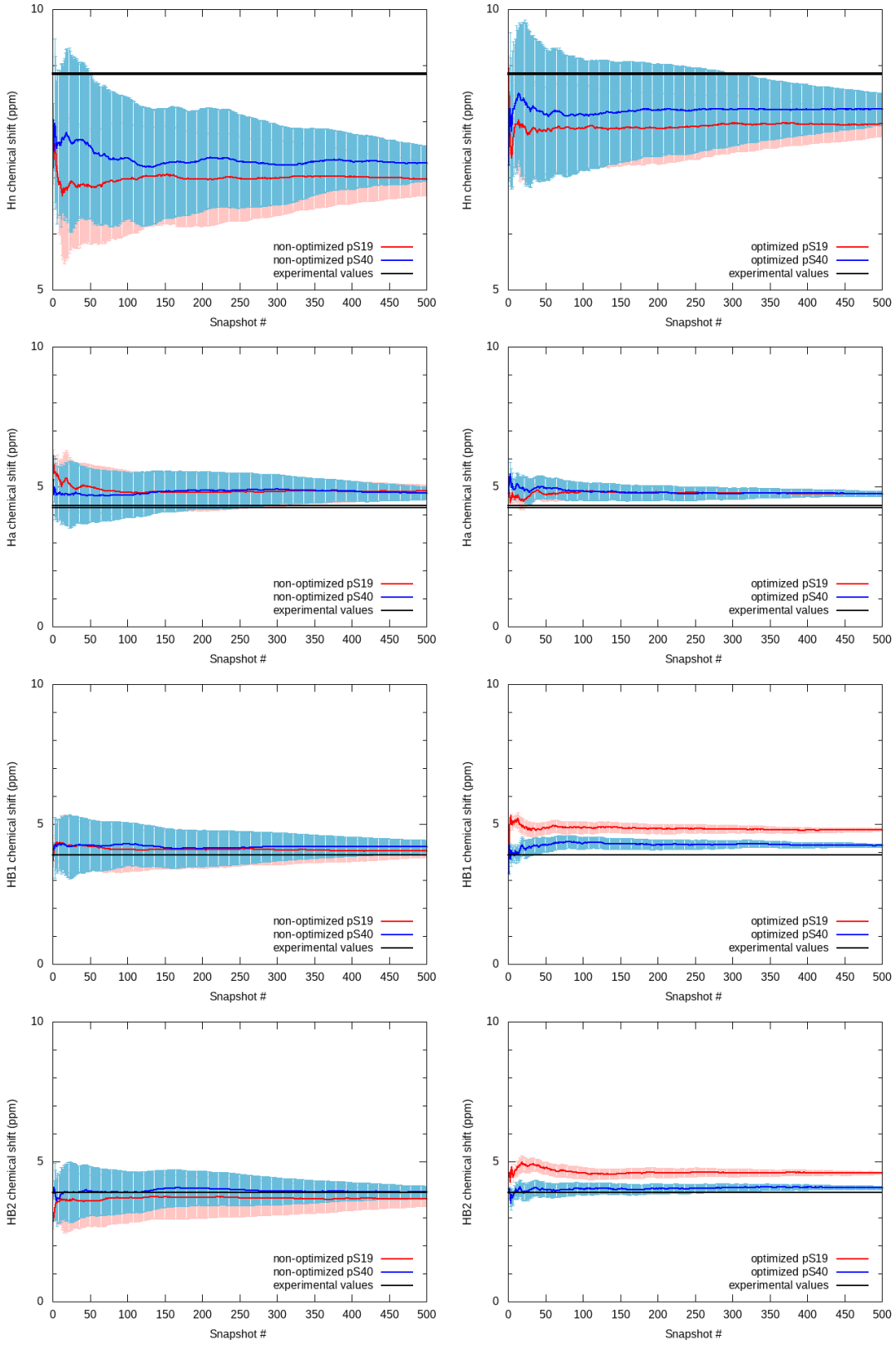

Figure S2: Running MEEs for the optimized REGULAR ensemble with 500 frames for Hydrogen Atoms: (first) H^N, (second) HA, (third) HB1, and (fourth) HB2 CSs

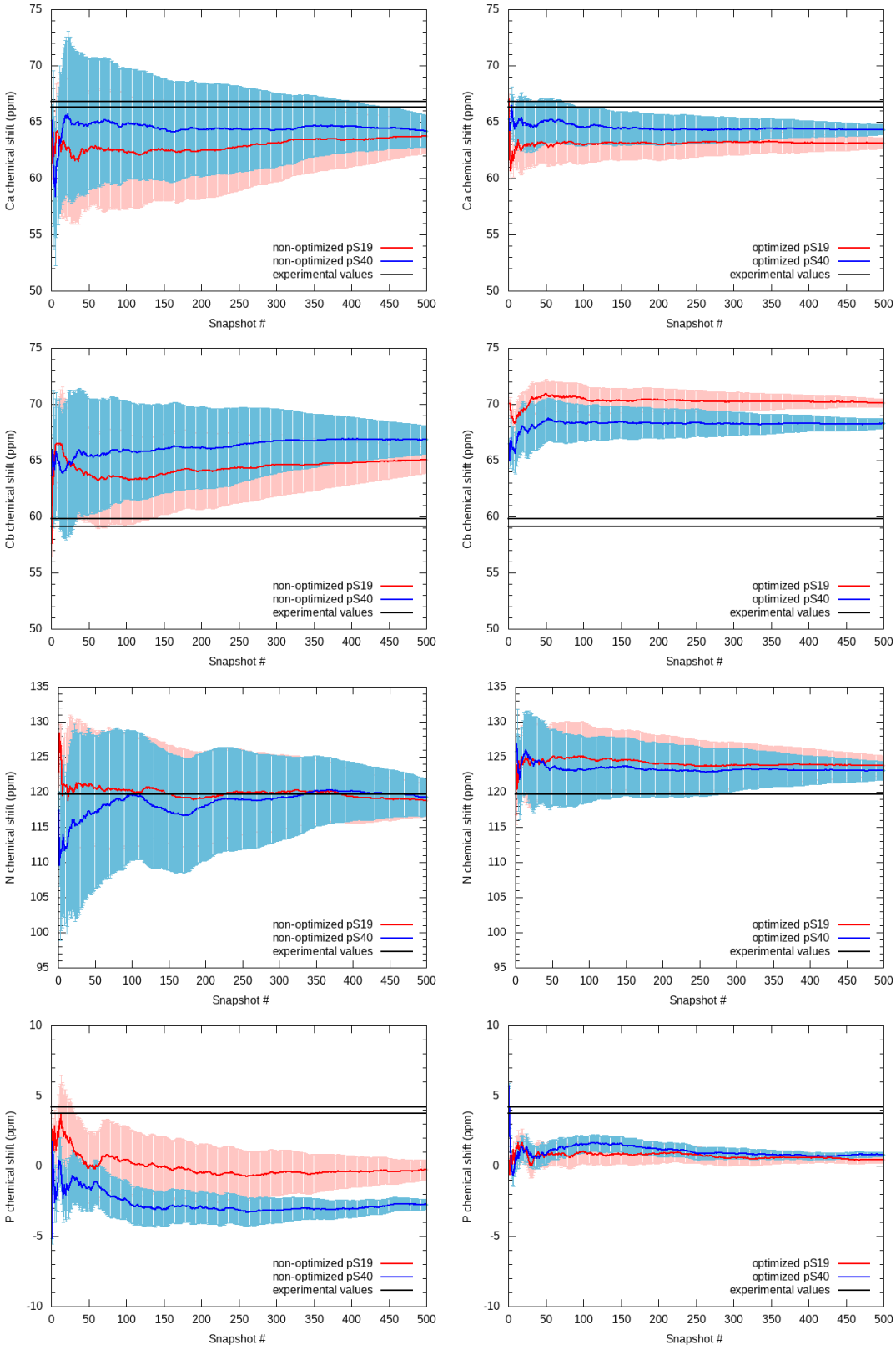

Figure S3: Running MEEs for the optimized REGULAR ensemble with 500 frames for Hydrogen Atoms: (first) CA, (second) CB, (third) N, and (fourth) P CSs

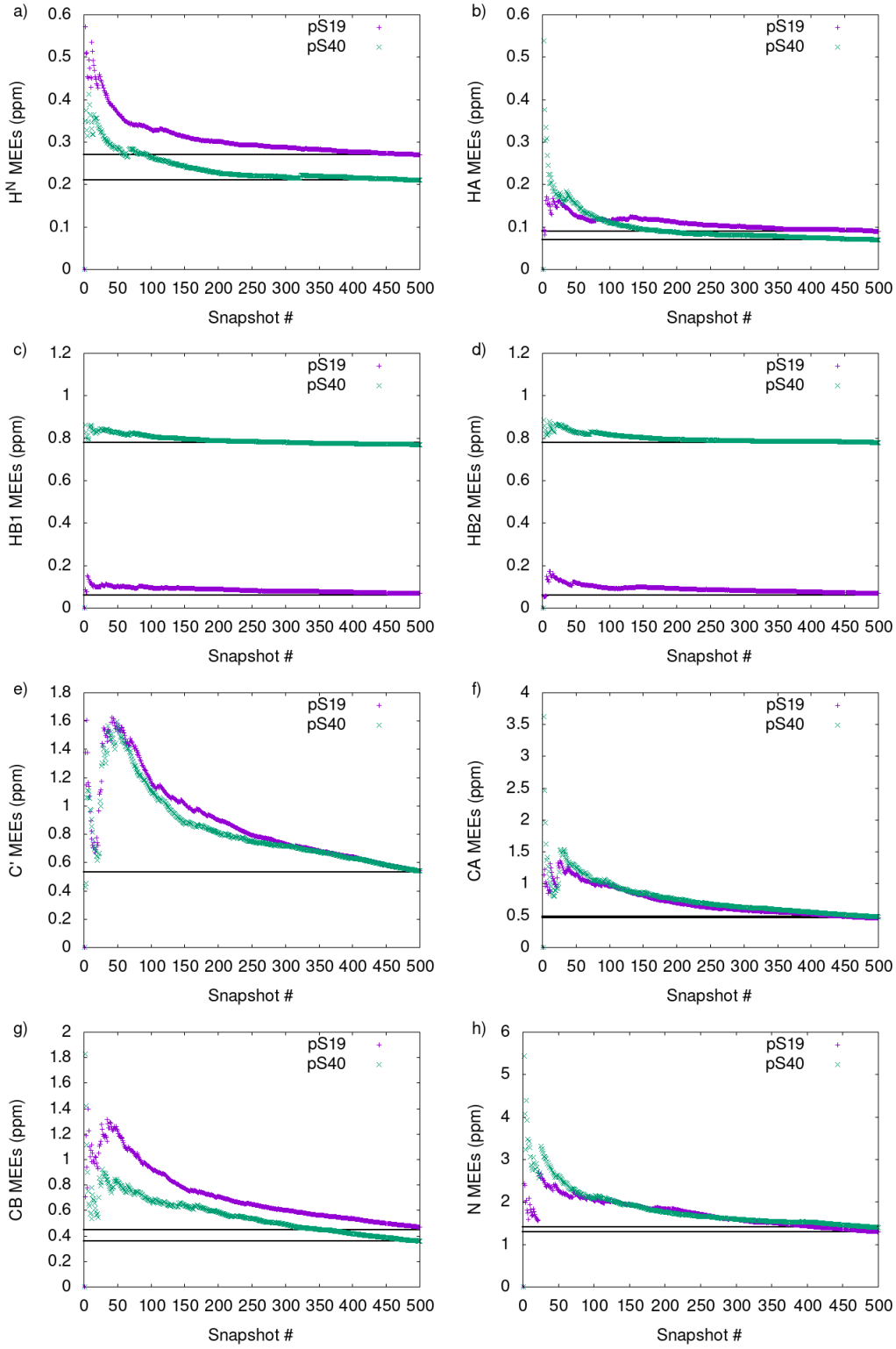

Figure S4: Running MEEs for the optimized REGULAR ensemble with 500 frames: (a) H^N, (b) HA, (c) HB1, (d) HB2, (e) C', (f) CA, (g) CB, and (h) N CSs

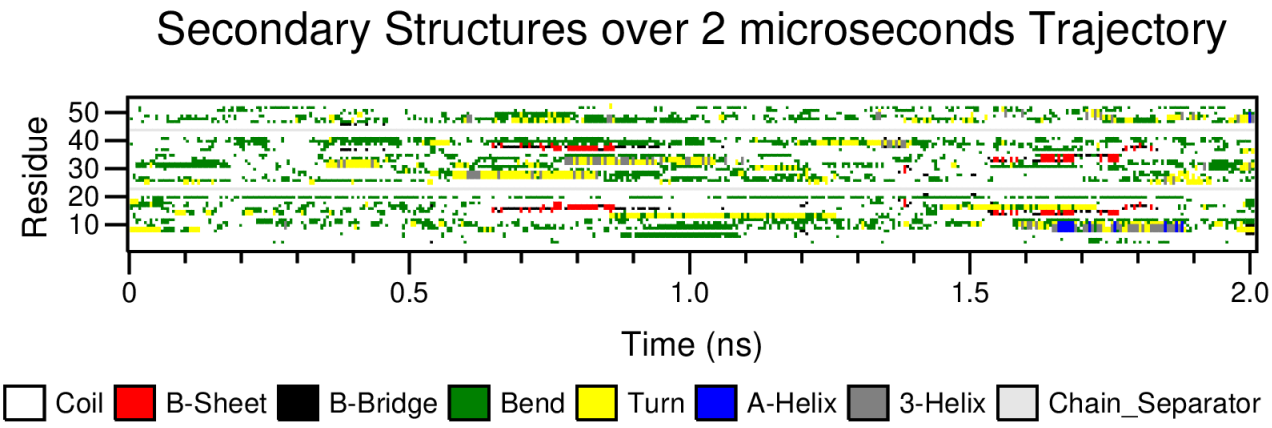

Figure S5: Secondary Structures analysis of a continuation (2 microseconds) trajectory. DSSP shows new novel secondary structures, such as beta sheets and alpha helixes, although these structures are short lived, represented the untrapped nature of the trajectory.

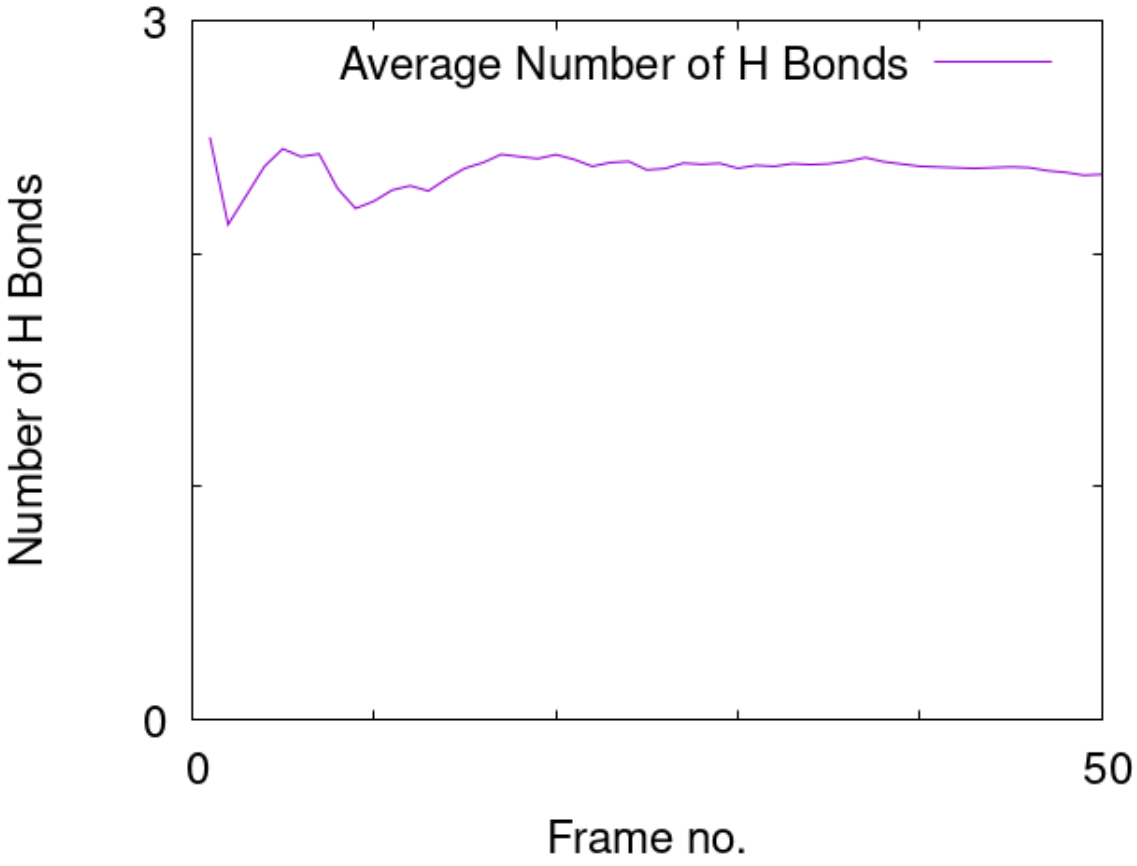

Figure S6: Average count of hydrogen bonds throughout each of the clusters. A hydrogen bond is only considered if the H in question is within 3 Angstroms of the nucleophilic site.

References

(S1) Louša, P.; Nedozrálová, H.; Župa, E.; Nováček, J.; Hritz, J. Phosphorylation of the Regulatory Domain of Human Tyrosine Hydroxylase 1 Monitored Using Non-Uniformly Sampled NMR. *Biophys. Chem.* **2017**, *223*, 25–29.

(S2) Hritz, J.; Byeon, I.-J. L.; Krzysiak, T.; Martinez, A.; Sklenář, V.; Gronenborn, A. M. Dissection of Binding between a Phosphorylated Tyrosine Hydroxylase Peptide and 14-3-3: A Complex Story Elucidated by NMR. *Biophys. J.* **2014**, *107*, 2185–2194.