Supporting Information for Are Protein Force Fields Getting Better? A Systematic Benchmark on 524 Diverse NMR Measurements

Kyle A. Beauchamp,[†] Yu-Shan Lin,[‡] Rhiju Das,[¶] and Vijay S. Pande[*,‡]

Biophysics Program, Chemistry Department, Stanford University, Stanford, CA, and Biochemistry Department, Stanford University, Stanford, CA

E-mail: pande@stanford.edu

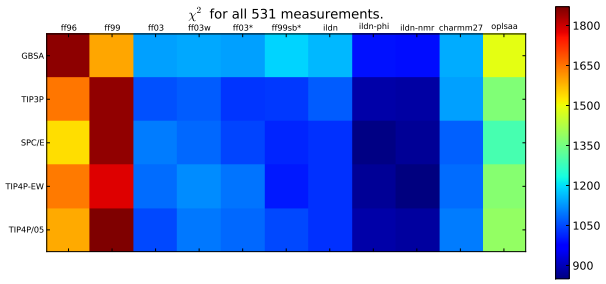
[*]To whom correspondence should be addressed
[†]Biophysics Program
[‡]Chemistry Department
[¶]Biochemistry Department

Supporting Information Available

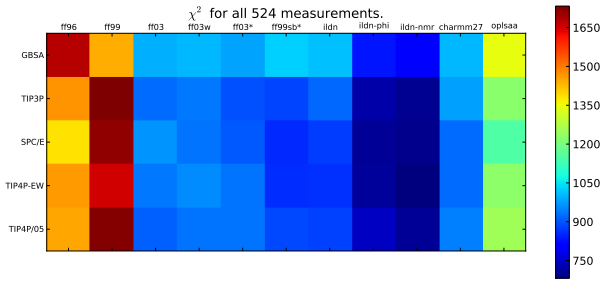
1 Table of Contents

2 Alternative J Coupling and Chemical Shift Models

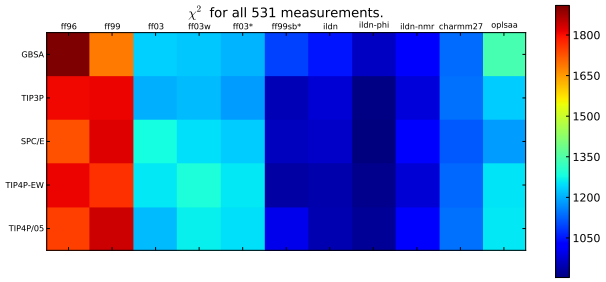
Because several published models exist for calculating J couplings and chemical shifts, we repeated the overall χ^2 analysis using alternative models (Figure S1). For J coupling calculations, we considered two models, which we label 1997[1] and 1999.[2] For chemical shift calculations, we considered both SPARTA+[3] and ShiftX.[4] We found that the combination of SPARTA+ and the 1997 J coupling parameterization gave the lowest values of χ^2; that choice of models is used in the main text. All four model combinations suggest highest accuracy by either ff99sb-ildn-nmr or ff99sb-ildn-phi. This suggests that the present results are robust to differences in the calculation of NMR observables. In Table S2, we list all Karplus parameters used in this work.



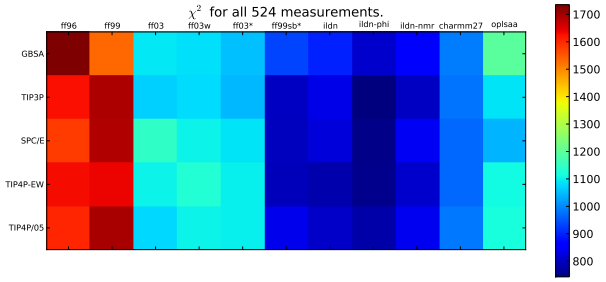
(a)



(b)

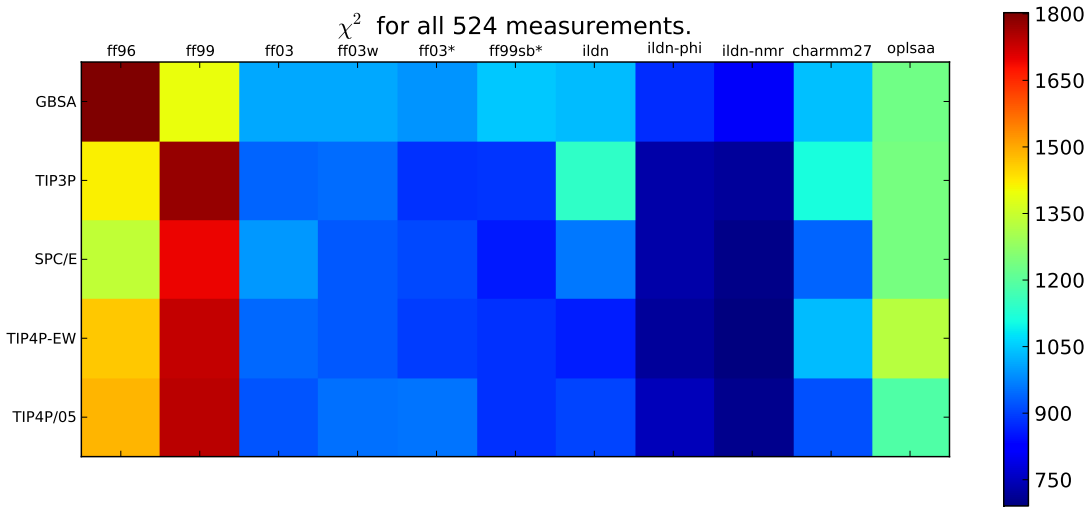


(c)

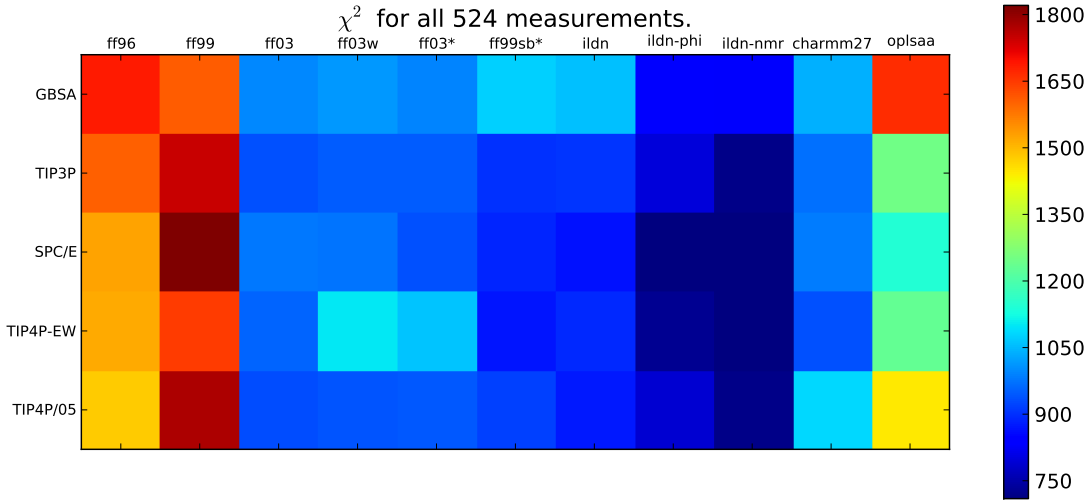


(d)

Figure S1: The overall χ^2 quantifies the agreement with all 524 experimental measurements. (a) shows the 1997 J coupling model with the ShiftX chemical shift model. (b) shows the 1997 J coupling model with the SPARTA+ chemical shift model. (c) shows the 1999 J coupling model with the ShiftX chemical shift model. (d) shows the 1999 J coupling model with the SPARTA+ chemical shift model. To ensure consistency between the four panels, we used the SPARTA+ and 1997 error estimates. The total number of comparisons depends on the choice of chemical shift model because SPARTA+ is unable to estimate shifts on terminal atoms.

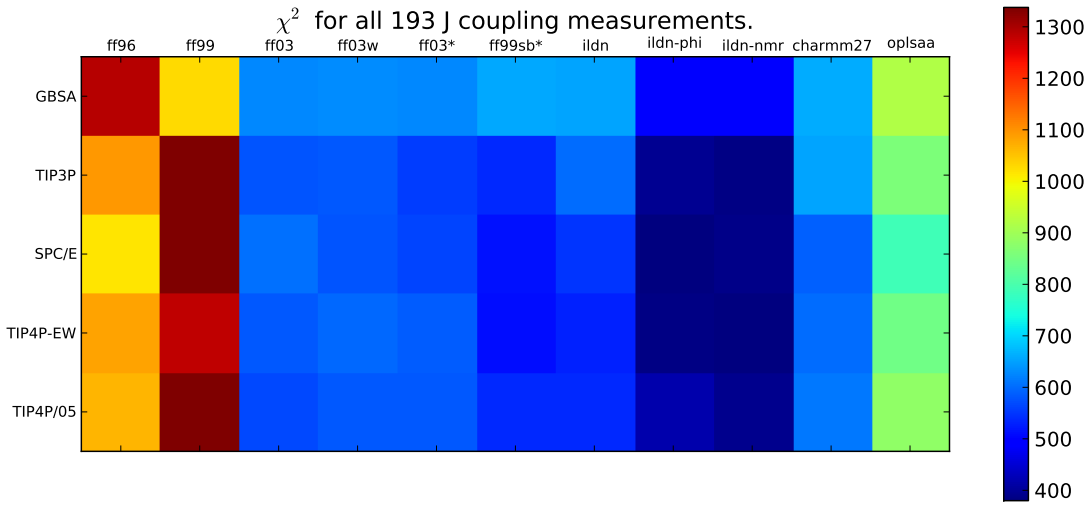


(a)

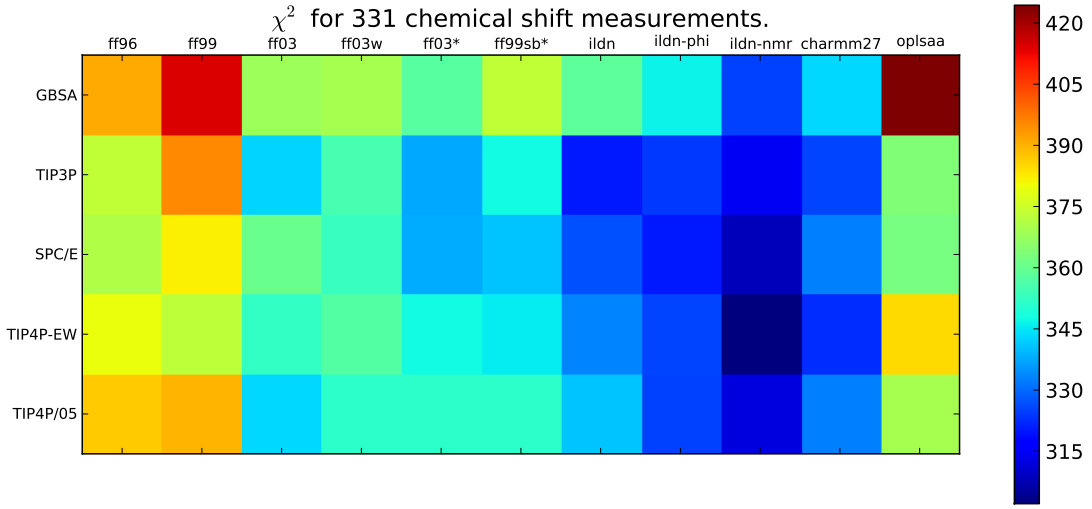


(b)

Figure S2: The overall χ^2 was calculated using two sets of independent simulations. The key conclusions are robust to the statistical uncertainty present in the data. In particular, sampling leads to approximately 6% uncertainty in the values of χ^2. In contrast, the poorly-performing forcefields show χ^2 values that are 20-100% larger than the well-performing ildn-phi and ildn-nmr force fields.



(a)



(b)

Figure S3: The overall χ^2 was calculated for J coupling (a) and chemical shift (b) experiments. The key conclusions of this work do not depend on which experiment is used as benchmark.

3 Raw Comparisons, Sorted by Experiment

Here, we show the raw results of each class of experiment. Simulation results are shown for the TIP4P-EW water model and the ff99sb-ildn, ff99sb-ildn-phi, ff03w, ff99sb-ildn-nmr, and ff99 force fields. Error bars on experimental results are the RMS errors in the J coupling and chemical shift models; these errors are much larger than the reported experimental uncertainties. Error bars on simulated quantities are taken to be the standard deviation estimated from comparing the two independent datasets.

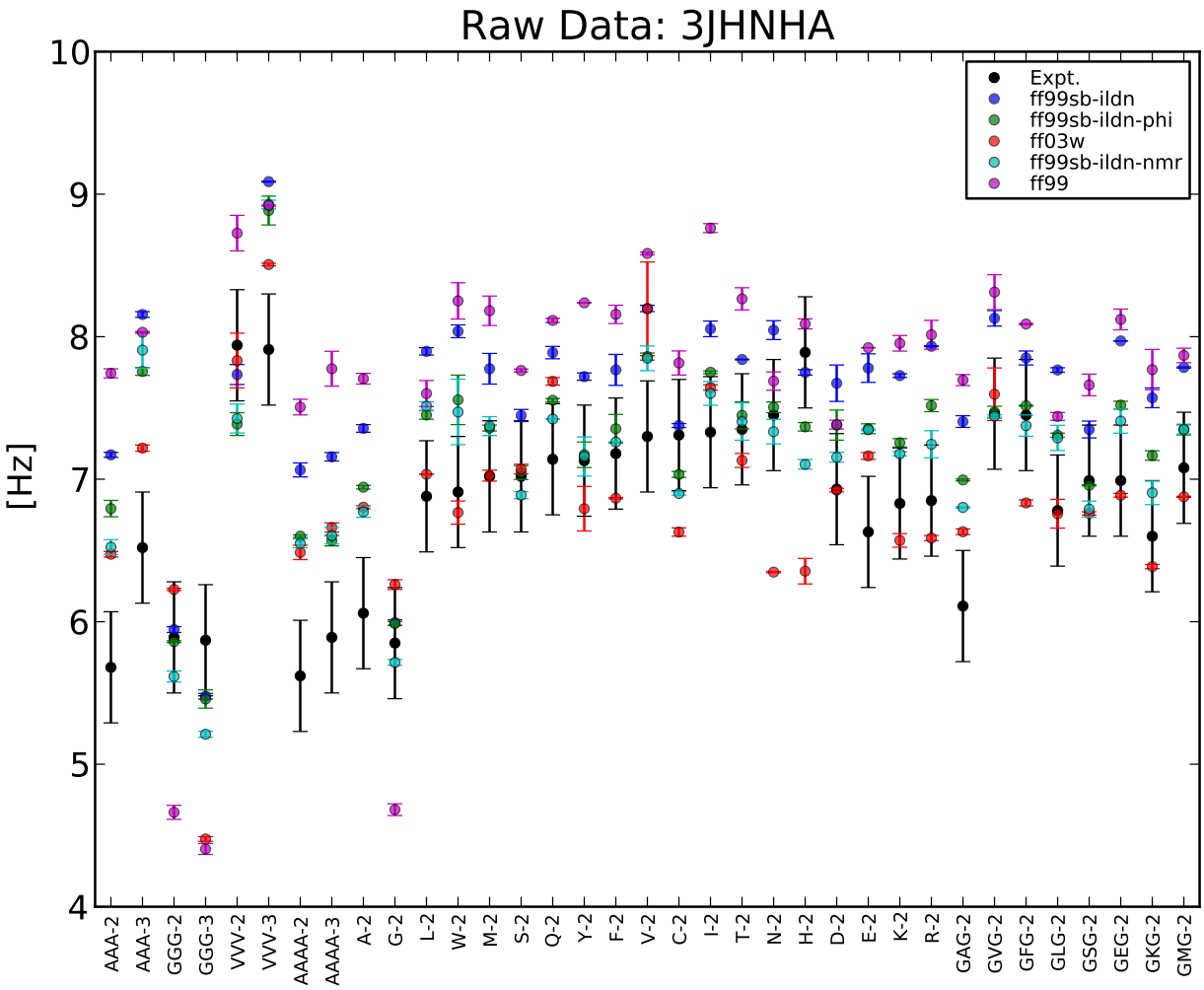


Figure S4: All $^3JH_NH_\alpha$ experiments.

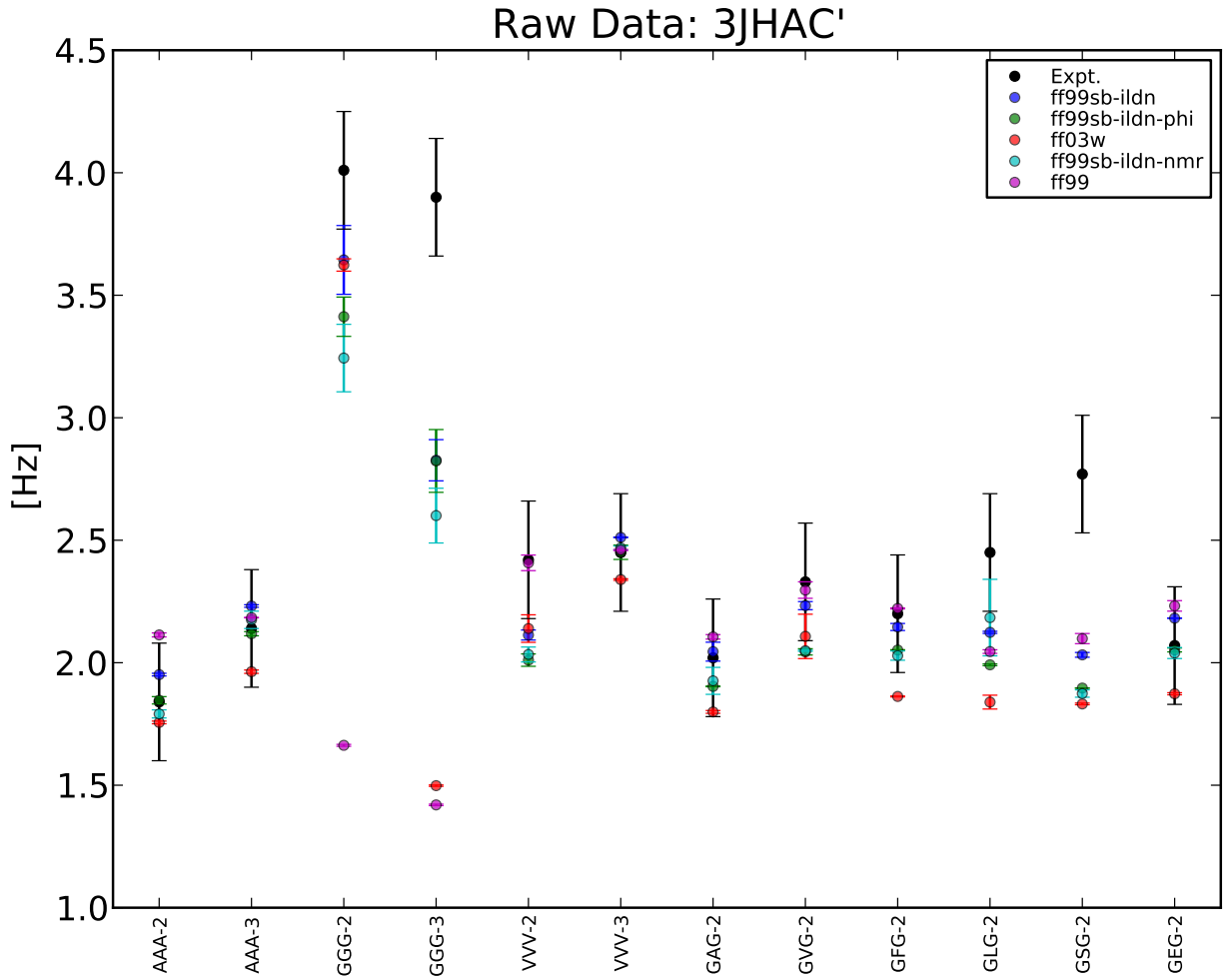


Figure S5: All $^{3}J(H^{\alpha}C')$ experiments.

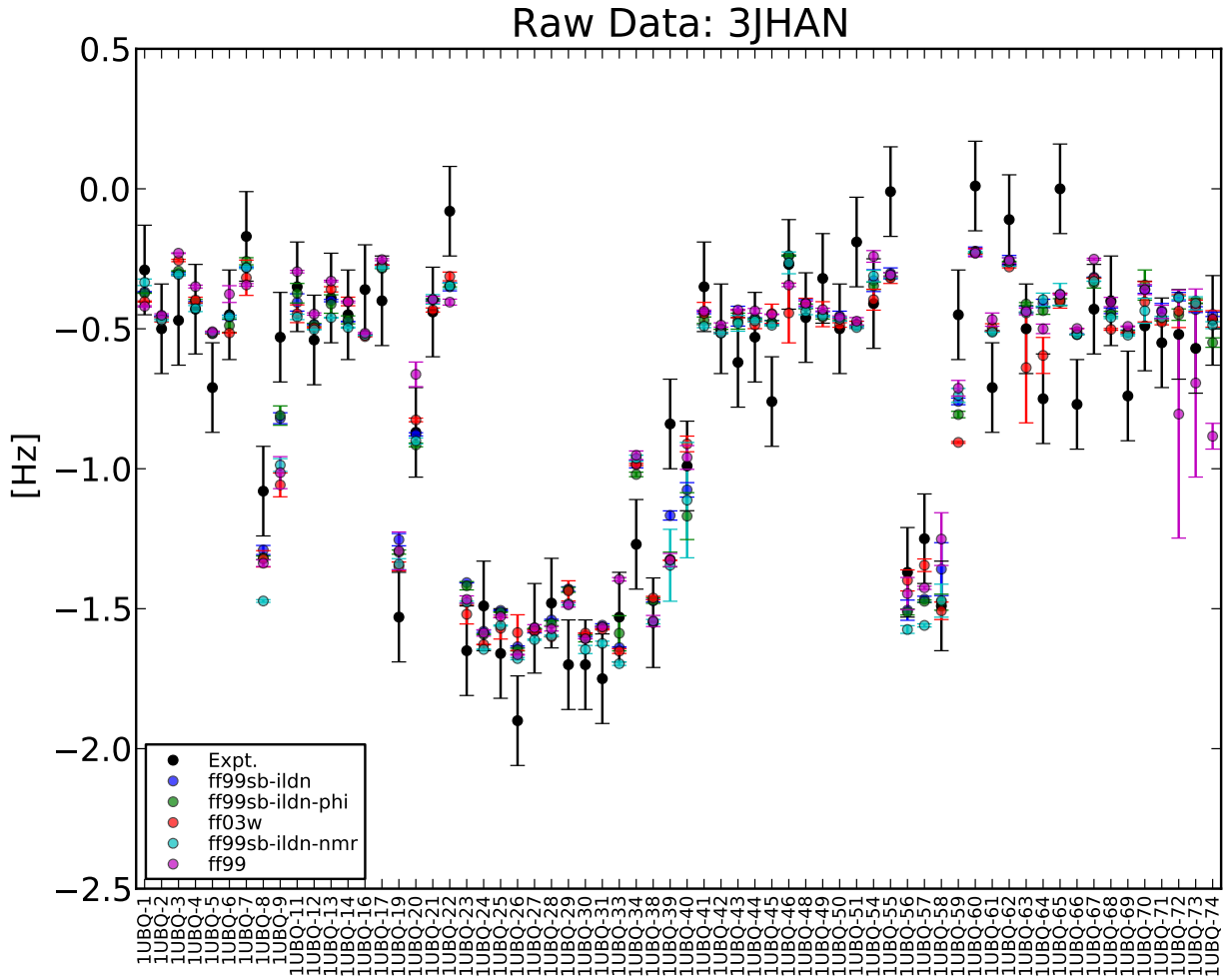


Figure S6: All $^{3}J(H^{\alpha}N)$ experiments.

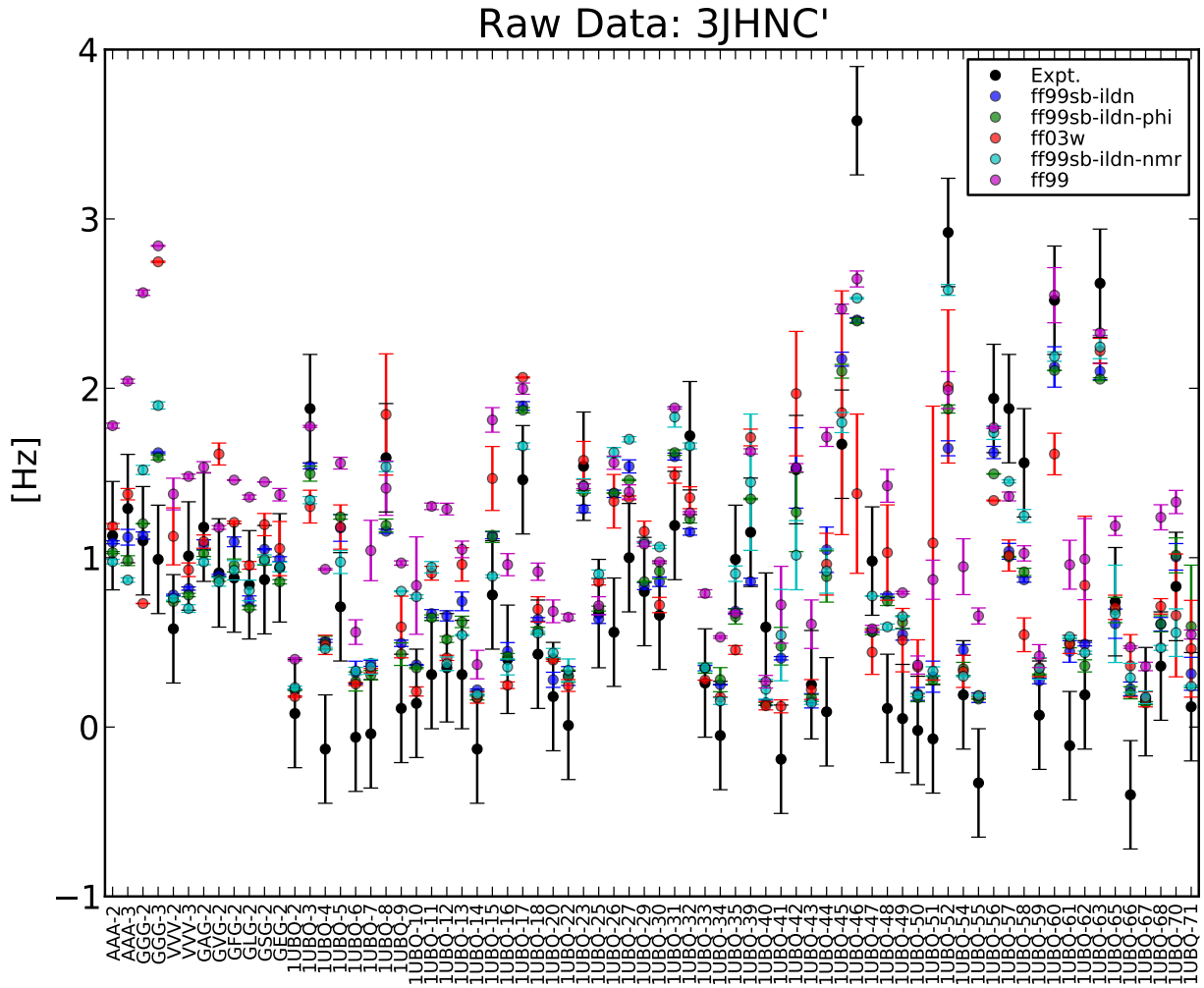


Figure S7: All $^{3}J(H^{N}C')$ experiments.

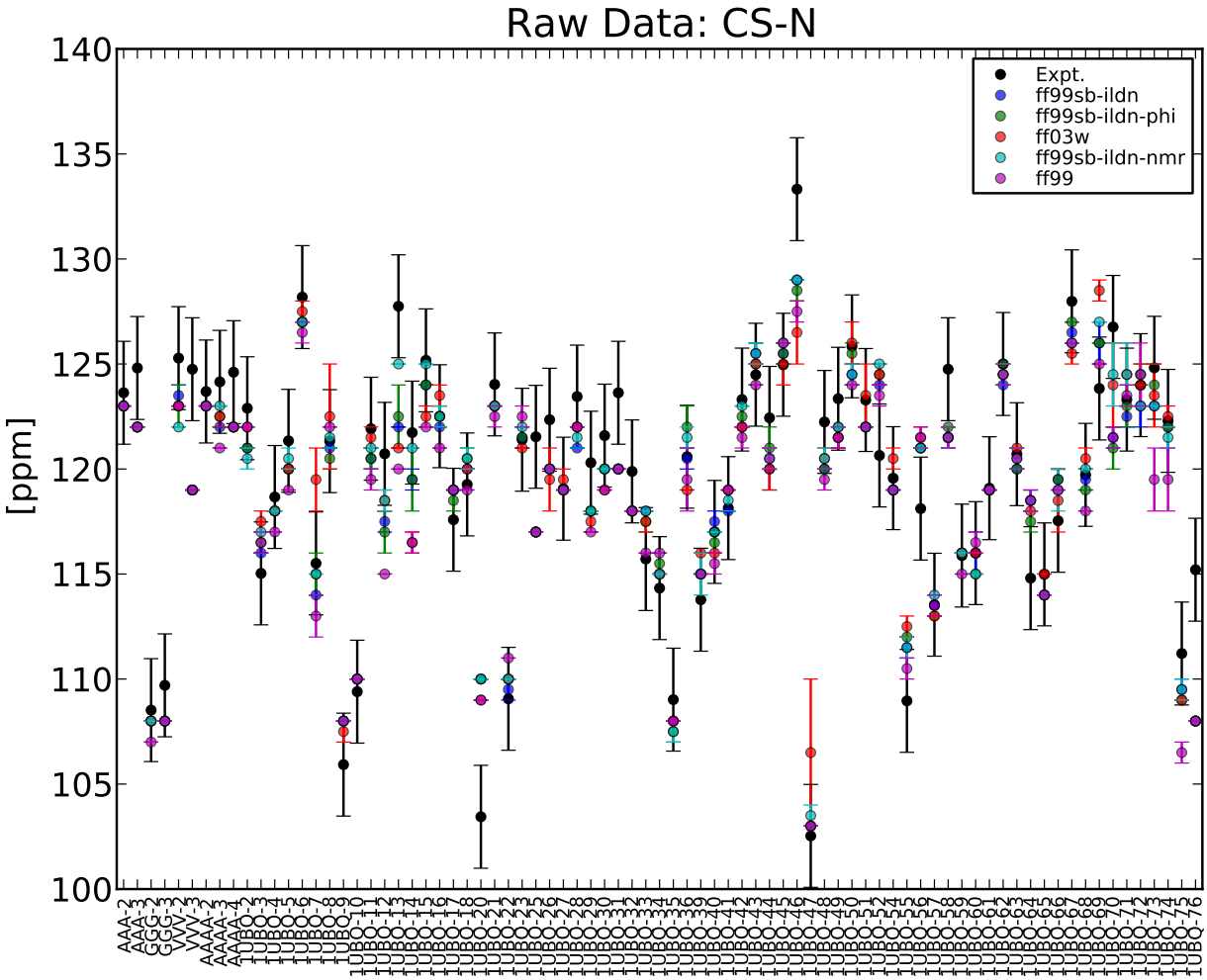


Figure S8: All CS-N experiments.

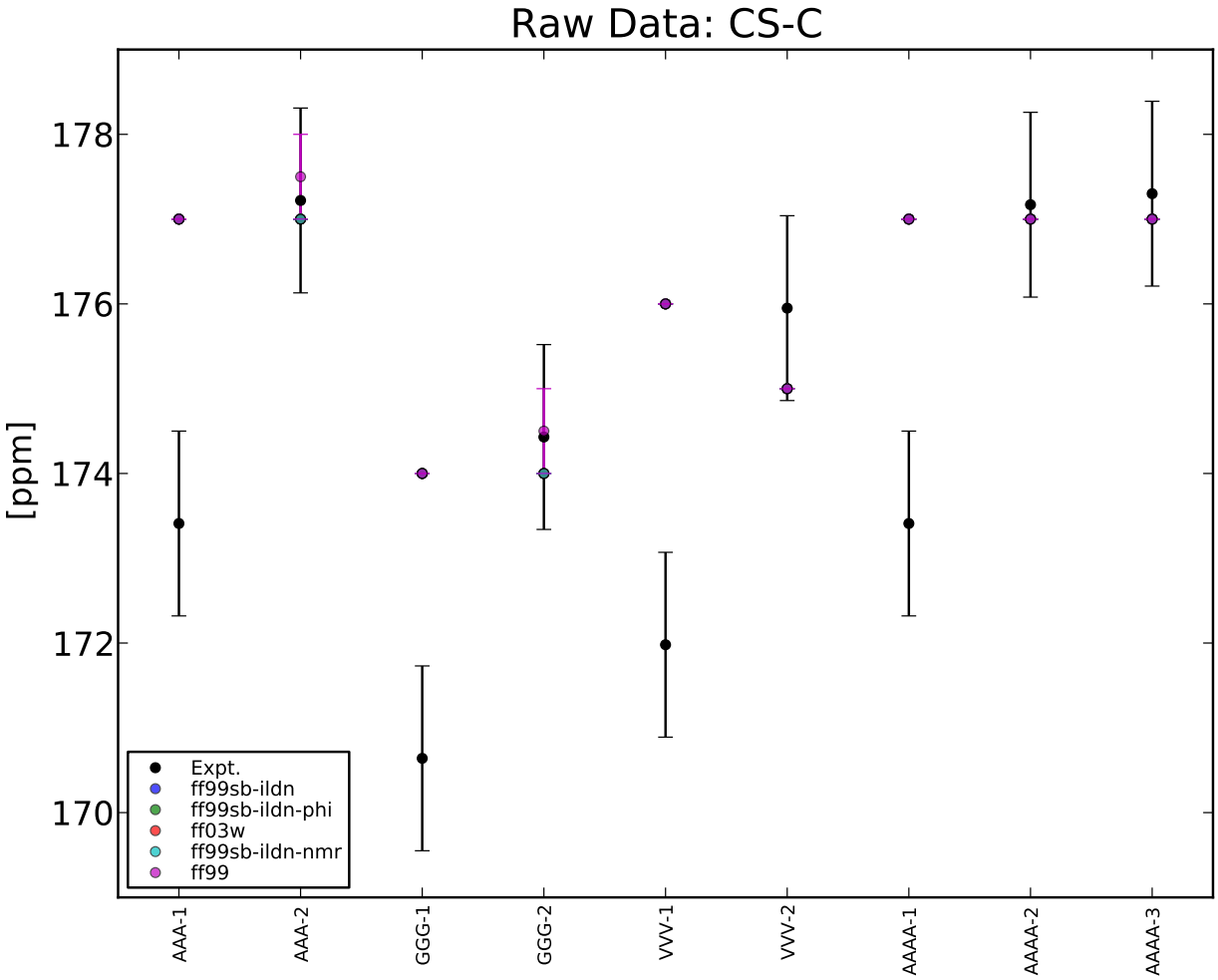


Figure S9: All CS-C experiments.

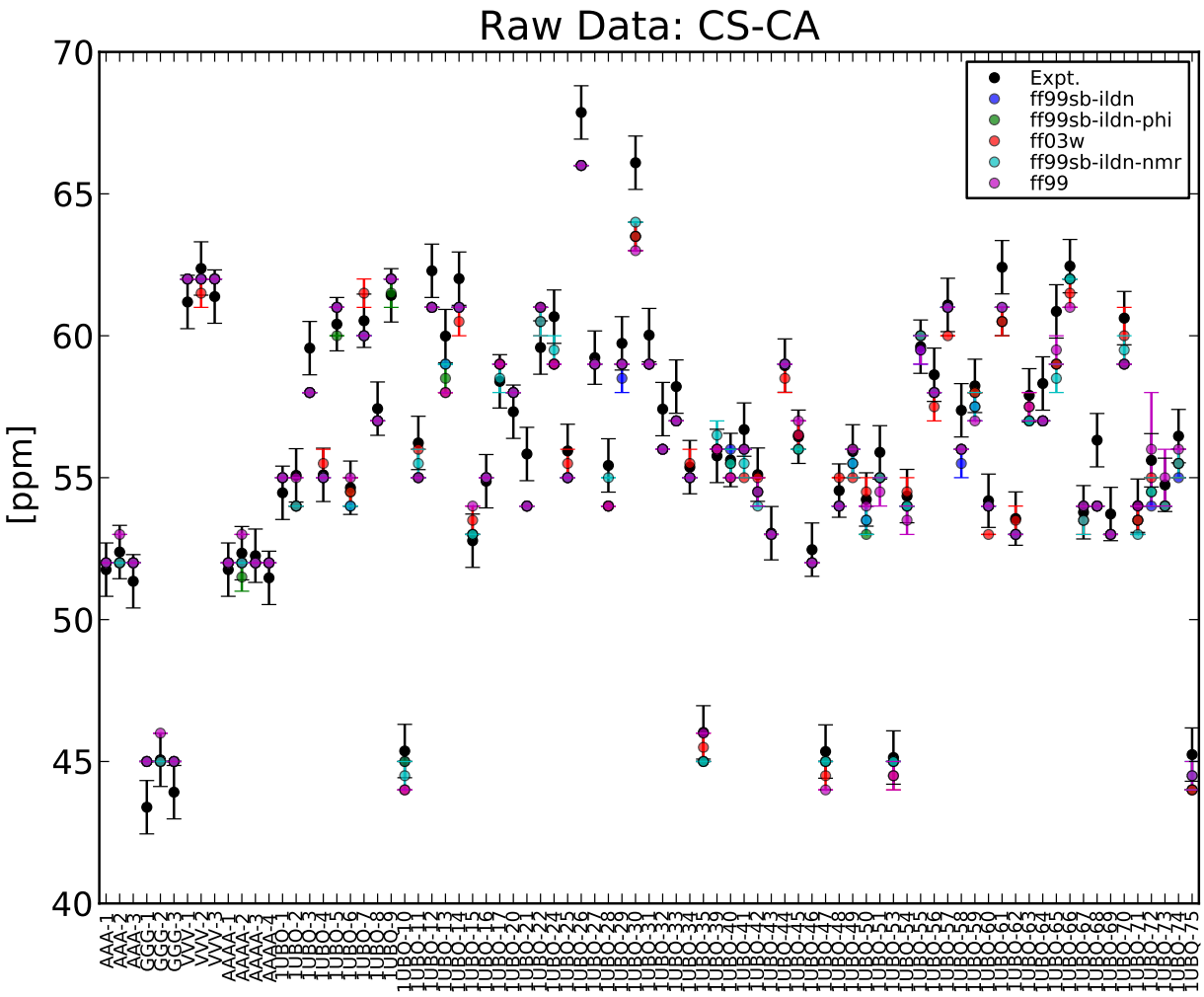


Figure S10: All CS-Cα experiments.

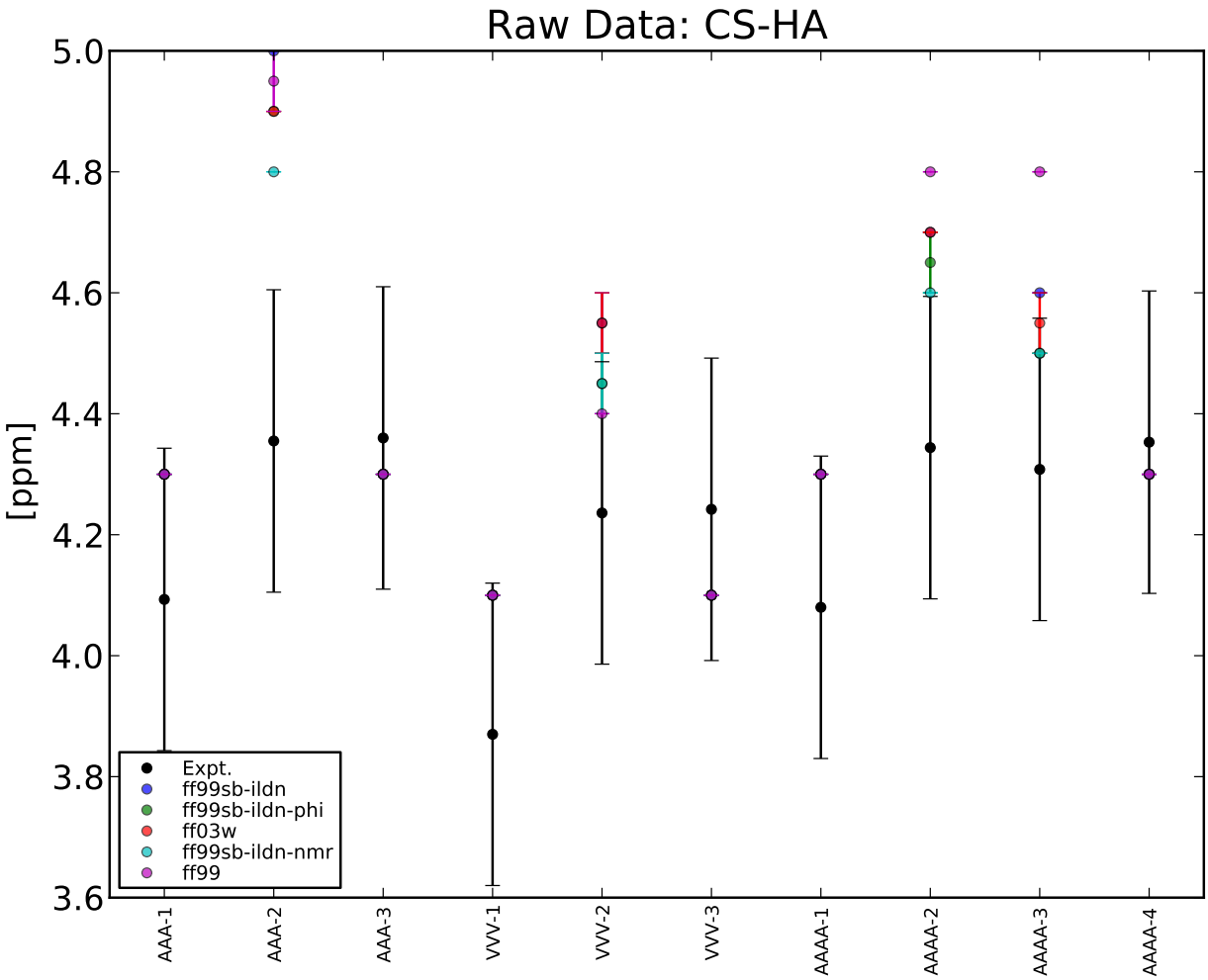


Figure S11: All CS-Hα experiments.

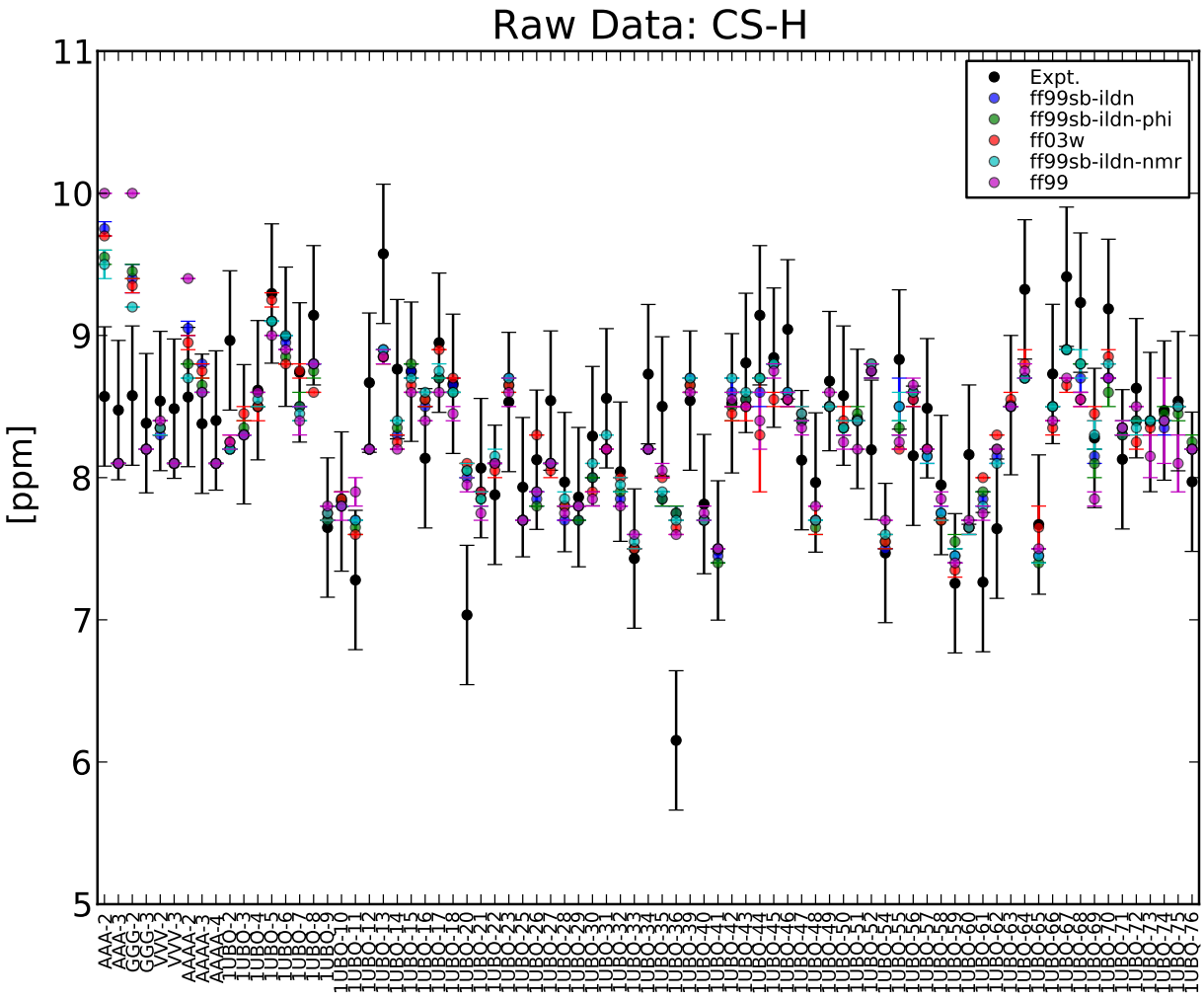


Figure S12: All CS-H experiments.

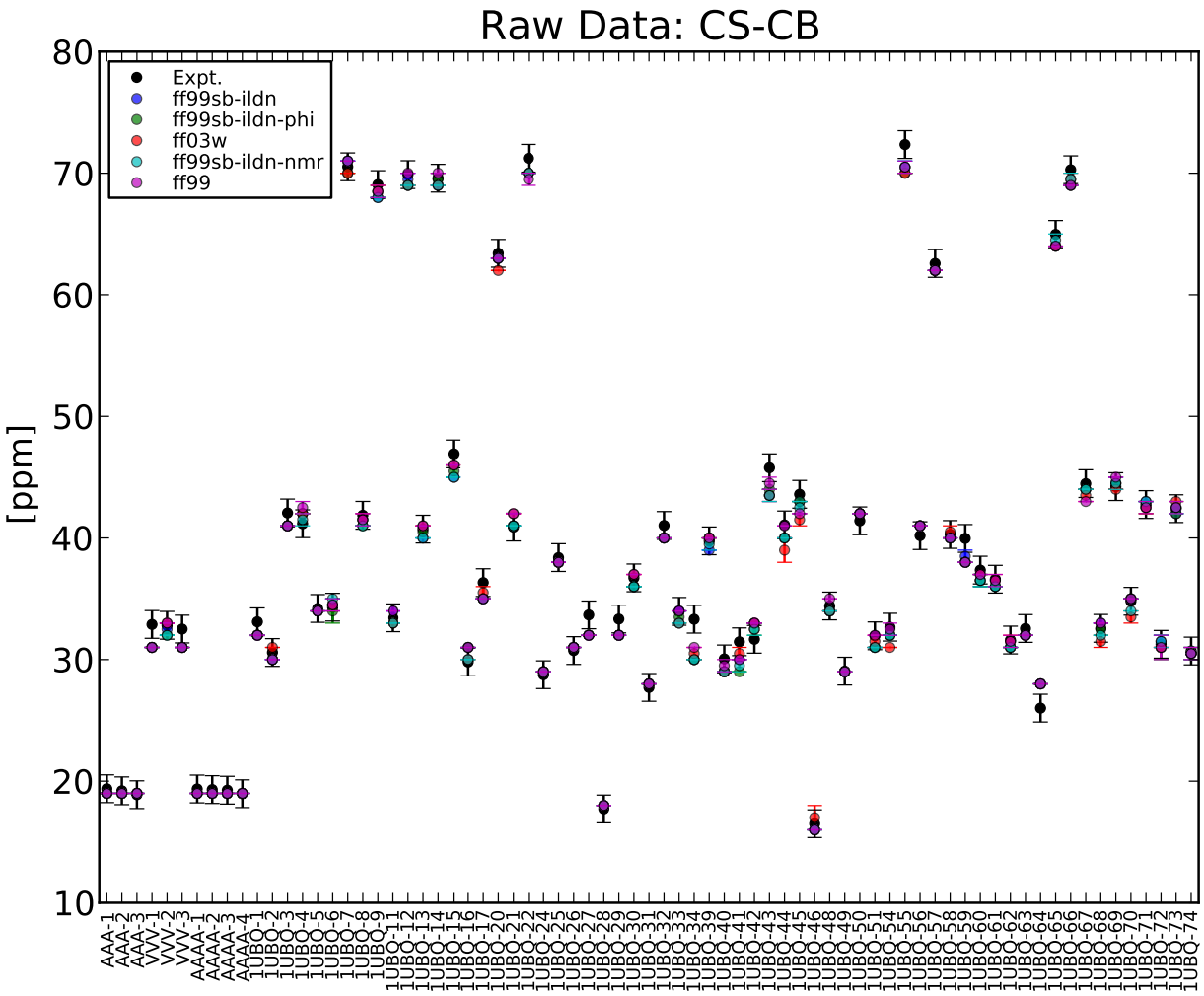


Figure S13: All CS-Cβ experiments.

4 RMS Errors, Sorted by Experiment

Here, we show the RMS errors for each experiment class.

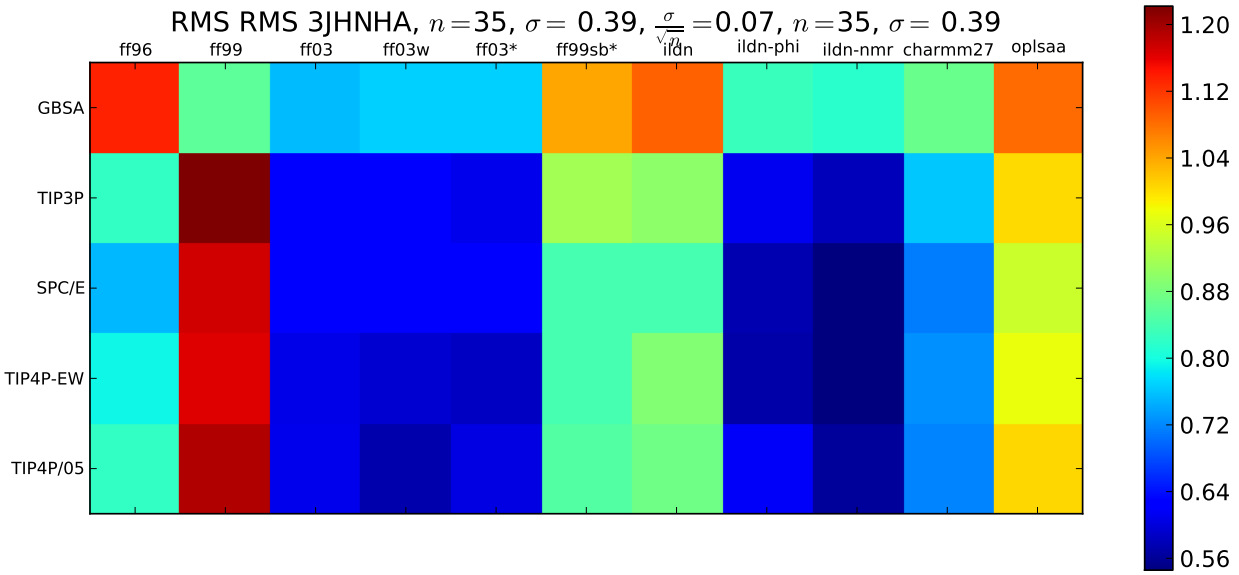


Figure S14: RMS errors of ${}^3JH_NH_\alpha$ experiments.

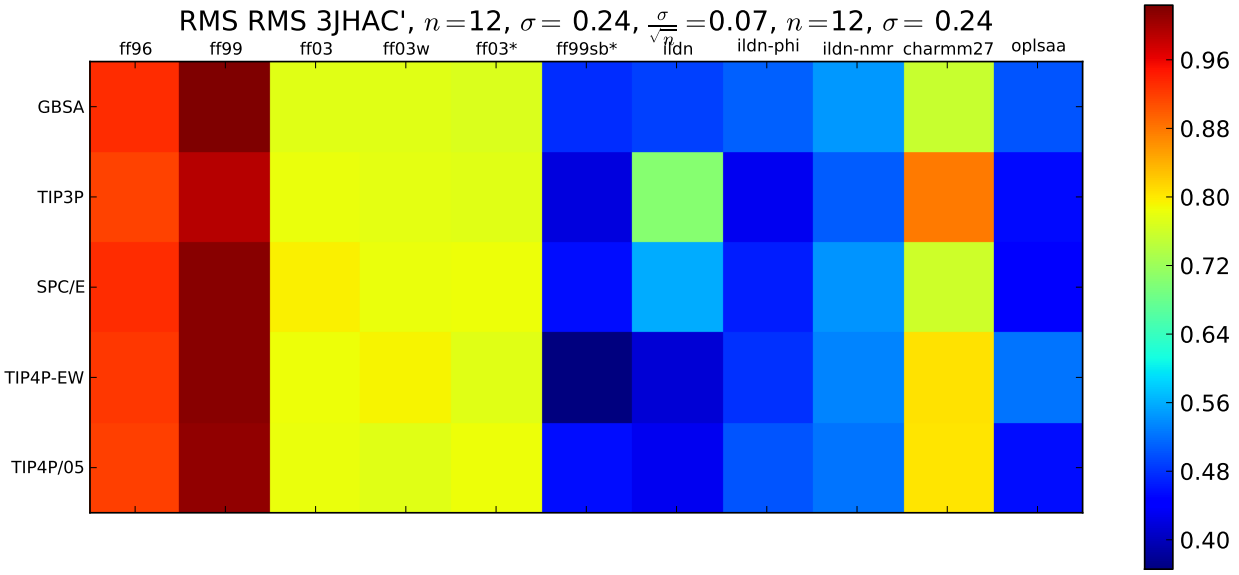


Figure S15: RMS errors of ${}^3J(H^\alpha C')$ experiments.

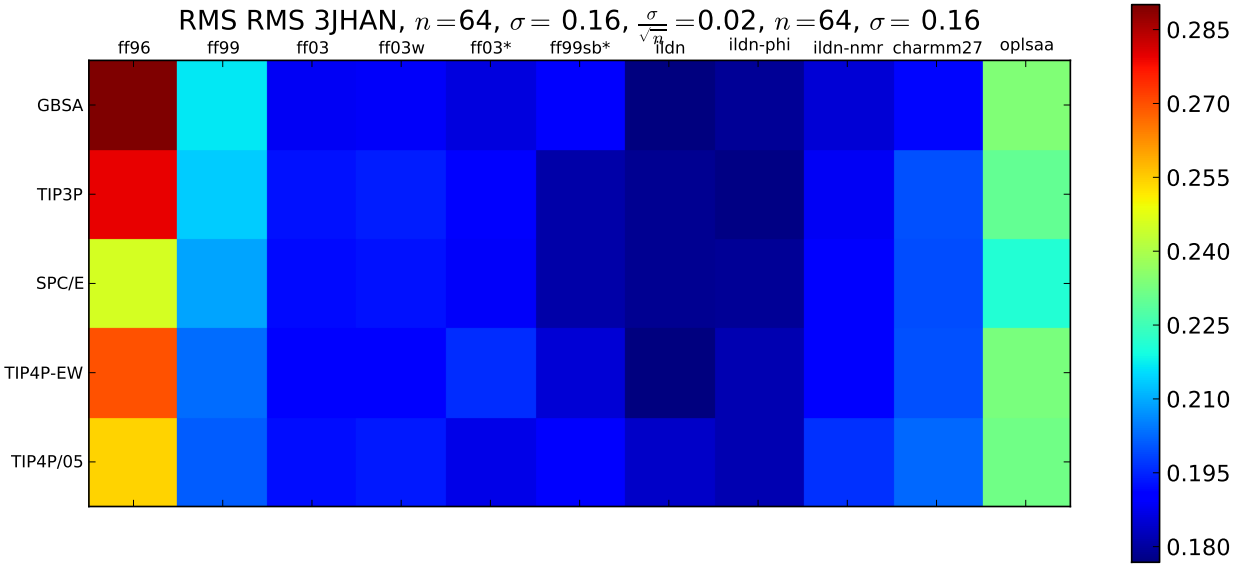


Figure S16: RMS errors of ${}^3J(H^\alpha N)$ experiments.

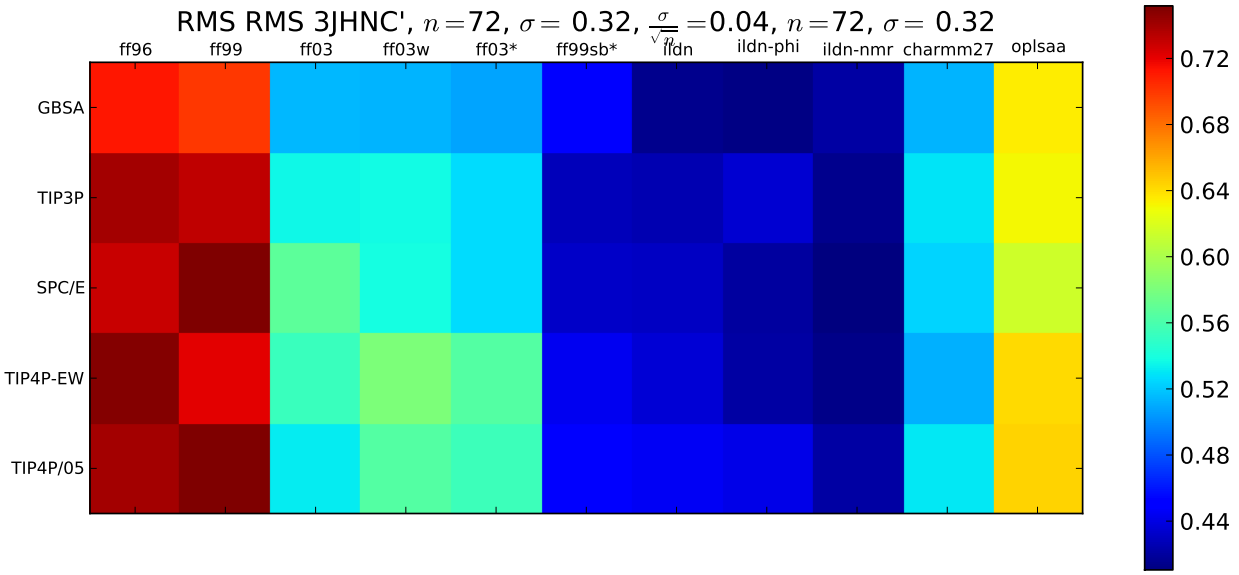


Figure S17: RMS errors of ${}^3J(H^NC')$ experiments.

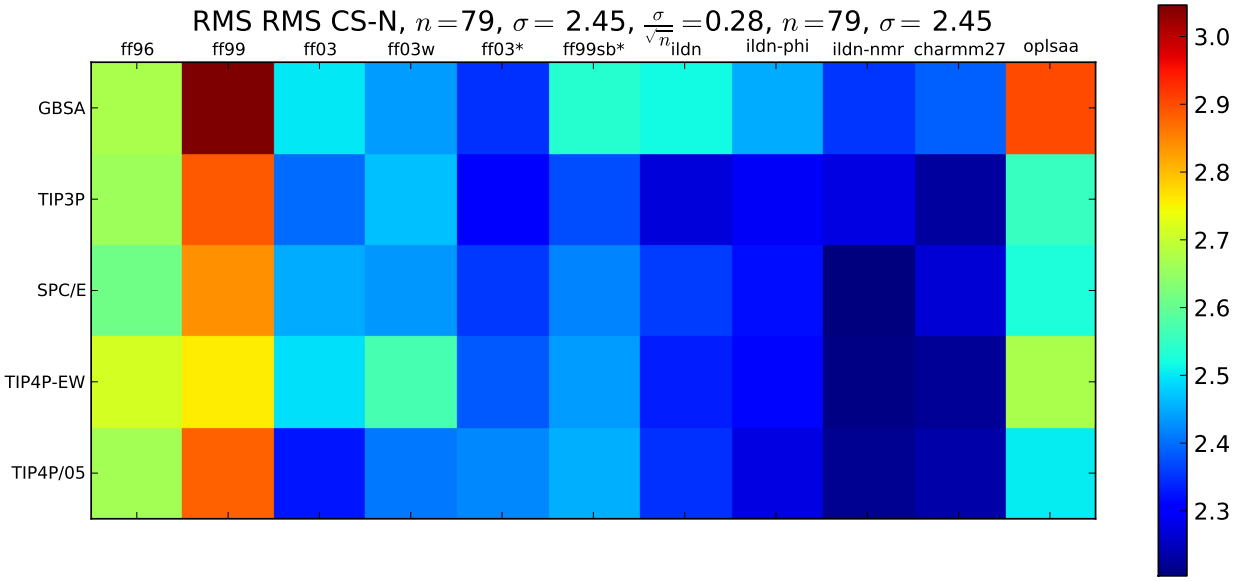


Figure S18: RMS errors of CS-N experiments.

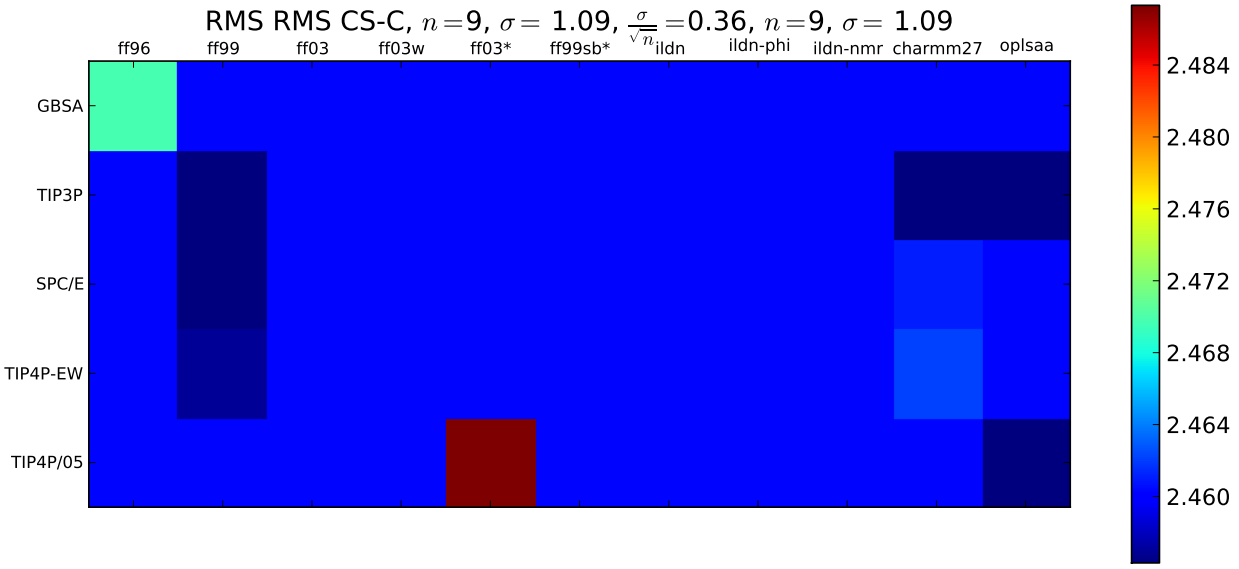


Figure S19: RMS errors of CS-C experiments.

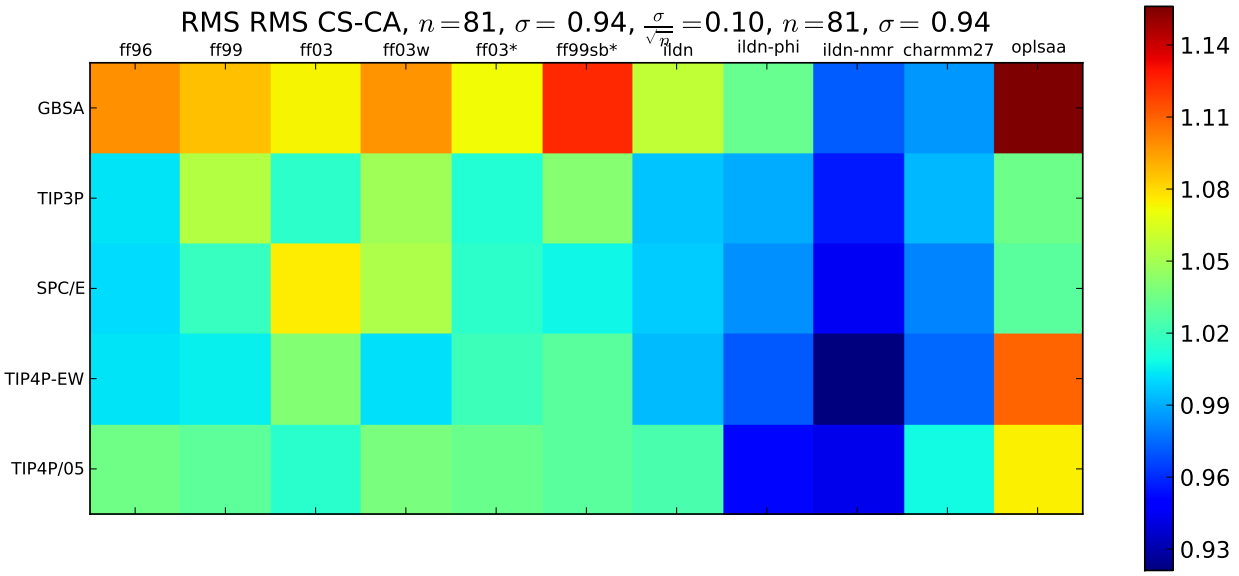


Figure S20: RMS errors of CS-Cα experiments.

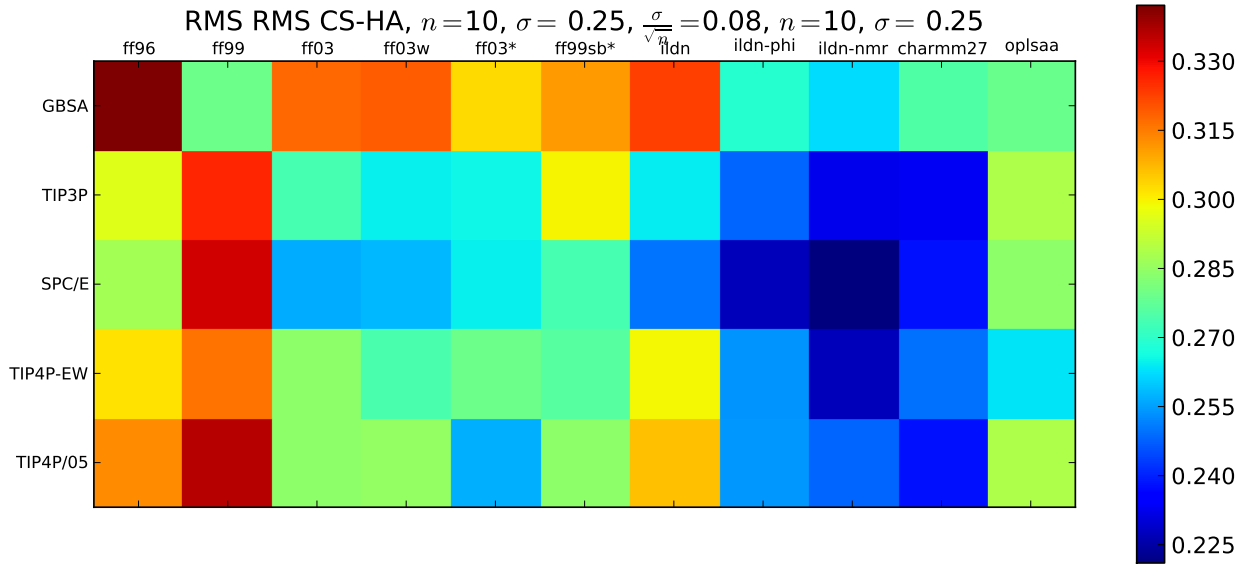


Figure S21: RMS errors of CS-Hα experiments.

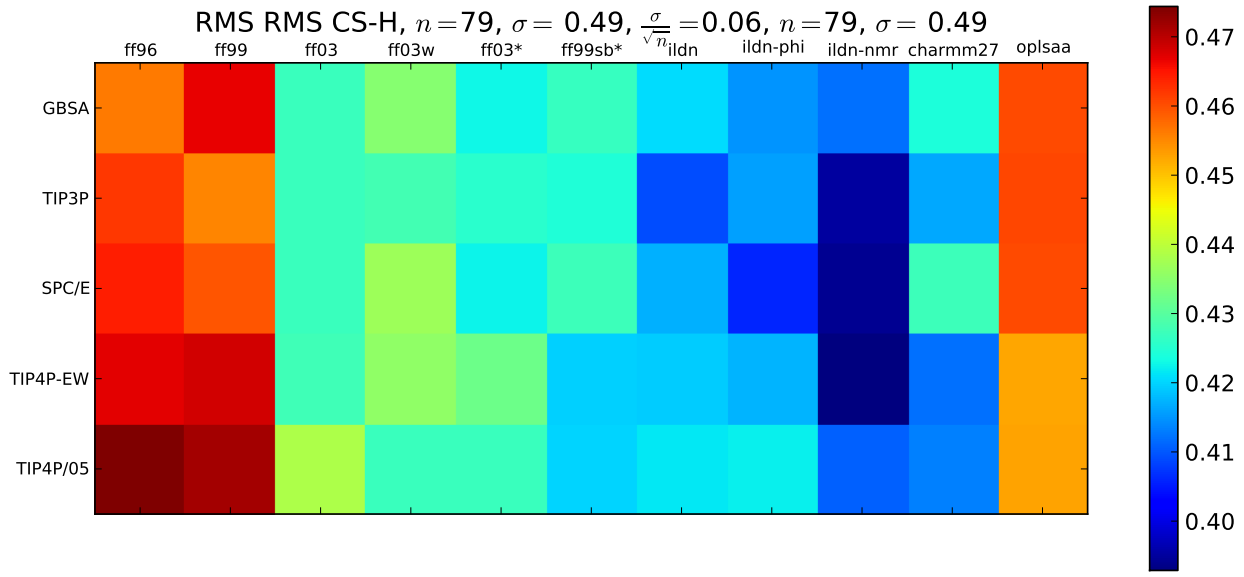


Figure S22: RMS errors of CS-H experiments.

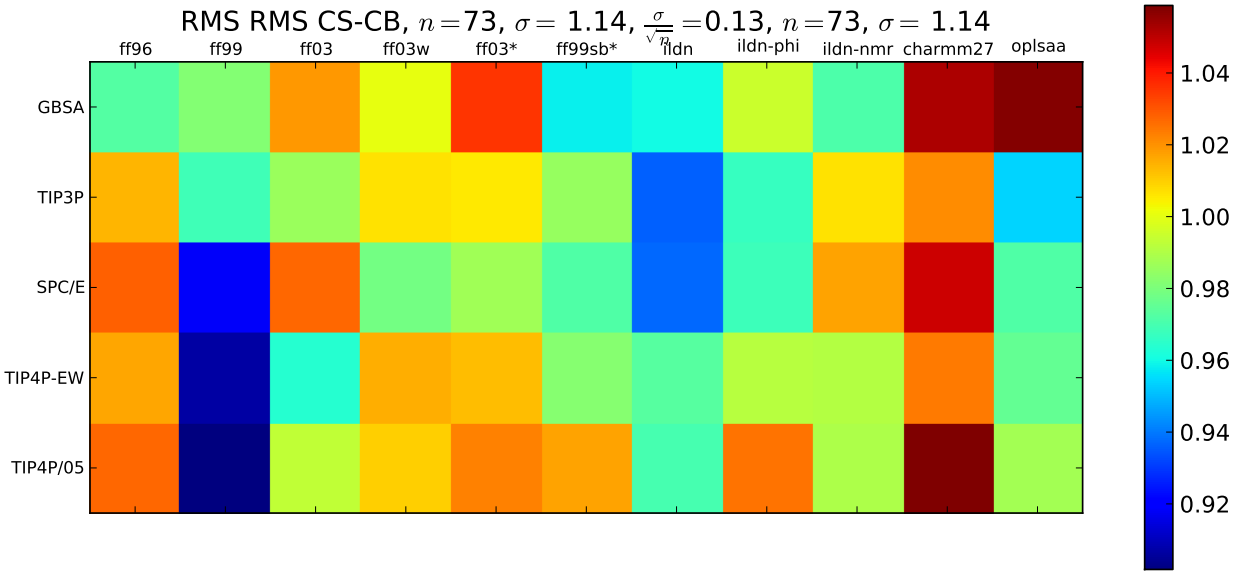


Figure S23: RMS errors of CS-Cβ experiments.

5 $^3J(H^\alpha C')$ comparison.

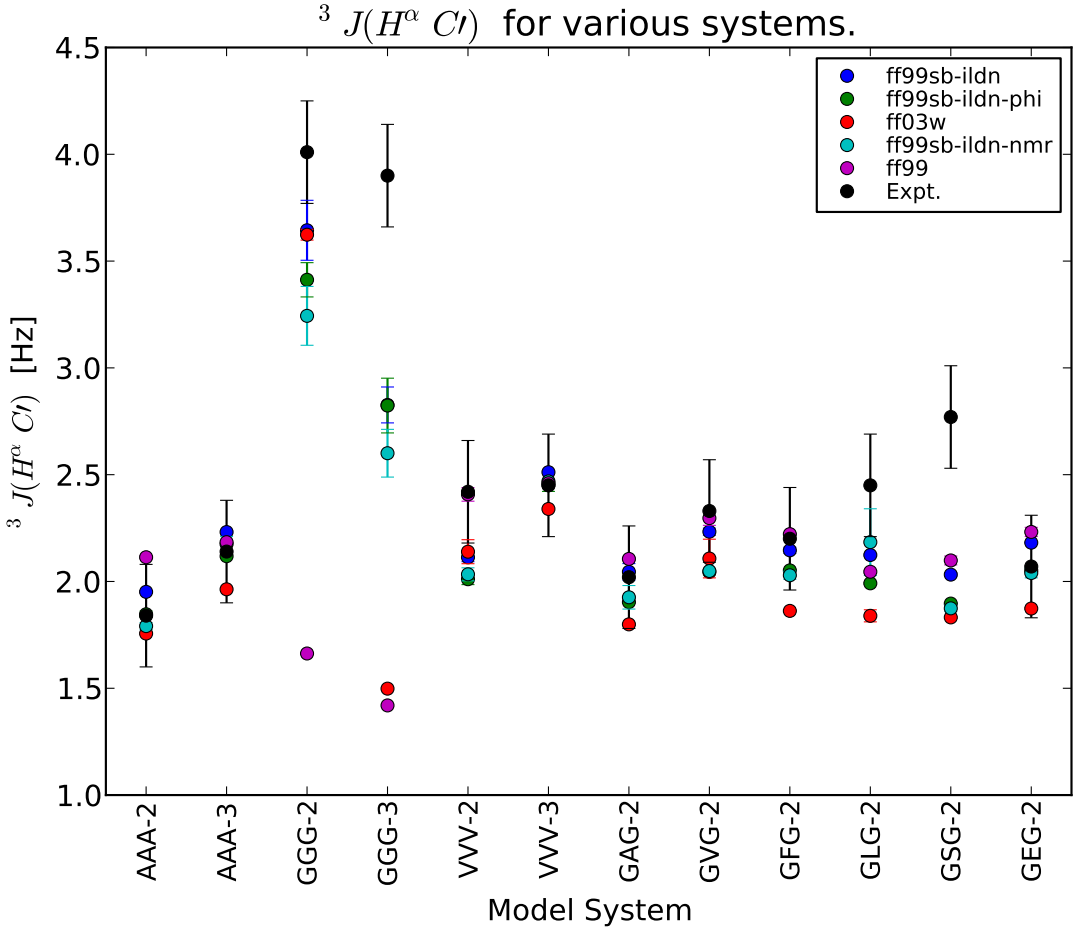


Figure S24: Values are shown for TIP4P-EW with five force fields.

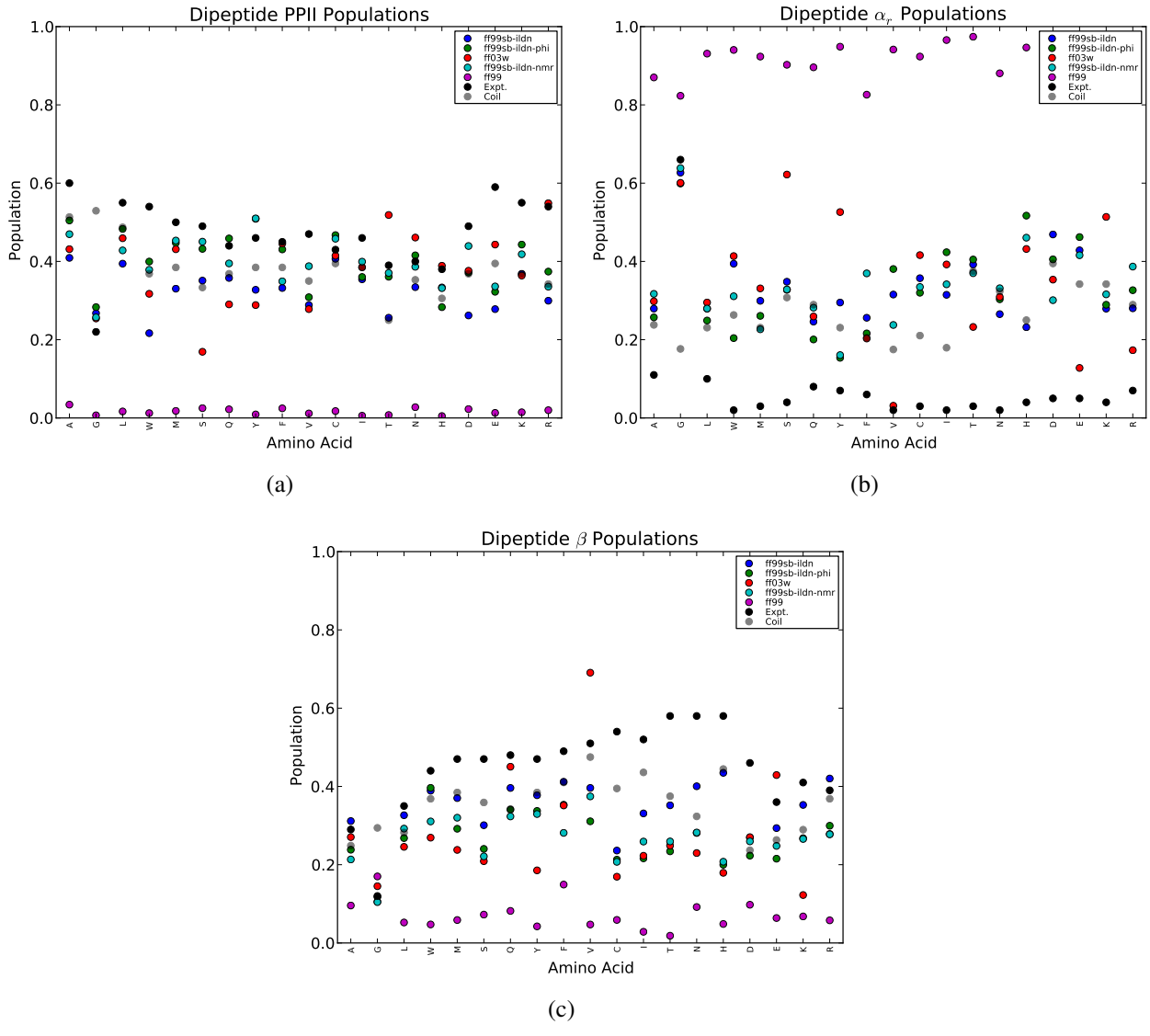


Figure S25: Populations of all 19 dipeptides for five force fields (with TIP4P-EW). This analysis suggests that amber99 is significantly improved by amber99sb-ildn, which is modestly improved by the amber99sb-ildn-phi and amber99sb-ildn-nmr force fields. The PII populations of amber99sb-ildn-phi and amber99sb-ildn-nmr appear approximately correct, while their beta and alpha populations may benefit from further refinement.

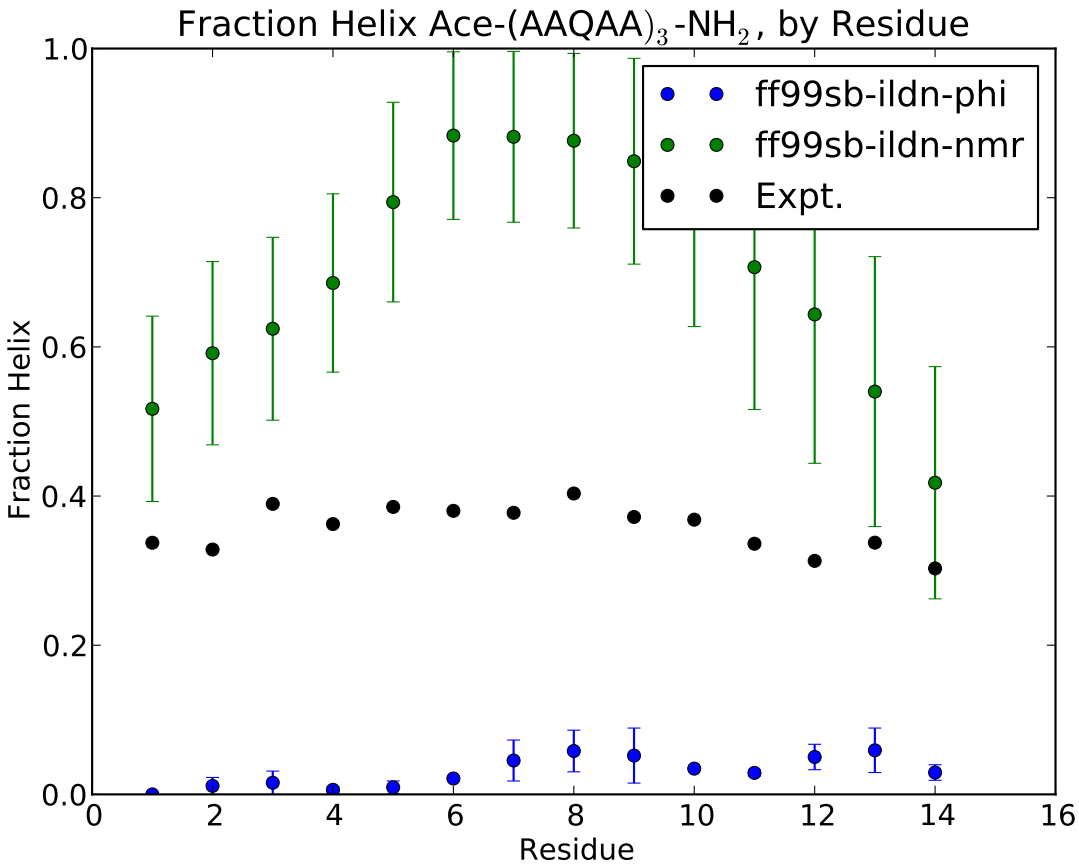


Figure S26: The helix forming peptide $Ace-(AAQAA)_3-NH_2$ has previously been used to illustrate limitations[5] in secondary structure propensities of force fields. Thus, we considered a long simulation (>1000 ns; 300K, started from helical conformation) of this peptide using each of the two best-performing force fields (ff99sb-ildn-nmr and ff99sb-ildn-phi, with TIP4P-EW) in this work. We find that ff99sb-ildn-nmr somewhat overestimates helical content, while ff99sb-ildn-phi inherits ff99sb's under-helical tendencies. Secondary structure were estimated using Stride;[6] errors estimated via splitting each simulation into two blocks of equal length.

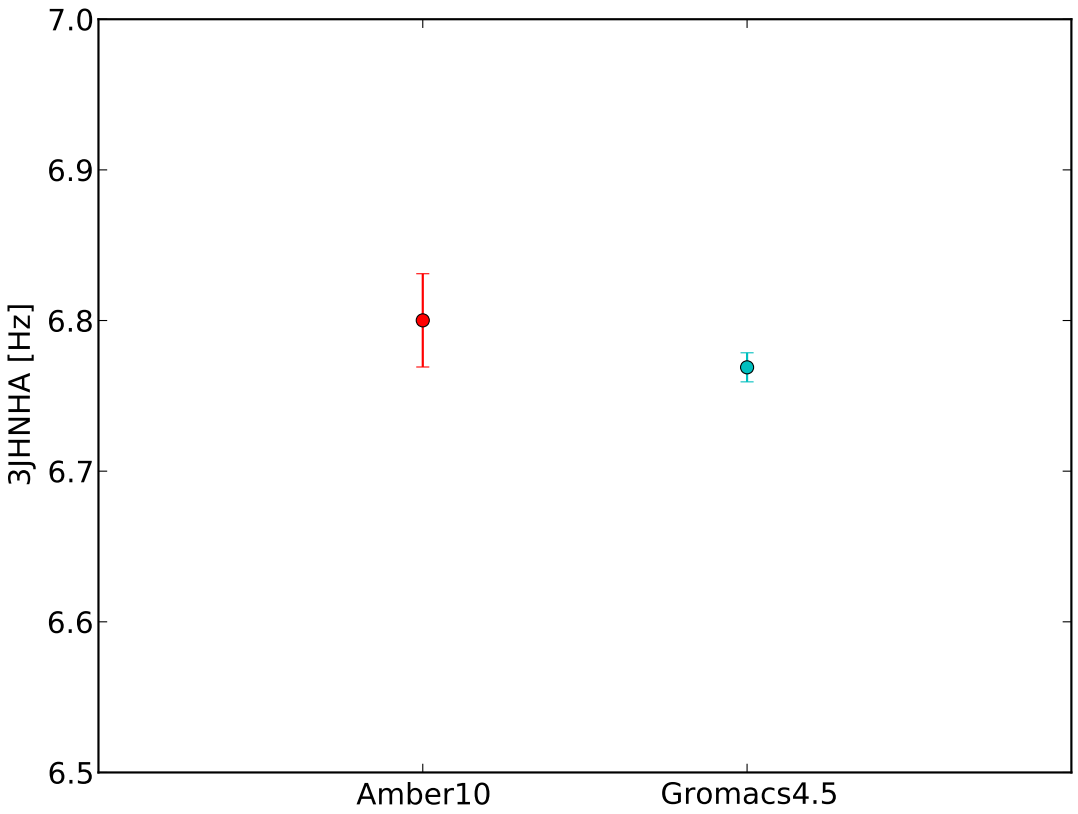


Figure S27: To verify the consistency of simulations performed using Gromacs, we used Amber10[7] to perform a 10 ns simulation of alanine dipeptide in TIP3P with the amber99sb-ildn force field. We then compared the J coupling estimate using this simulation. We find that the result is in good agreement with the Gromacs simulations.

Table S1: Number of measurements, sorted by experiment type and model system. Experiments are either J couplings or chemical shifts (CS). The 19 dipeptides (rows 1-19) are capped dipeptides of the form NME-X-ACE, consistent with the experiments in ref. [8,9] The remaining constructs have charged termini.

System	$^3JH_NH_\alpha$	$^3JH_\alpha C'$	$^3JH_\alpha N$	$^3JH_N C'$	CS-N	CS-C	CS-Cα	CS-Hα	CS-H	CS-Cβ
A	1	0	0	0	0	0	0	0	0	0
G	1	0	0	0	0	0	0	0	0	0
L	1	0	0	0	0	0	0	0	0	0
W	1	0	0	0	0	0	0	0	0	0
M	1	0	0	0	0	0	0	0	0	0
S	1	0	0	0	0	0	0	0	0	0
Q	1	0	0	0	0	0	0	0	0	0
Y	1	0	0	0	0	0	0	0	0	0
F	1	0	0	0	0	0	0	0	0	0
V	1	0	0	0	0	0	0	0	0	0
C	1	0	0	0	0	0	0	0	0	0
I	1	0	0	0	0	0	0	0	0	0
T	1	0	0	0	0	0	0	0	0	0
N	1	0	0	0	0	0	0	0	0	0
H	1	0	0	0	0	0	0	0	0	0
D	1	0	0	0	0	0	0	0	0	0
E	1	0	0	0	0	0	0	0	0	0
K	1	0	0	0	0	0	0	0	0	0
R	1	0	0	0	0	0	0	0	0	0
AAA	2	2	0	2	2	2	3	3	2	3
GGG	2	2	0	2	2	2	3	0	2	0
VVV	2	2	0	2	2	2	3	3	2	3
AAAA	2	0	0	0	3	3	4	4	3	4
GAG	1	1	0	1	0	0	0	0	0	0
GVG	1	1	0	1	0	0	0	0	0	0
GFG	1	1	0	1	0	0	0	0	0	0
GLG	1	1	0	1	0	0	0	0	0	0
GSG	1	1	0	1	0	0	0	0	0	0
GEG	1	1	0	1	0	0	0	0	0	0
GKG	1	0	0	0	0	0	0	0	0	0
GMG	1	0	0	0	0	0	0	0	0	0
1UBQ	0	0	63	60	70	0	70	0	70	63

Table S2: The parameters for the Karplus relations used in this work are listed. Note that both 1997 and 1999 share the same $^3JH_\alpha N(\psi)$ parameterization, which was taken from ref.[10] The reported uncertainties are the RMS errors calculated during the model parameterizations. All Karplus relations assume the form $J(x) = A\cos(x+\theta)^2 + B\cos(x+\theta) + C$.

1997	θ	A	B	C	σ [Hz]
$^3JH_NH_\alpha$	-60	7.09	-1.42	1.55	0.39
$^3JH_NC_\beta$	60	3.06	-0.74	0.13	0.21
$^3JH_\alpha C\prime$	120	3.72	-2.18	1.28	0.24
$^3JH_NC\prime$	180	4.29	-1.01	0.00	0.32
$^3JH_\alpha N$	60	-0.88	-0.61	-0.27	0.16
1999	θ	A	B	C	σ [Hz]
$^3JH_NH_\alpha$	-60	7.90	-1.05	0.65	0.78
$^3JH_NC_\beta$	60	2.90	-0.56	0.18	0.25
$^3JH_\alpha C\prime$	120	3.76	-1.63	0.89	0.48
$^3JH_NC\prime$	180	4.41	-1.36	0.24	0.43
$^3JH_\alpha N$	60	-0.88	-0.61	-0.27	0.16

Table S3: Error estimates [ppm] for chemical shift predictions from SPARTA+ and ShiftX. The reported uncertainties are the RMS errors calculated during the model parameterizations.

Package / Atom	N	C	Cα	Hα	H	Cβ
SPARTA+	2.45	1.09	0.94	0.25	0.49	1.14
ShiftX	2.43	1.16	0.98	0.23	0.49	1.10

Table S4: The χ^2 contribution of each experiment is tabulated. Because the number of measurements differs for each experiment type, the various experiments are weighted differently in the overall χ^2. Thus, we also consider the reduced χ^2 ($\frac{\chi^2}{n}$), which indicates whether (e.g. if $\frac{\chi^2}{n} \leq 1$) the simulation error is within the uncertainty for each experiment class. Given values of χ^2 and $\frac{1}{n}\chi^2$ are for ff99sb-ildn-phi with TIP4P-EW.

Experiment	FF+WM (top performer of class)	RMS (top)	RMS (ildn-phi)	RMS (ildn-nmr)	σ	χ^2	$\frac{\chi^2}{n}$
${}^3J(H^N H^\alpha)$	ff99sb-ildn-nmr + TIP4P-EW	0.55 [Hz]	0.58	0.55	0.39	75.1	2.15
${}^3J(H^\alpha C')$	ff99sb* + TIP4P-EW	0.37 [Hz]	0.47	0.54	0.24	48.2	4.01
${}^3J(H^\alpha N)$	ff99sb-ildn + GBSA	0.18 [Hz]	0.18	0.19	0.16	82.6	1.29
${}^3J(H^N C')$	ff99sb-ildn-nmr + SPC/E	0.41 [Hz]	0.42	0.41	0.32	125.1	1.73
$CS-N$	ff99sb-ildn-nmr + SPC/E	2.20 [ppm]	2.32	2.20	2.45	70.4	0.89
$CS-C$	ff99sb-ildn-nmr + GBSA (tied)	2.46 [ppm]	2.46	2.46	1.09	45.9	5.10
$CS-C\alpha$	ff99sb-ildn-nmr + TIP4P-EW	0.92 [ppm]	0.98	0.95	0.94	86.4	1.07
$CS-H\alpha$	ff99sb-ildn-nmr + SPC/E	0.22 [ppm]	0.23	0.22	0.25	10.3	1.03
$CS-H$	ff99sb-ildn-nmr + TIP4P-EW	0.39 [ppm]	0.41	0.39	0.49	57.3	0.73
$CS-CB$	ff99 + TIP4P/05	0.90 [ppm]	0.97	1.01	1.14	55.2	0.76

This material is available free of charge via the Internet at `http://pubs.acs.org`.

References

(1) Hu, J.-S.; Bax, A. *J. Am. Chem. Soc.* **1997**, *119*, 6360–6368.

(2) Schmidt, J.; Blümel, M.; Löhr, F.; Rüterjans, H. *J. Biomol. NMR* **1999**, *14*, 1–12.

(3) Shen, Y.; Bax, A. *J. Biomol. NMR* **2010**, *48*, 13–22.

(4) Neal, S.; Nip, A. M.; Zhang, H.; Wishart, D. S. *J. Biomol. NMR* **2003**, *26*, 215–240.

(5) Best, R.; Hummer, G. *J. Phys. Chem. B* **2009**, *113*, 9004–9015.

(6) Frishman, D.; Argos, P. *Proteins* **1995**, *23*, 566–579.

(7) Case, D.; Darden, T.; Cheatham Iii, T.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Crowley, M.; Walker, R.; Zhang, W. et al.Case, DA and Darden, TA and Cheatham Iii, TE and Simmerling, CL and Wang, J. and Duke, RE and Luo, R. and Crowley, M. and Walker, R.C. and Zhang, W. and others *University of California, San Francisco* **2008**, *32*, year.

(8) Avbelj, F.; Grdadolnik, S.; Grdadolnik, J.; Baldwin, R. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 1272.

(9) Grdadolnik, J.; Mohacek-Grosev, V.; Baldwin, R.; Avbelj, F. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 1794.

(10) Wang, A. C.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 1810–1813.