Supporting information

Performance Evaluation of Balanced Force Field ff03CMAP for Intrinsically Disordered and Ordered Proteins

Yuxin Jiang[1], Hai-Feng Chen[1,2,*]

[1]State Key Laboratory of Microbial metabolism, Joint International Research Laboratory of Metabolic & Developmental Sciences, Department of Bioinformatics and Biostatistics, National Experimental Teaching Center for Life Sciences and Biotechnology, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, 200240, China

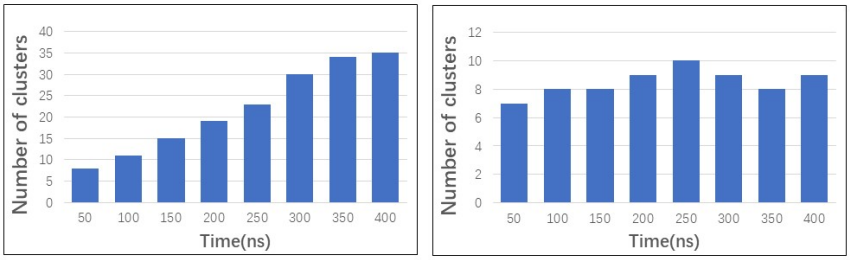
[2]Shanghai Center for Bioinformation Technology, 200240, Shanghai, China.

***Corresponding Author**
Hai-Feng Chen (Full Professor)
State Key Laboratory of Microbial metabolism, Joint International Research Laboratory of Metabolic & Developmental Sciences, Department of Bioinformatics and Biostatistics, National Experimental Teaching Center for Life Sciences and Biotechnology, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, 200240, China
Tel: 86-21-34204073; **Fax**: 86-21-34204073; **Email**: haifengchen@sjtu.edu.cn

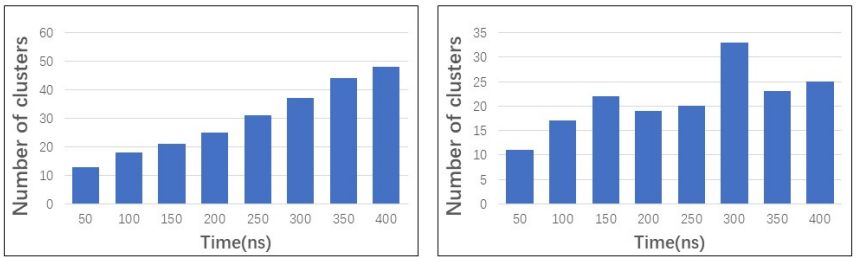
Notes

The authors declare that there is no conflict of interest.

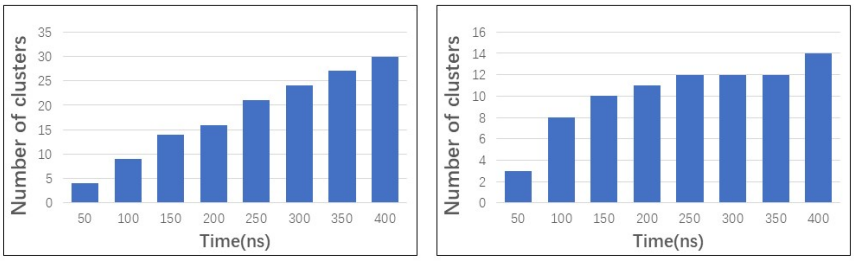
A


Number of clusters
Time(ns)
50 100 150 200 250 300 350 400
0 5 10 15 20 25 30 35 40
0 2 4 6 8 10 12

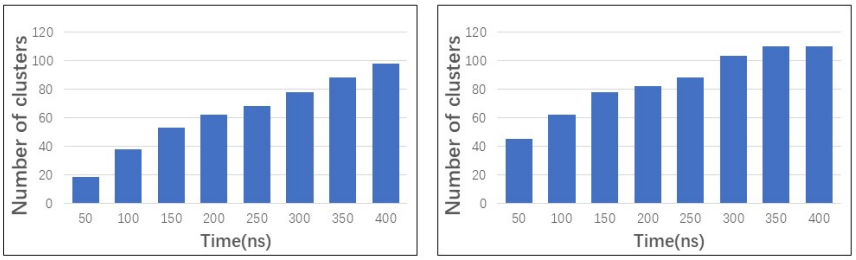
B


Number of clusters
Time(ns)
50 100 150 200 250 300 350 400
0 10 20 30 40 50 60
0 5 10 15 20 25 30 35

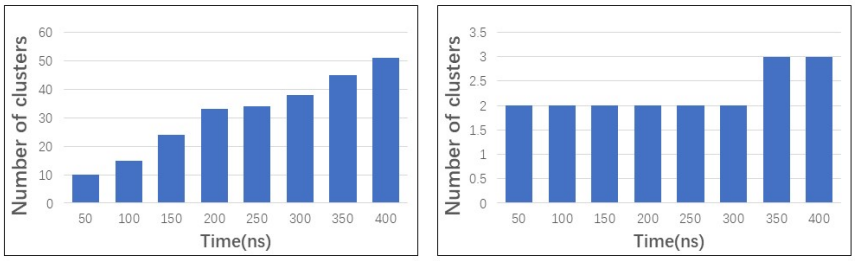
C


Number of clusters
Time(ns)
50 100 150 200 250 300 350 400
0 5 10 15 20 25 30 35
0 2 4 6 8 10 12 14 16

D


Number of clusters
Time(ns)
50 100 150 200 250 300 350 400
0 20 40 60 80 100 120

E


Number of clusters
Time(ns)
50 100 150 200 250 300 350 400
0 10 20 30 40 50 60
0 0.5 1 1.5 2 2.5 3 3.5

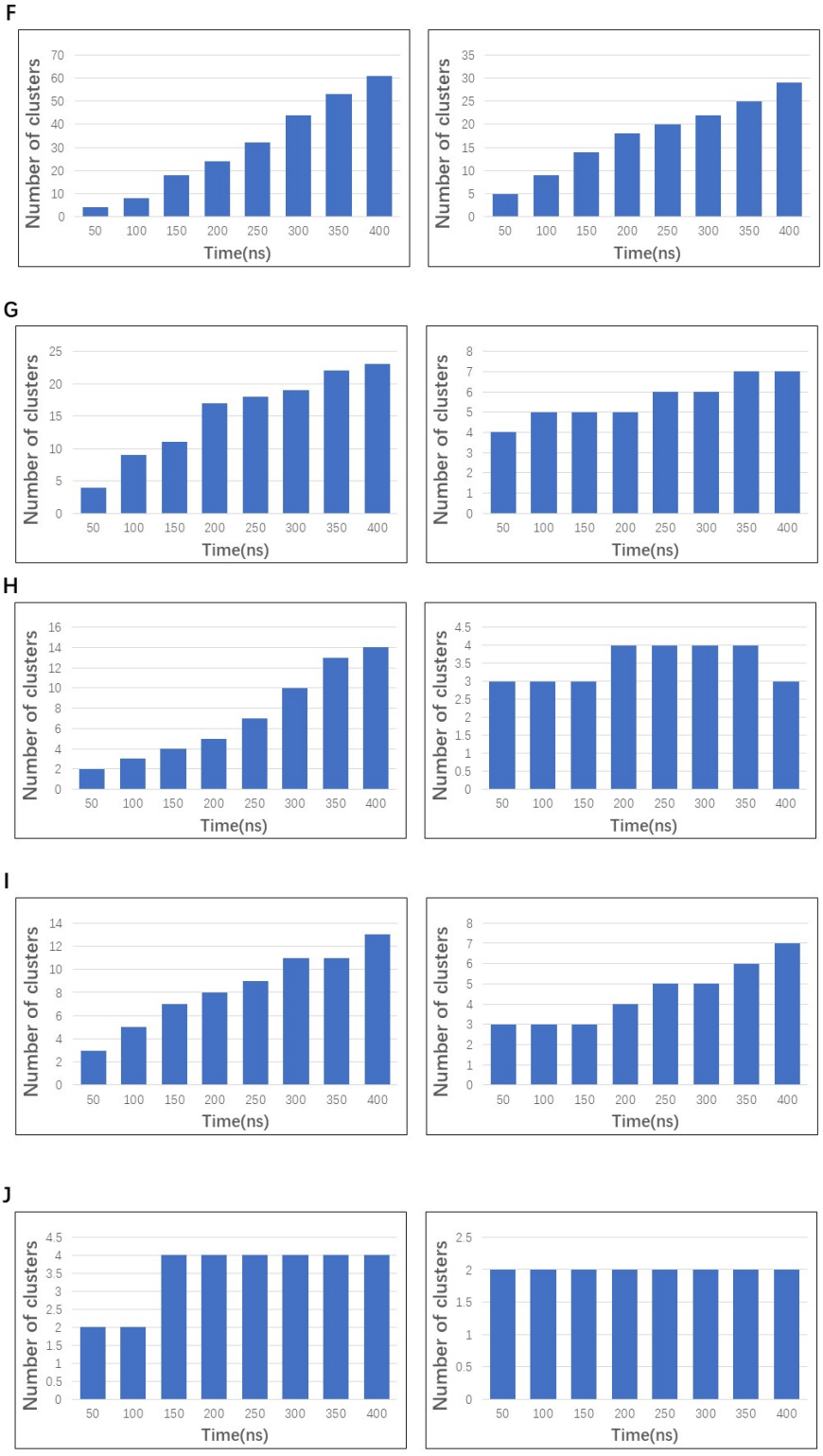
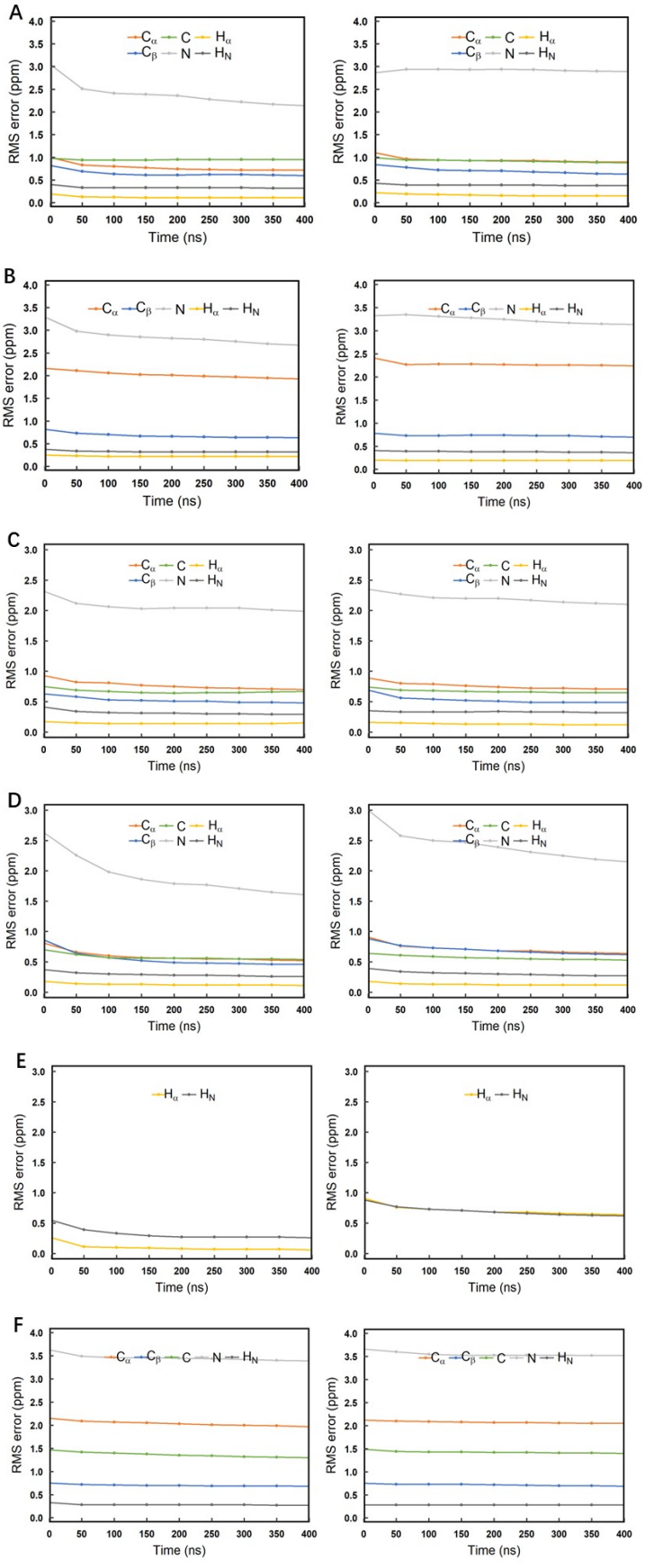


Figure S1. Time-dependent cumulative numbers of conformational clusters. The number of conformational clusters at a certain time is calculated with the simulated ensembles from 0 ns to this certain time. The left panel stands for ff03cmap and the right panel stands for ff03（a）drkN（b）Sic1 (c)PaaA2 (d)p15^{PAF} (e)Histain5 (f)synuclein (g)hIAPP (h)MeVN (i)2kpp (j)2krk


A
B
C
D
E
F
RMS error (ppm)
Time (ns)
Cα
Cβ
C
N
Hα
HN

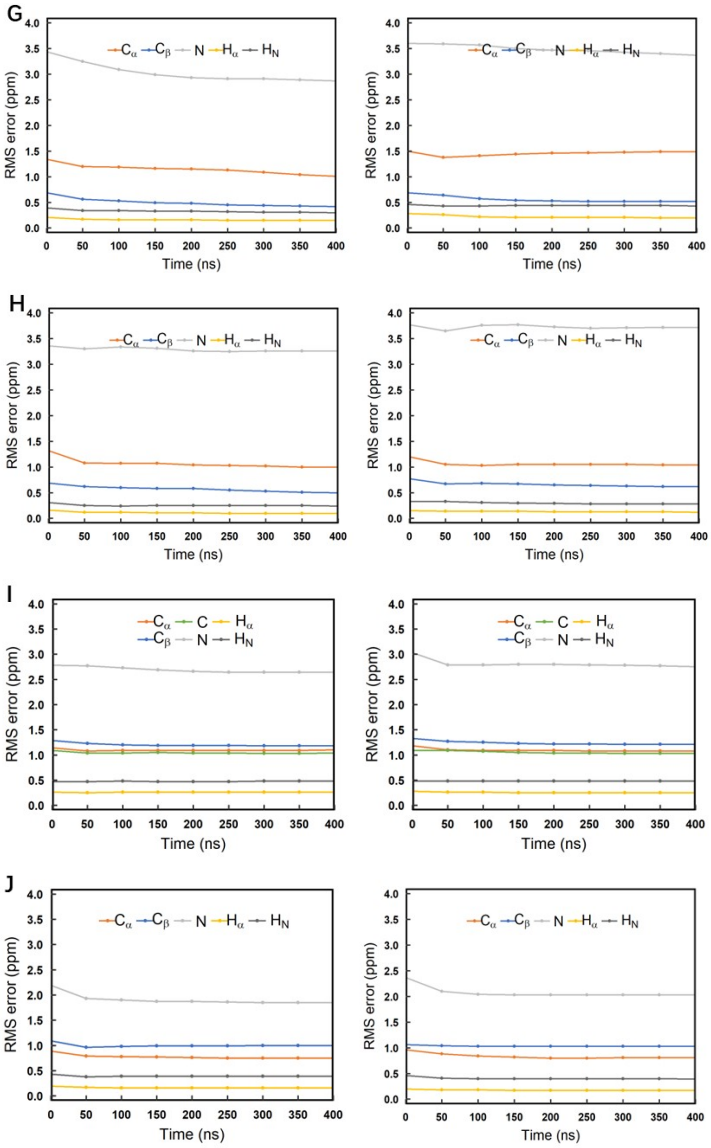


Figure S2. Time-dependent RMS error between experimental data and cumulative average chemical shifts simulated. The left panel stands for ff03cmap and the right panel stands for ff03（a）drkN（b）Sic1 (c)PaaA2 (d)p15^{PAF} (e)Histain5 (f)synuclein (g)hIAPP (h)MeVN (i)2kpp (j)2krk

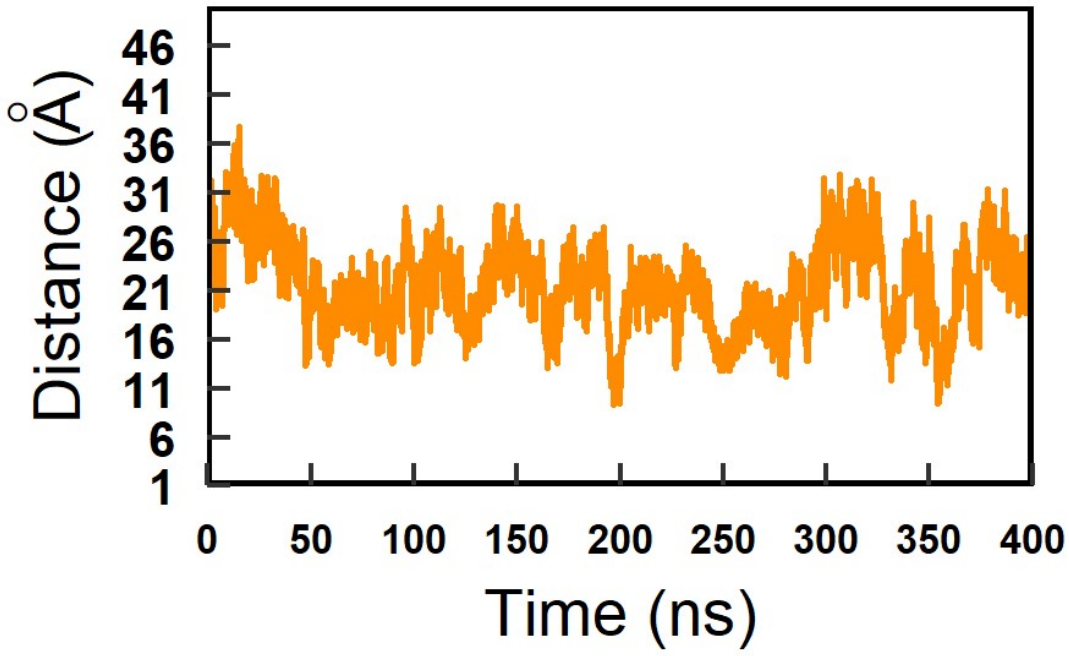


Figure S3. The end-to-end distance of this Histain5 during MD.

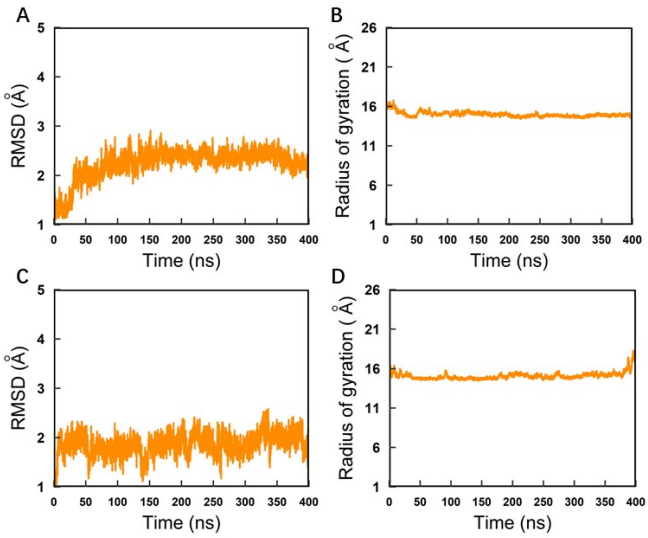


Figure S4. Structure stability of folded proteins 2kpp in simulations with the ff03cmap and ff03 force field. (A)The RMSD of 2kpp with ff03cmap (B) The Rg of 2kpp with ff03cmap (C)The RMSD of 2kpp with ff03 (D) The Rg of 2kpp with ff03

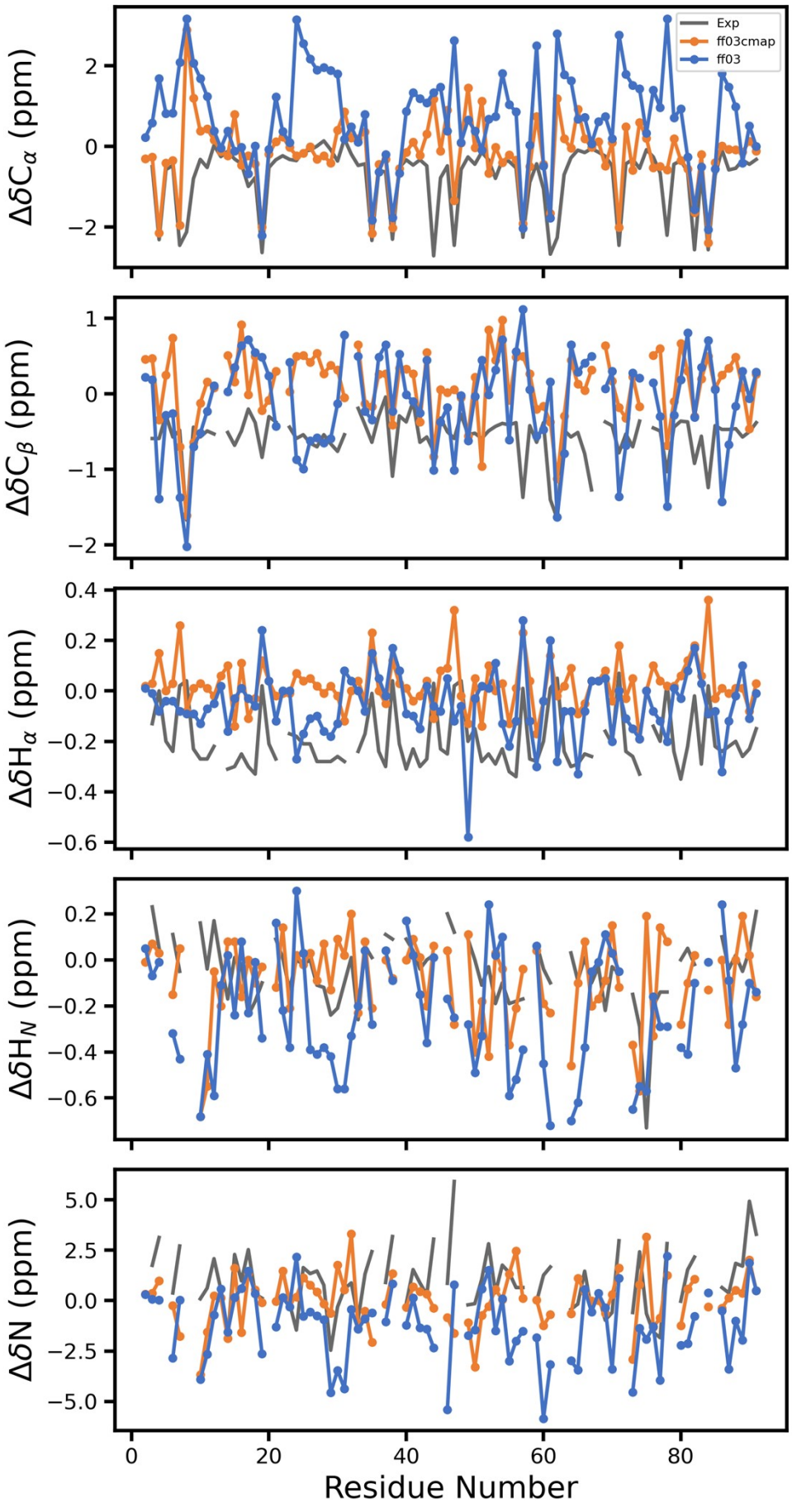

Exp
ff03cmap
ff03
ΔδCα (ppm)
ΔδCβ (ppm)
ΔδHα (ppm)
ΔδHN (ppm)
ΔδN (ppm)
Residue Number

Figure S5. Secondary chemical shifts of simulation and experimental data for Sic1. Simulated and experimental secondary chemical shifts for (A) Cα, (B) Cβ, (C) N, (D) Hα and (E) HN.

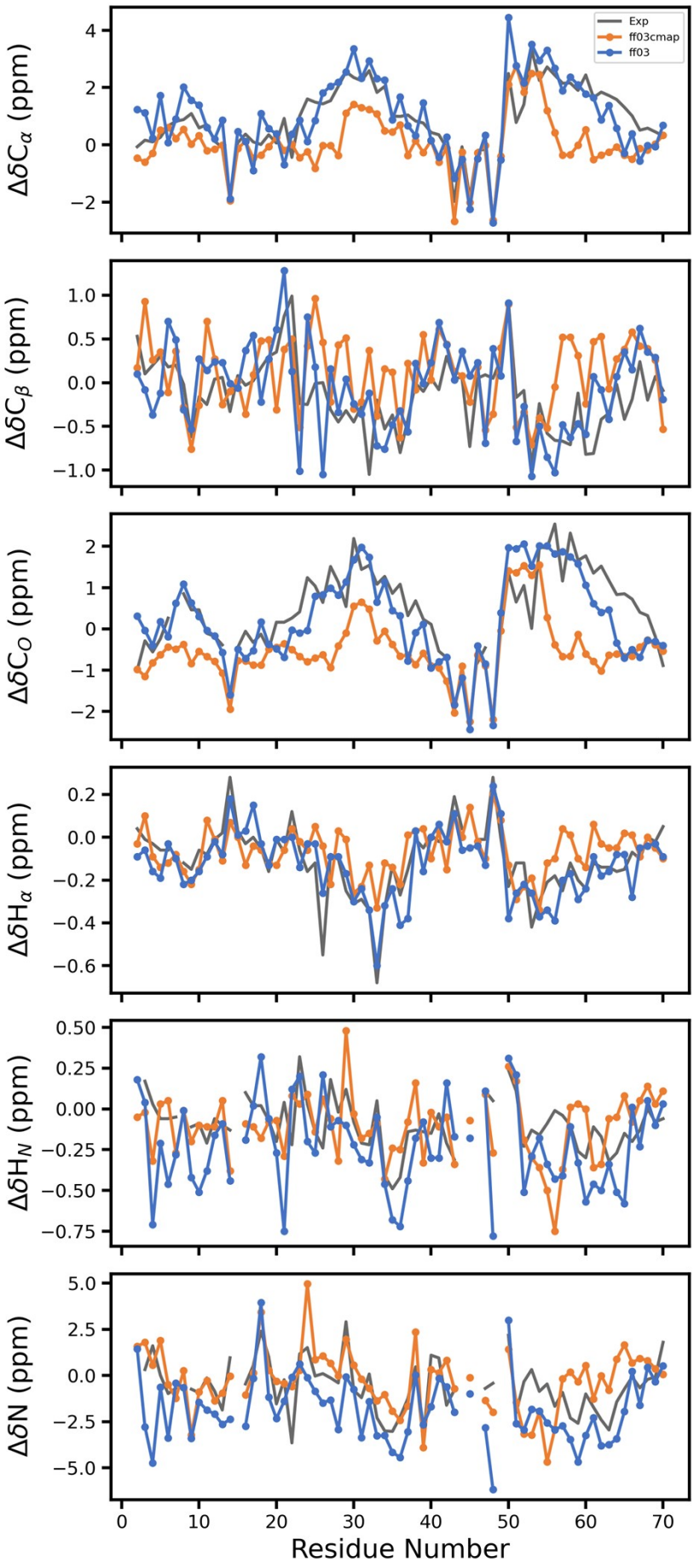

Exp
ff03cmap
ff03
ΔδCα (ppm)
ΔδCβ (ppm)
ΔδCO (ppm)
ΔδHα (ppm)
ΔδHN (ppm)
ΔδN (ppm)
Residue Number

Figure S6. Secondary chemical shifts of simulation and experimental data for PaaA2. Simulated and experimental secondary chemical shifts for (A) Cα, (B) Cβ, (C) C, (D) N, (E) Hα and (F) HN.

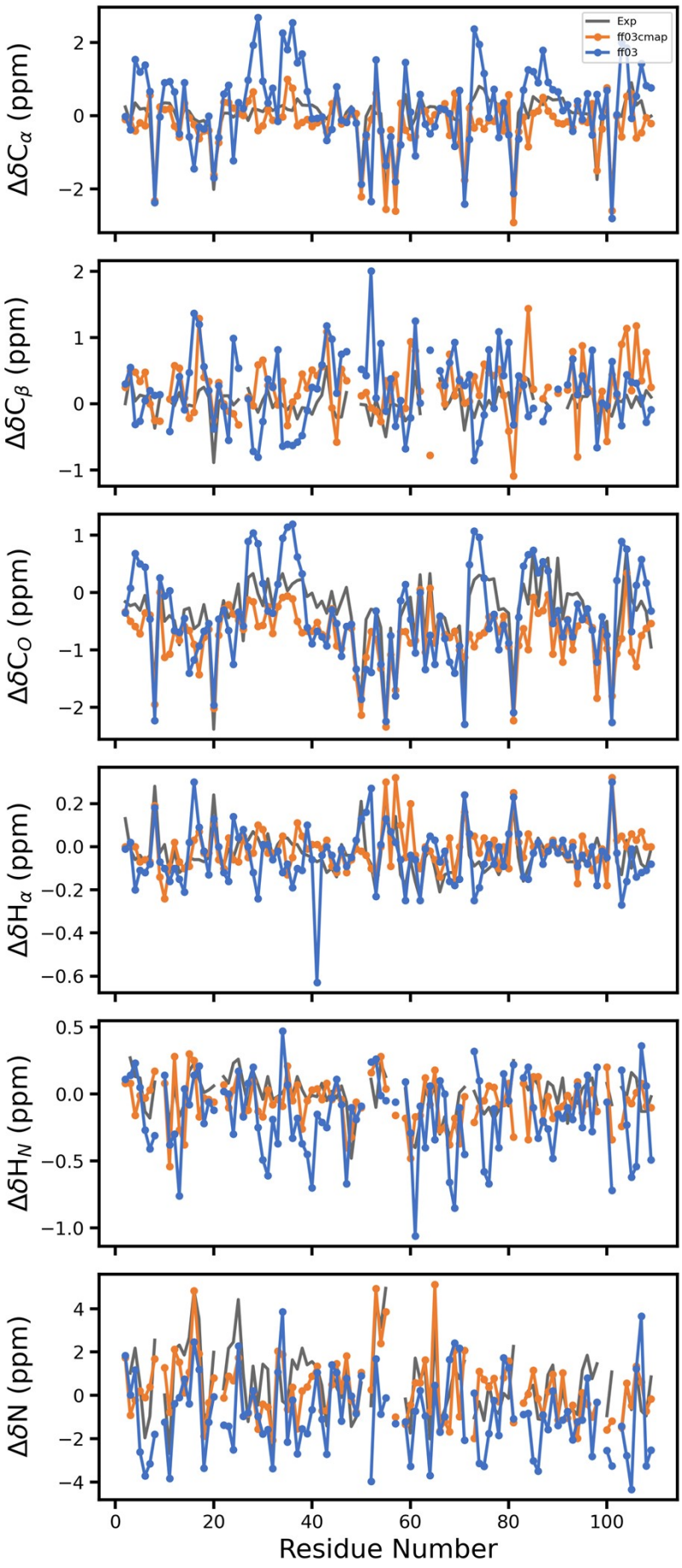

Exp
ff03cmap
ff03
ΔδCα (ppm)
ΔδCβ (ppm)
ΔδCO (ppm)
ΔδHα (ppm)
ΔδHN (ppm)
ΔδN (ppm)
Residue Number

Figure S7. Secondary chemical shifts of simulation and experimental data for p15^{PAF}. Simulated and experimental secondary chemical shifts for (A) Cα, (B) Cβ, (C) C, (D) N, (E) Hα and (F) HN.

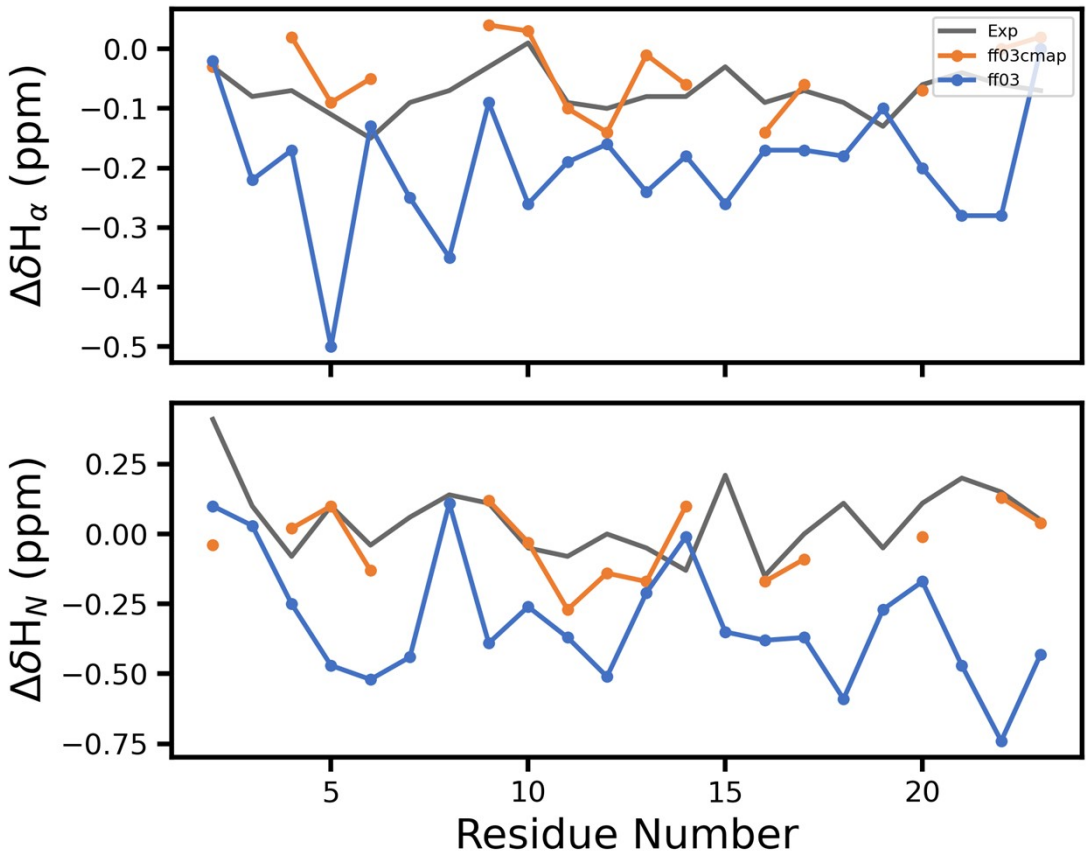


Figure S8. Secondary chemical shifts of simulation and experimental data for Histain5. Simulated and experimental secondary chemical shifts for (A) Cα, (B) Cβ, (C) C, (D) N, (E) Hα and (F) HN.

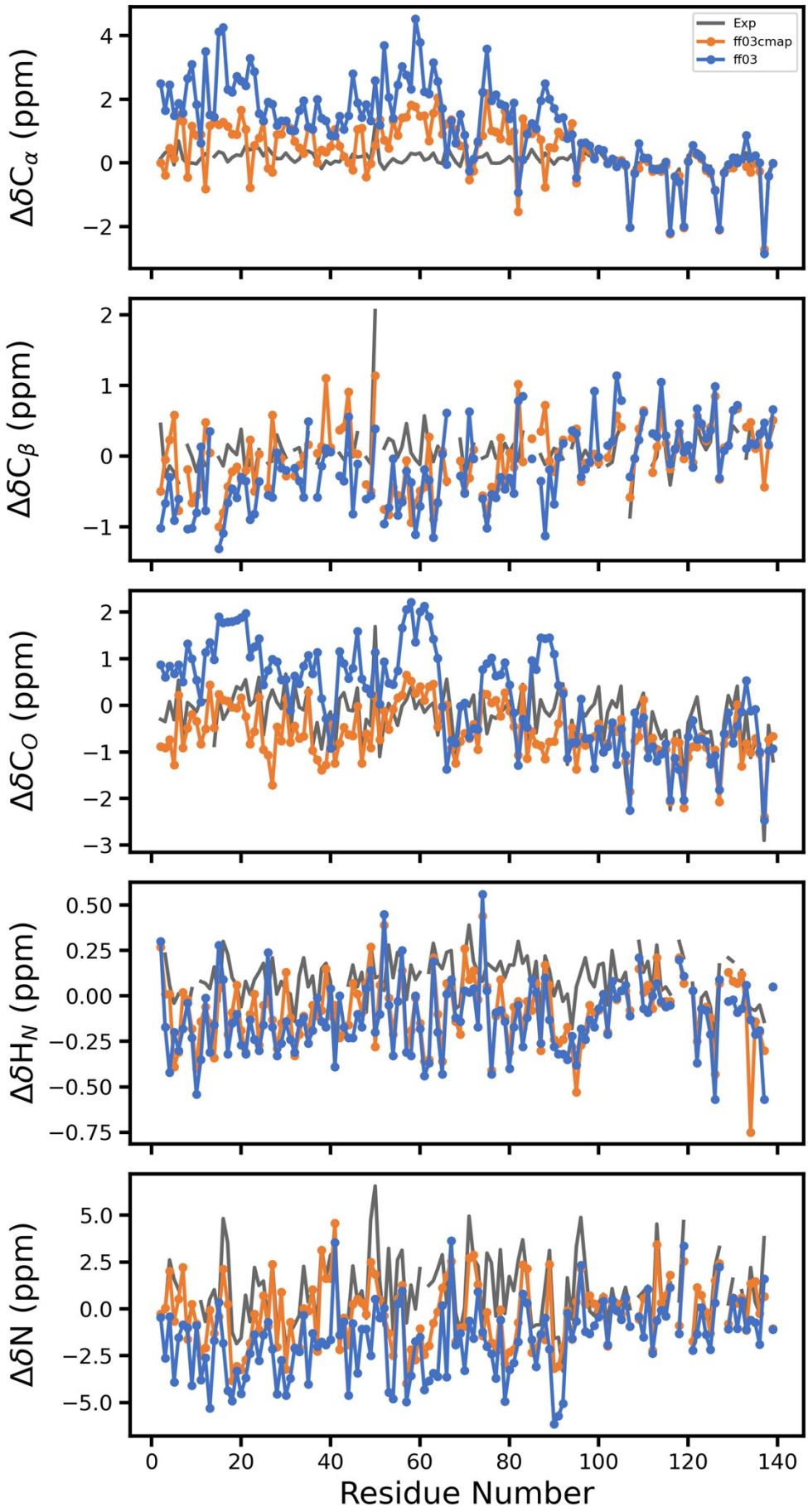

Exp
ff03cmap
ff03
ΔδCα (ppm)
ΔδCβ (ppm)
ΔδCO (ppm)
ΔδHN (ppm)
ΔδN (ppm)
Residue Number

Figure S9. Secondary chemical shifts of simulation and experimental data for synuclein. Simulated and experimental secondary chemical shifts for (A) Cα, (B) Cβ, (C) C, (D) N, and (E) HN

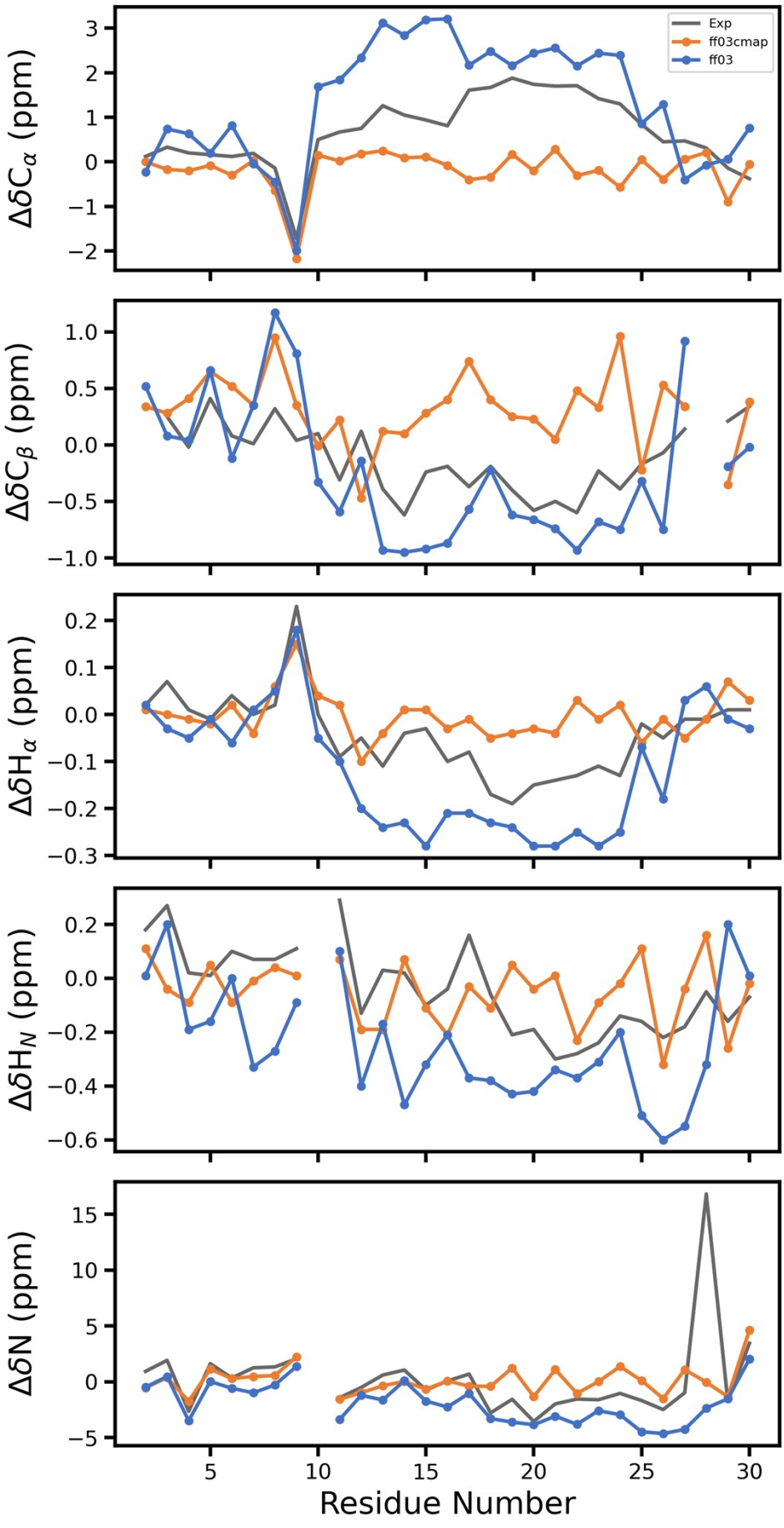

Exp
ff03cmap
ff03
ΔδCα (ppm)
ΔδCβ (ppm)
ΔδHα (ppm)
ΔδHN (ppm)
ΔδN (ppm)
Residue Number

Figure S10. Secondary chemical shifts of simulation and experimental data for MeVN. Simulated and experimental secondary chemical shifts for (A) Cα, (B) Cβ, (C) N, (D) Hα and (E) HN

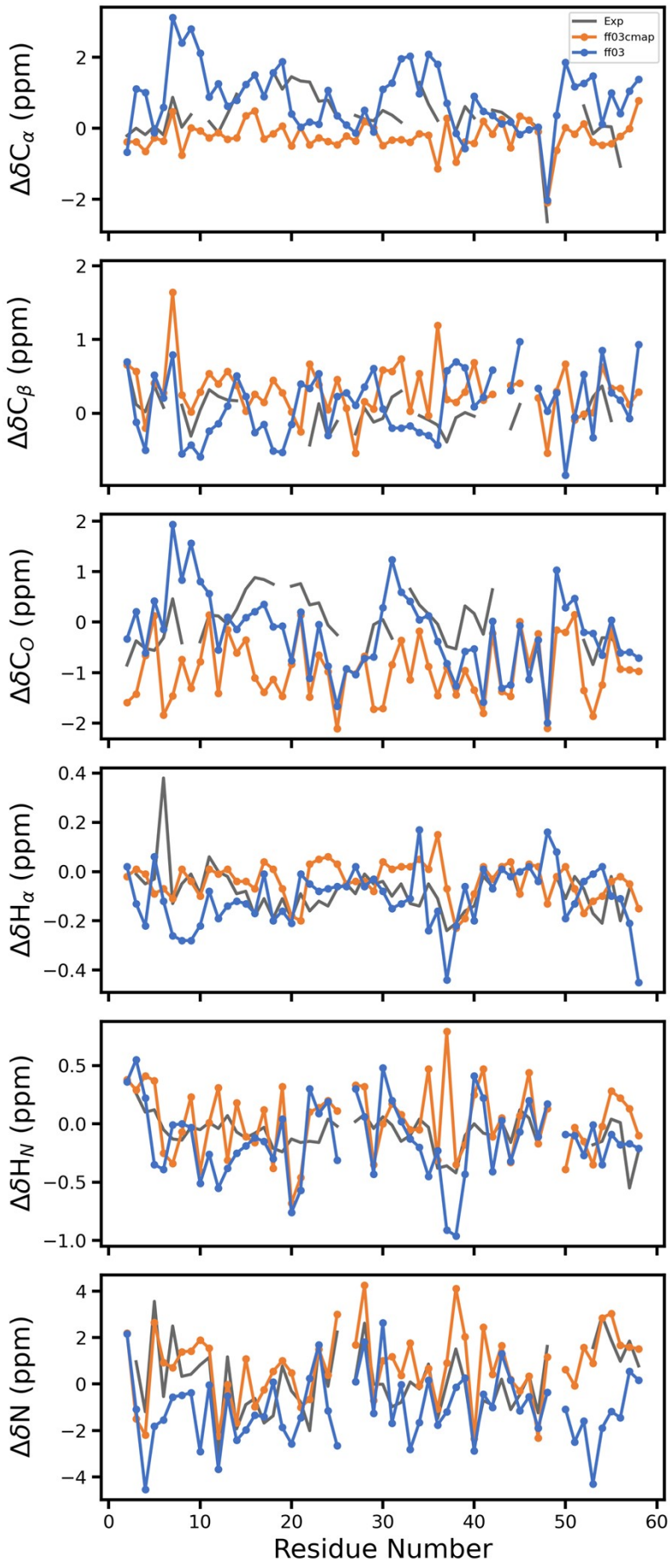

Exp
ff03cmap
ff03
ΔδCα (ppm)
ΔδCβ (ppm)
ΔδCO (ppm)
ΔδHα (ppm)
ΔδHN (ppm)
ΔδN (ppm)
Residue Number

Figure S11. Secondary chemical shifts of simulation and experimental data for drkN. Simulated and experimental secondary chemical shifts for (A) Cα, (B) Cβ, (C) C, (D) N, (E) Hα and (F) HN.

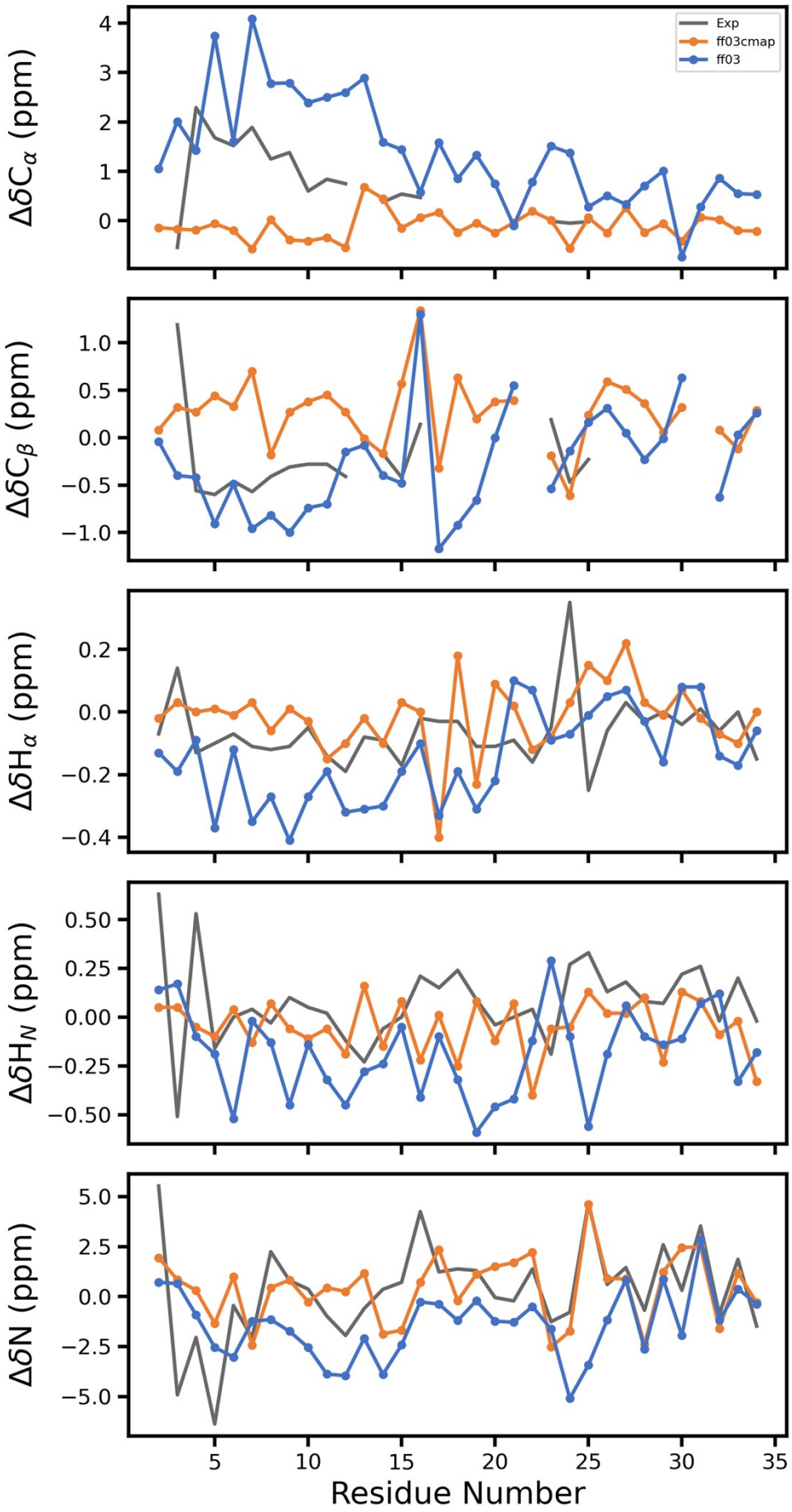

Exp
ff03cmap
ff03
ΔδCα (ppm)
ΔδCβ (ppm)
ΔδHα (ppm)
ΔδHN (ppm)
ΔδN (ppm)
Residue Number

Figure S12. Secondary chemical shifts of simulation and experimental data for hIAPP. Simulated and experimental secondary chemical shifts for (A) Cα, (B) Cβ, (C) N, (D) Hα and (E) HN.

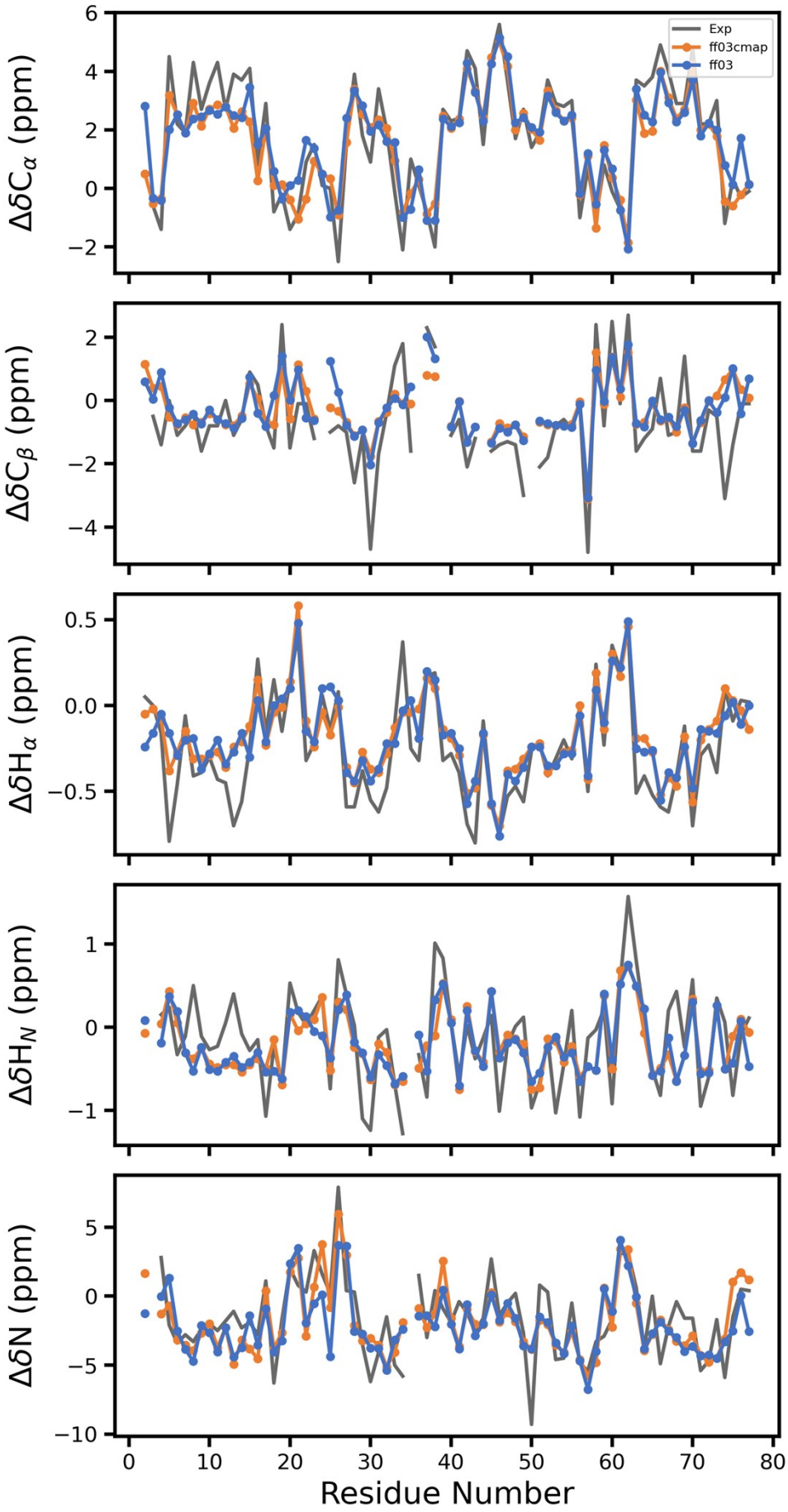

Exp
ff03cmap
ff03
ΔδCα (ppm)
ΔδCβ (ppm)
ΔδHα (ppm)
ΔδHN (ppm)
ΔδN (ppm)
Residue Number

Figure S13. Secondary chemical shifts of simulation and experimental data for 2krk. Simulated and experimental secondary chemical shifts for (A) Cα, (B) Cβ, (C) N, (D) Hα and (E) HN.

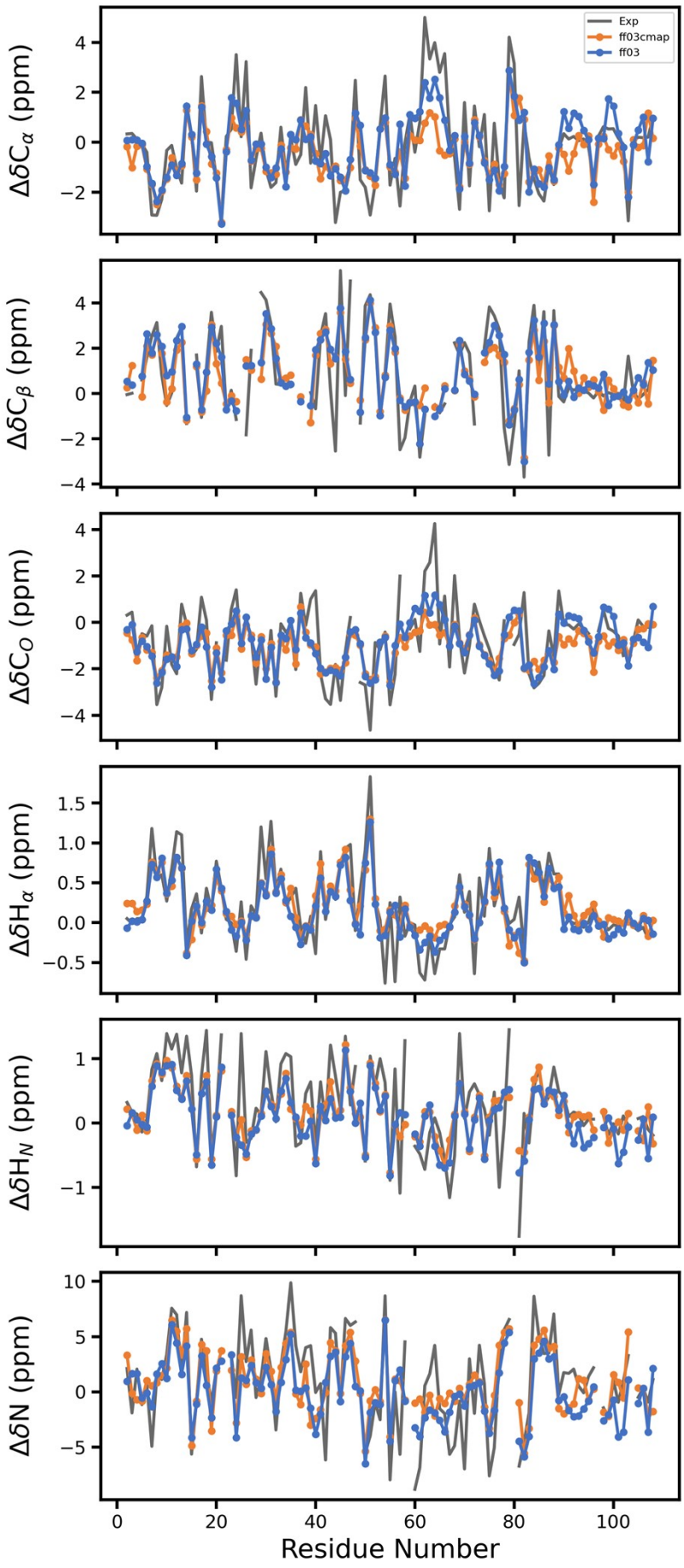
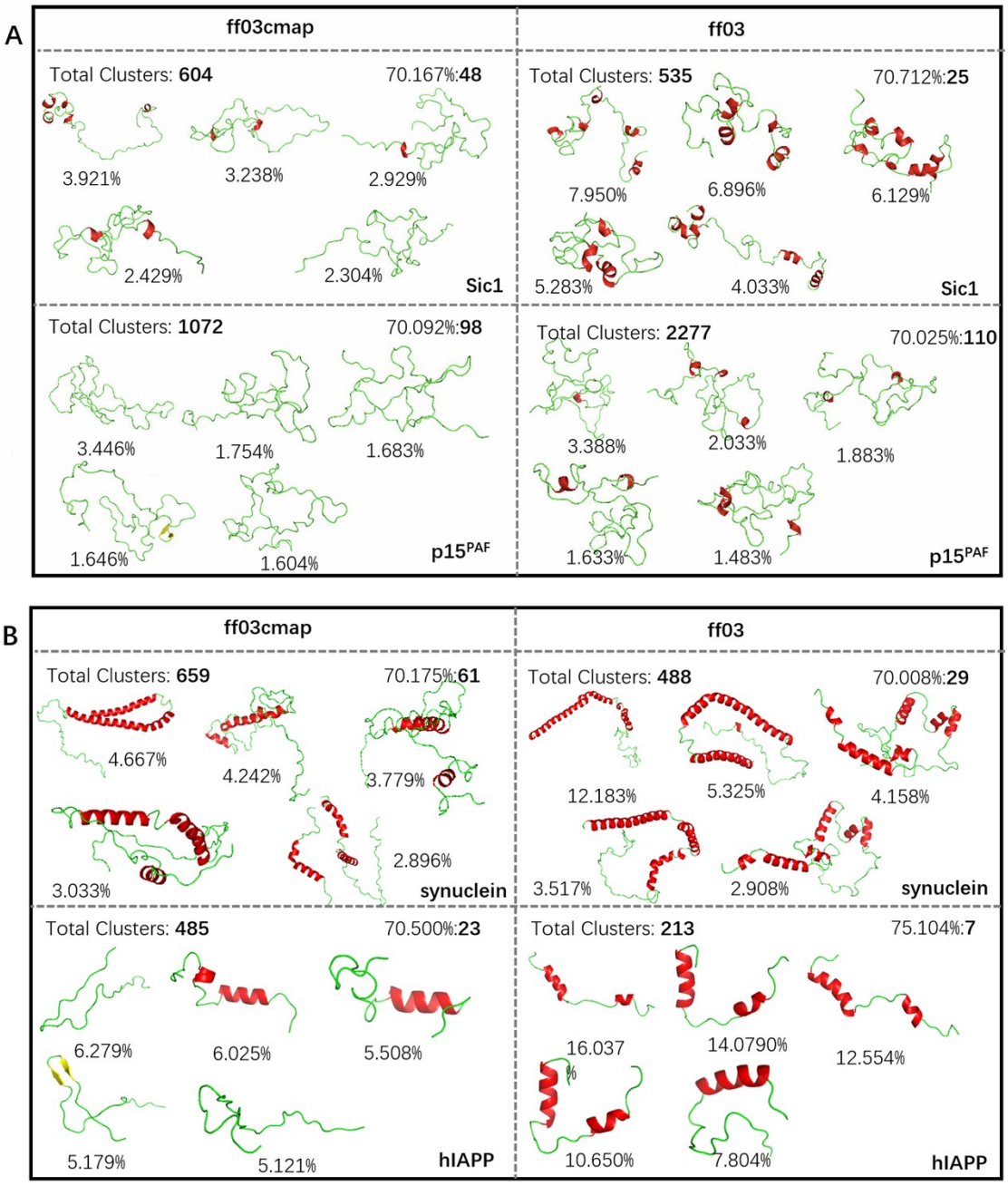

Exp
ff03cmap
ff03
ΔδCα (ppm)
ΔδCβ (ppm)
ΔδCO (ppm)
ΔδHα (ppm)
ΔδHN (ppm)
ΔδN (ppm)
Residue Number

Figure S14. Secondary chemical shifts of simulation and experimental data for 2kpp. Simulated and experimental secondary chemical shifts for (A) Cα, (B) Cβ, (C) C, (D) N, (E) Hα and (F) HN.



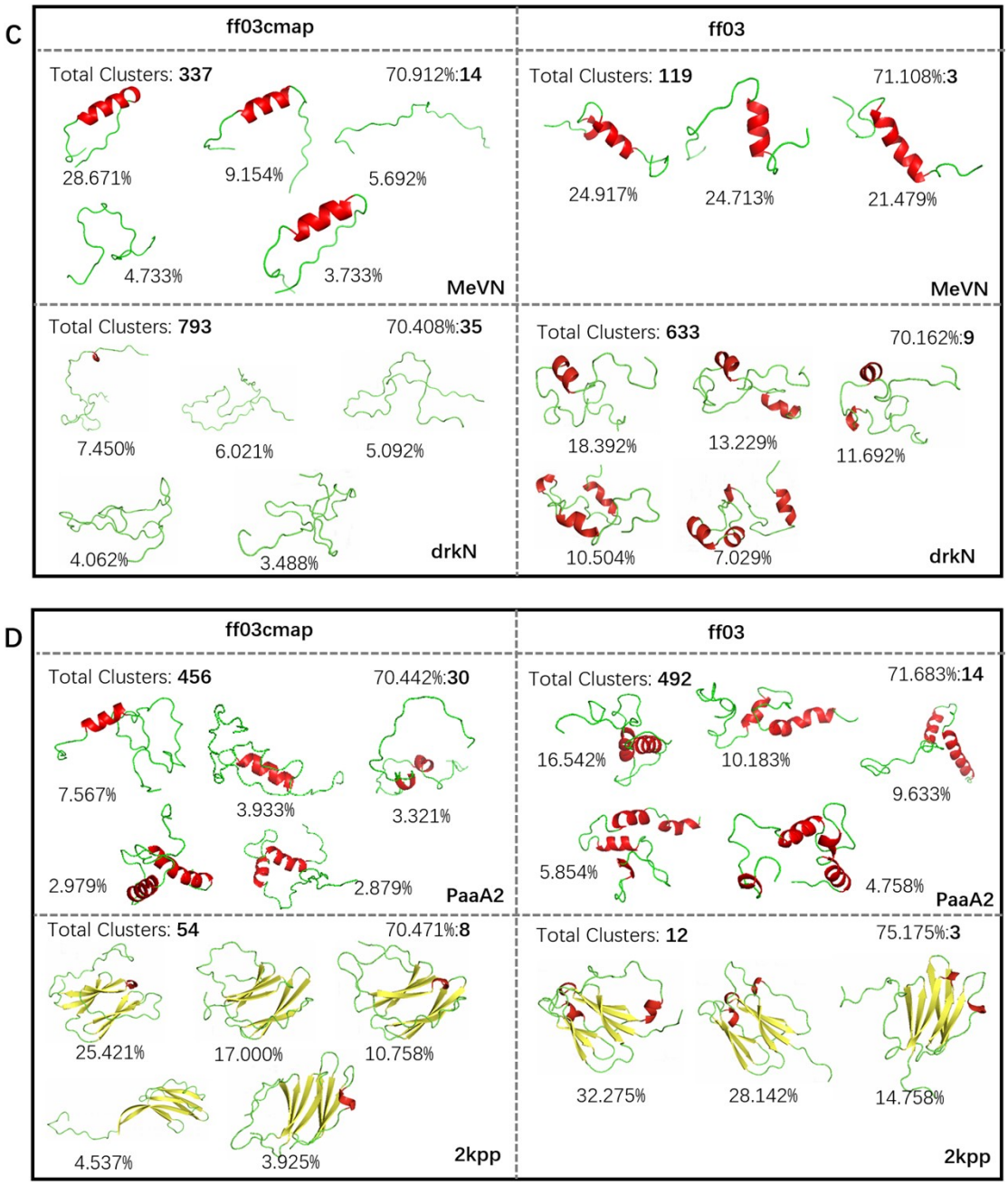


Figure S15. Representative conformers of the top-five clusters between ff03cmap and ff03 force fields for ten proteins. (A) Clusters of disordered proteins Sic1 and p15^{PAF}. (B) Clusters of disordered proteins synuclein and hIAPP. (C) Clusters of disordered proteins MeVN and drkN. (D) Clusters of ordered proteins PaaA2 and 2kpp. The conformations on the left were simulated by ff03cmap, and those on the right were simulated by ff03. The numbers below the conformations show the populations for each cluster. The numbers on the conformations show the total number of cluster conformations and the total number of cluster conformations that account for the top 70 percent.

Table S1. Simulation condition for all tested protein system.

Protein	force filed	no.of traj.	Simulation time(ns)	Temperature/K	ions	waters	BMRB accession no.
drkN	ff03cmap	3	400	298	7	7325	25501
	ff03	3	400				
Sic1	ff03cmap	3	400	298	68	25096	16657
	ff03	3	400				
PaaA2	ff03cmap	3	400	298	86	9624	18840
	ff03	3	400				
p15PAF	ff03cmap	3	400	298	17	18959	19332
	ff03	3	400				
Histain5	ff03cmap	3	400	303	0	3990	-
	ff03	3	400				
synuclein	ff03cmap	3	400	273	87	32236	19337
	ff03	3	400				
hIAPP	ff03cmap	3	400	278	0	3522	34069
	ff03	3	400				
MeVN	ff03cmap	3	400	293	5	2797	6566
	ff03	3	400				
2krk	ff03cmap	3	400	298	0	7514	16640
	ff03	3	400				
2kpp	ff03cmap	3	400	298	33	9434	16563
	ff03	3	400				