Journal of the Korean Magnetic Resonance Society **2022**, *26*, 1-9
DOI 10.6564/JKMRS.2022.26.1.001

Refinement of protein NMR structures using atomistic force field and implicit solvent model: Comparison of the accuracies of NMR structures with Rosetta refinement

Jun-Goo Jee *

Research Institute of Pharmaceutical Sciences, College of Pharmacy, Kyungpook National University
80 Daehak-ro, Buk-gu, Daegu 41566, Republic of Korea

Received Jan 21, 2022; Revised Jan 24, 2022; Accepted Jan 25, 2022

Abstract There are two distinct approaches to improving the quality of protein NMR structures during refinement: all-atom force fields and accumulated knowledge-assisted methods that include Rosetta. Mao *et al*. reported that, for 40 proteins, Rosetta increased the accuracies of their NMR-determined structures with respect to the X-ray crystal structures (Mao *et al*., *J. Am. Chem. Soc*. **136**, 1893 (2014)). In this study, we calculated 32 structures of those studied by Mao *et al*. using all-atom force field and implicit solvent model, and we compared the results with those obtained from Rosetta. For a single protein, using only the experimental NOE-derived distances and backbone torsion angle restraints, 20 of the lowest energy structures were extracted as an ensemble from 100 generated structures. Restrained simulated annealing by molecular dynamics simulation searched conformational spaces with a total time step of 1-ns. The use of GPU-accelerated AMBER code allowed the calculations to be completed in hours using a single GPU computer—even for proteins larger than 20 kDa. Remarkably, statistical analyses indicated that the structures determined in this way showed overall higher accuracies to their X-ray structures compared to those refined by Rosetta (p-value < 0.01). Our data demonstrate the capability of sophisticated atomistic force fields in refining NMR structures, particularly when they are coupled with the latest GPU-based calculations. The straightforwardness of the protocol allows its use to be extended to all NMR structures.

Keywords NMR structure refinement, all-atom force field, generalized-Born implicit solvent

Introduction

The structure determination of biomolecules using NMR data couples NOE assignments and structure calculations in an iterative way.[1] The experimental restraints guide the conformational search, and the calculated structures provide additional experimental restraints. The latest algorithms can automate the procedure for NMR structure calculations without any manual intervention and interpretation. However, chemical shifts must be known and assigned to most of the atoms. Furthermore, good quality NOE data must be available from which sufficient distance restraints can be obtained.[2-3] There is now a standard protocol for structure determination, but there are few available experimental restraints compared with X-ray crystallography. Therefore, the determination of structures with the same level of accuracy and precision as X-ray crystallography is still difficult.

* Address correspondence to: **Jun-Goo Jee,** Research Institute of Pharmaceutical Sciences, College of Pharmacy, Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu 41566, Republic of Korea. Tel: +82-53-950-8568; E-mail: jjee@knu.ac.kr

One computational approach to overcome these difficulties and improve the accuracy and precision of NMR structure refinement is to use sophisticated force fields. Traditional software programs for NMR structure calculations use simplified force fields compared to those used in atomistic molecular dynamics (MD) simulations.[4-6] This simplification principally occurs by neglecting the Lennard-Jones potentials and solvation energies terms. Whereas these simplifications enable the efficient survey of conformational space by high-temperature annealing, the geometries of the regions that lack structural restraints often diverge and are occasionally inaccurate. All-atom force field (AAFF) MD calculations are helpful for characterizing the areas where experimental restraints are insufficient. Most atomistic MD-driven calculations for NMR structure refinement approximate the solvation effects using the generalized-Born implicit solvent (GBIS) model. Using GBIS yields structures of the quality expected from explicit solvation calculations without the large computational overhead. Simulated annealing by MD simulation can accelerate the search of conformational spaces combined with AAFF-GBIS. Several researchers have reported the successful application of AAFF-GBIS in protein, protein-protein complex, and membrane protein NMR structure refinement.[7-13]

Another computational approach is to extend the accumulated knowledge of the experimental structures. Conformational space annealing can belong to this type.[14] But, the most well-known algorithm is Rosetta.[15] Once developed for de novo protein structure prediction, Rosetta has evolved into a tool for homology modeling, protein design, and the calculation of NMR structures. Rosetta extracts fragments from a database of 3D structures and then assembles the fragments. This method improves the efficiency of searching conformational space. Indeed, several reports have demonstrated the success of this type of structure calculation from NMR data even in large proteins greater than 20 kDa.[16-19] Besides the improvements in determining the protein's global fold, these recent results demonstrated that Rosetta is useful for refining local structures. Mao *et al*. reported that Rosetta-refined structures are more suitable for molecular replacement in X-ray crystal structure phasing, which is indicative that these refined structures are more accurate with respect to the X-ray crystal structures of the corresponding protein.[20]

Studies indicate that the improvements made by atomistic MD simulation and Rosetta are complimentary and can be used in a synergetic way.[21-26] Despite these two computational approaches being popular, there are few direct comparisons of the resulting NMR structures. To develop synergistic methods in NMR structure calculations, it is a prerequisite that we understand the strengths and weaknesses of both approaches. In this study, we focus on the comparison of NMR structures refined by AAFF-GBIS and Rosetta. We selected the 32 protein NMR structures for which both the information from Rosetta-refined structures and experimental distance and torsion angle restraints are available. Atomistic MD simulations need long calculation times; however, here we have used the GPU-accelerated AMBER code,[27-29] and this has increased the refinement speed and reduced the computational time required to the level of conventional NMR structure calculations.

Experimental Methods

NMR structure calculations—Table 1 contains a list of the proteins modeled in this study and their various statistical properties. The distance and backbone torsion angle restraints extracted from the PDB-deposited restraints were used to calculate the structures. However, the hydrogen bond and side-chain torsion angle restraints were excluded. Also, the restraints from residual dipolar couplings (RDCs) were omitted. Structure calculations consisted of initial calculations by CYANA[6] followed by simulated annealing refinement using atomistic MD simulations with AMBER.[30] Firstly, 300 structures were calculated for each protein using the experimental distance and torsion angle restraints with CYANA. Then, using an iterative process, any restraints that showed significant violation from predetermined values (> 0.5 Å for distances and > 5° for torsion angles) were removed. The top 100 CYANA structures that had the lowest target functions were chosen for further refinement. The ff14SB AAFF and the GBIS solvation model as implemented in the PMEMD code[27-28] were used from the AMBER package (version 14) running on

CUDA-enabled GPUs.[30] The GBIS was set to '*igb* = 8'. AMBER refinement comprised three stages: first, a 1500-step minimization; then, a 1-ns restrained simulated annealing; and, final, a second 1500-step minimization. The 1-ns restrained simulated annealing was divided into three stages. Initially, for the first 250 s, the temperature was increased to 1000 K. For the second 250 s, the temperature was held at 1000 K. Finally, a stepwise cooling to 0 K was carried out for 500 s. The force constants for distance and torsion angle restraints were 20 kcal/(mol·Å^2) and 200 kcal/(mol·rad^2), respectively. The integration time step for restrained simulated annealing was 2 fs with SHAKE restraints. Of 100 structures, 20 structures that had the lowest AMBER energies were selected as an ensemble for analysis. All the computations were carried out with single-precision floating-point (SPFP) corrections[31] on personal computers equipped with GTX-780 graphic cards.

Quantitative analyses of calculated NMR structures—The calculated structures were compared on the basis of their backbone Cα atoms root mean square deviations (RMSDs) to the mean structure in the ensemble. The ranges for the RMSD comparison were determined with CYRANGE.[32] The similarity to the corresponding X-ray structure was measured by the template modeling score (TM-score) and global distance test total score (GDT-TS) values. The TM-score is a measure of the global fold similarity and is less sensitive than other structure-metrics to local structural variations.[33-34] GDT-TS is a measure of the 3D similarity of two structures[35-37] and is the arithmetic mean of four values: the percentage of structurally equivalent pairs of Cα atoms within specified distance cutoffs of d = 1, 2, 4, and 8 Å. The TM-score program by Zhang and Skolnick was used to calculate both the TM-score and GDT-TS values. In the metrics of TM-score and GDT-TS, the averaged values from the 20 structures in each protein were represented. Structural quality scores were analyzed using the protein structure validation software suite (PSVS),[38] which generated Verify3D, Procheck (phi-psi), Procheck (all), and MolProbity Clashscore metrics.

Results and Discussion

Mao *et al*. reported Rosetta-refined structures for 40 proteins.[20] Because eight of the proteins are either complexes or dimers, we omitted them for convenience with respect to format conversion. However, the exclusion of these structures does not imply difficulty in calculating the structures of complex proteins using AAFF-GBIS. We refined the 32 protein structures (Table 1) by generating 300 structures using CYANA and then extracting 100 low-target scoring structures for each protein. From these 100 structures, an ensemble of the top 20 structures with the lowest AMBER energies was identified. The high speed of the CYANA calculations enabled the removal of any structures that violated restraints via iterative runs. This resulted in no significant restraint violations (> 0.5 Å for distance and > 5° for torsion angle restraints) in the ensembles generated by the AAFF-GBIS method. These structures differed from the results of Rosetta calculations, which started from the deposited PDB structures and contained several input restraint violations.[20] In the case of a 20 kDa protein, an AAFF-GBIS calculation run on a single CPU (Intel E5-2650, 2.60 GHz) with a total time step of 1-ns takes approximately 8 h per structure. In stark contrast, the CUDA-encoded SPFP version of PMEMD finished the same calculation within 6 min on a GTX-780 graphic card. This represents a speedup that is similar to reported benchmarks.[28] Although there are concerns about the imprecise handling of floating-point numbers in the GPU, the use of experimental restraints decreases potential errors. Indeed, even if the errors increase, the potential deviations would be smaller than those caused by the use of the simplified force fields used in conventional NMR structure calculations.[39]

Table 1. Statistics for the NMR structures refined by AAFF-GBIS

NMR	X-ray	ID	AA	eRMSD[¶] (Å)	TM-score	GDT-TS	Verify3D	Procheck (ϕ–ψ)	Procheck (all)	Mol-Probity
1PQX	2FFM	ZR18	91	0.81	0.86	0.87	-0.64	-0.55	-0.18	1.51
1XPV	1TTZ	XCR50	78	0.56	0.86	0.88	-0.80	-0.08	-0.06	1.48
1XPW	1TVG	HR1958	143	0.72	0.85	0.81	-1.28	-1.73	-1.42	1.5
2HFI*	2IM8	SGR145	123	0.31	0.92	0.92	-0.64	3.07	2.72	1.51
2JN0	3FIF	ER382A	50	0.49	0.78	0.84	-1.44	-1.18	-1.06	1.25
2JN8	2ES9	STR65	109	0.84	0.90	0.89	-1.28	1.77	1.54	1.41
2JPU	2Q00	SSR10	129	0.5	0.92	0.92	-1.61	2.71	2.48	1.50
2JQN	2O0Q	CCR55	116	0.71	0.92	0.9	-0.32	-0.24	0.0	1.46
2JVD*	3BHP	SGR209C	48	0.13	0.75	0.8	-2.25	2.56	3.08	1.53
2JZ2*	3C4S	SR213	66	0.28	0.93	0.97	-3.85	-1.3	-1.12	1.5
2JZT	2ES7	STR70	142	0.61	0.86	0.84	-0.32	0.67	0.0	1.44
2K07	3EVX	HR41	175	0.60	0.93	0.88	-1.12	0.16	0.18	1.51
2K2E	3CPK	BER31	158	0.62	0.90	0.89	-0.32	-0.28	-0.41	1.51
2K5P	3CWI	GMR137	78	0.61	0.75	0.8	-1.44	-1.53	-2.25	1.42
2K5V	3E0E	MRR110B	98	0.31	0.94	0.95	0.32	-1.49	-1.3	1.53
2KCU	3E0H	CTR107	166	1.27	0.83	0.74	0.0	-0.28	-0.47	1.46
2KCV	3MA5	SRR115C	99	0.22	0.94	0.94	-0.64	2.01	1.89	1.44
2KCZ	3GGN	DRR147D	155	1.31	0.60	0.56	-2.41	-0.04	-0.47	1.43
2KFP	3H9X	PSR293	125	0.91	0.88	0.82	0.32	-0.2	-0.83	1.45
2KHN	3FIA	HR3646E	121	0.77	0.84	0.82	-1.61	1.18	0.59	1.44
2KKZ	2RHK	OR8C	134	0.38	0.93	0.93	-1.12	-0.55	-0.77	1.50
2KL6	3IDU	PfR193A	108	0.13	0.87	0.86	-1.28	-1.3	-1.36	1.42
2KPP	3LD7	LKR112	114	0.27	0.95	0.97	-1.44	-1.46	-1.01	1.52
2KPU	3LYW	DHR29B	96	0.64	0.89	0.89	-0.8	-1.26	-1.01	1.50
2KPW	3JT0	HR5546A	122	0.77	0.88	0.85	-0.8	-1.46	-1.6	1.51
2KRK	3KW6	HR3102A	86	0.32	0.90	0.92	-1.12	1.49	1.66	1.51
2KRT	3K63	UUR17A	121	0.70	0.82	0.78	-2.41	-0.9	-1.42	1.49
2KW2	3LMO	RPR324	101	0.27	0.85	0.84	-0.32	1.18	1.36	1.53
2KW5*	3MER	SOR77	202	0.91	0.76	0.61	-0.48	-1.14	-1.36	1.14
2L05	3NY5	HR4694F	86	0.35	0.88	0.9	-1.44	-0.55	0.0	1.52
2L1P	3NZL	HR4435B	83	1.26	0.79	0.82	-0.48	2.16	2.01	1.50
2L33	3P1X	HR4527E	91	0.11	0.87	0.91	-2.09	0.0	0.18	1.40

[¶]Ensemble RMSD
*Omitted for comparison with Rosetta-refined structures due to the disagreements in the reported values

Structures refined by AAFF-GBIS are more precise than PDB-deposited structures. No single parameter can be used to judge whether a calculated NMR structure is accurate or not. The currently accepted idea in NMR community is to use several metrics simultaneously.[40] Accordingly, we extracted the metrics recommended for the validation of NMR structures. Ensemble RMSD reflects the precision of the coordinates. Previously, we have reported that if the structures are refined using an atomistic force field and the ensemble is selected from the viewpoints of AMBER energies then there is a correlation between precision and accuracy.[7, 41] Because the flanking regions in an ensemble can bias the calculation of the RMSD, we first obtained the core residues of the PDB structures using the CYRANGE program.[32] Within this range, we calculated the RMSD of the Cα atoms in the ensembles. Figure 1 showed the comparison of ensemble RMSD in the PDB-deposited and AAFF-GBIS-refined structures. Data points positioned under the diagonal line give an estimate of the increased precision of structures generated by AAFF-GBIS. The figure clearly reveals that AAFF-GBIS-generated NMR structures are more precise with respect to the X-ray structures, as reported in our previous studies.[7, 10-11] Of the 32 AAFF-GBIS structures, five — 1PQX, 1XPV, 2KFP, 2KW5, and 2L1P — showed an increase in RMSD values. However, the deposited NMR structures of 1PQX, 1XPV, 2KFP and 2KW5 were refined with additional restraints except for 2L1P, where the decrease in RMSD (0.02 Å) is not statistically significant. The restraints for 1PQX, 2KFP, and 2KW5 included 54, 80, and 92 hydrogen bonds, respectively. The 78 torsion angle restraints confining side-chain geometries were added to those of 1XPV. The restraints of 2KW5 consisted of 201 restraints derived from residual dipolar couplings (RDC). In most cases, the existence of a hydrogen bond is judged based on the secondary structure. However, because this can miss abnormal conformations such as bulges in the protein structure, it would be better to determine the precise structure without adding restraints. Besides these four cases, of the 32 cases, 20 structures also used RDCs as additional restraints. This indicates that AAFF-GBIS improved the precision of the structures even when not restrained by RDC.

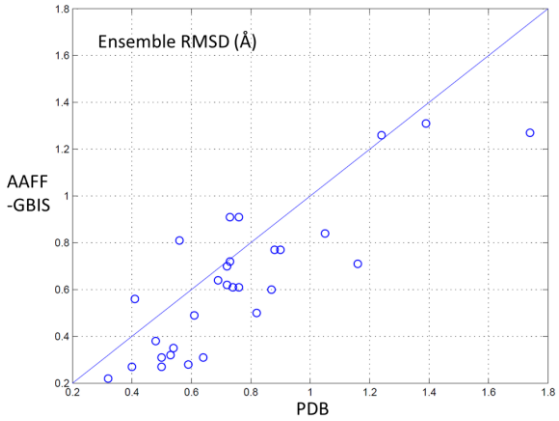

Figure 1. Comparison of ensemble RMSDs between PDB-deposited and AAFF-GBIS-refined NMR structures. CYRANGE was used to extract the core range to calculate the RMSD. Each data point indicates the mean value of the ensemble.

Structures refined by AAFF-GBIS have greater accuracy than PDB-deposited structures. We then compared TM-scores and GDT-TS values between PDB-deposited and AAFF-GBIS-refined NMR structures (Figure 2). Data points that lie above the diagonal line indicate increased accuracy with respect to the X-ray structures. The data clearly show the superior quality of the structures produced by AAFF-GBIS. The improvements by AAFF-GBIS are statistically significant by paired *t*-test (*p*-value < 10^{-5}). In particular, the GDT-TS increased by more than 0.1 in the cases of 2JQN and 2L1P. In the case of 2JQN, there was a substantial decrease in the backbone RMSD: a reduction from 1.16 to 0.71 Å. Again, this highlights the advantages of using AAFF-GBIS for refining NMR structures. We found the strongest correlation between the TM-scores and GDT-TS values of the AAFF-GBIS structures, which had a Pearson correlation coefficient of 0.88 (R^2= 0.77). The correlation coefficient between the RMSD and GDT-TS in each structure ensemble was −0.66 (R^2 = 0.44). This implies that the ensemble RMSD reflects the accuracy of structures to some degree, as we have shown previously.[7]

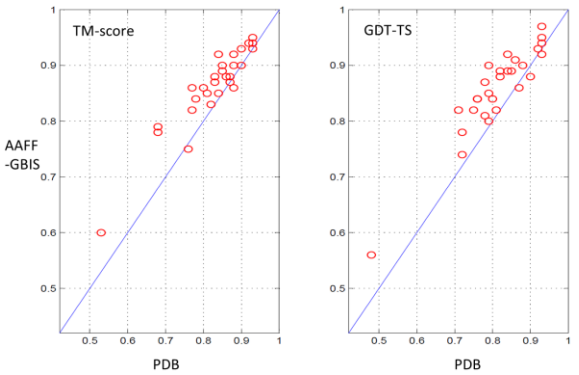

Figure 2. Comparison of the TM-scores and GDT-TS values between the PDB-deposited and AAFF-GBIS-refined NMR structures. The TM-score program calculated the values of TM-score and GDT-TS. Each data point indicates the mean value of the ensemble. The diagonal line indicates x = y.

Structures refined by AAFF-GBIS show higher accuracy than Rosetta-refined structures. Next, the results by AAFF-GBIS and Rosetta were compared. The Rosetta-refined structures are not publically available, which made it impossible to make a straightforward comparison of the AAFF-GBIS and Rosetta structures. Instead, we inspected the parameters that Mao *et al*. reported[20] and those that we calculated from our AAFF-GBIS structures. We found some disagreement between the values that likely arose from differences in parameter extraction. For example, the values of the TM-score and GDT-TS of 2HFI (NMR) with respect to 1TVG (X-ray) were 0.86 and 0.82, respectively, whereas the values reported by Mao *et al.* were 0.79 and 0.77, respectively. Therefore, in order to compare fairly the AAFF-GBIS and Rosetta-determined structures, we extracted the parameters in the cases where the values from PDB structures and from the paper were comparable. For instance, because the values for TM-score and GDT-TS in 1XPV were 0.88 and 0.90, respectively, both in the reported result and our calculation, we accepted all the parameters of 1XPV from Rosetta. Using our criteria, there are 28 cases for further comparison, omitting 4 — 2HFI, 2JVD, 2JZ2, and 2KW5 — due to the large deviations. Figure 3 shows the comparison of TM-score and GDT-TS between structures calculated by AAFF-GBIS and those by Rosetta. Remarkably, the structures generated by AAFF-GBIS are more accurate in 20 out of 28 cases with respect to the corresponding X-ray structures than Rosetta. In some cases, the Rosetta results were better; however, the overall results are sufficient to demonstrate the outperformance of AAFF-GBIS. The paired *t*-test reveals that the probability that AAFF-GBIS outperforms Rosetta by chance is 0.004 and 0.003 in TM-score and GDT-TS, respectively, proving that the improvements are statistically significant. In addition, it is noteworthy that the reported values of ensemble RMSD in the Rosetta-refined structures were comparable to those of the PDB-deposited structures.[20] This indicates that AAFF-GBIS structures were more precise than those calculated by Rosetta.

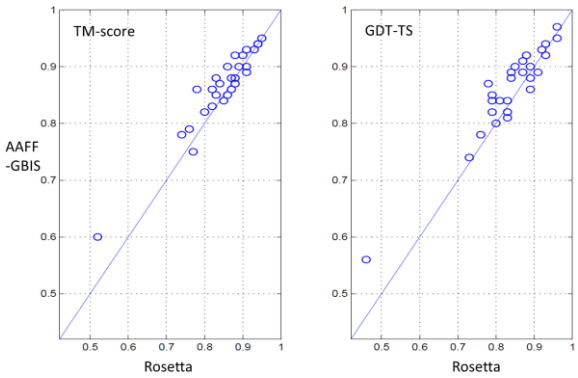

Figure 3. Comparison of TM-scores and GDT-TS values between Rosetta- and AAFF-GBIS-refined NMR structures. The same annotations as Figure 2 were used. Paired *t*-test indicates that the probability that AAFF-GBIS outperforms Rosetta by chance is lower than 0.01 both in TM-score and GDT-TS (p-value < 0.01).

Structures refined by AAFF-GBIS show better geometries than Rosetta-refined structures. How good are the geometries of the structures refined by AAFF-GBIS? We used the PSVS server (http://psvs-1_5-dev.nesg.org) [38] to extract four metrics—Verify3D, Procheck (phi-psi), Procheck (all), and MolProbity Clashscore—from the structures produced by AAFF-GBIS (Table 1). We compared these values with the reported values from both the PDB-deposited and Rosetta-refined structures. The metrics were normalized and are presented as Z-scores. Values close to zero indicate a good geometry. Figure 4 shows the values of the PDB-deposited, Rosetta-refined, and AAFF-GBIS-refined structures as boxplots. For all metrics except the MolProbity Clashscore, the Z-scores of the AAFF-GBIS structures have better statistical distributions. This demonstrates that the structures refined by AAFF-GBIS also show

improved geometries over Rosetta and PDB deposited structures. The largest improvements were found for 1PQX and 2KCZ, where the GDT-TS values increased from 0.78 to 0.87 and from 0.46 to 0.56, respectively. The reason that the MolProbity Clashscores of AAFF-GBIS structures were worse than those in Rosetta structures is unclear at this moment. To know whether one can improve the Clashscore by the force field or the solvent model of AAFF-GBIS will deserve future study. Nevertheless, it is noteworthy that the Clashscores by AAFF-GBIS-refined structures are still better than those found in PDB structures.

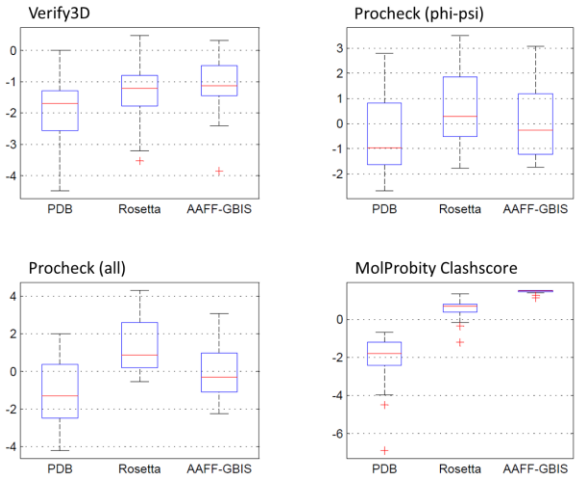

Figure 4. Boxplot-based comparison of geometric parameters extracted by PSVS version 1.5[38] in PDB-deposited, Rosetta- and AAFF-GBIS-refined NMR structures. The normalized Z-scores close to zero indicate a good geometry.

Conclusion

In this study, systematic calculations have convincingly shown that refinement of protein NMR structures with AAFF-GBIS improves the protein structures in terms of precision, accuracy, and geometry in comparison with other methods. To the best of our knowledge, this is the first example of a comparison between the performance of AAFF-GBIS and Rosetta for NMR structure determination. The PMEMD GPU-accelerated CUDA code allowed AAFF-GBIS refinements to run in several minutes even for time steps of up to 1-ns. This enabled many calculations, consisting of 3,200 runs (100 × 32), to be carried out. Recent studies have demonstrated the synergic use of AAFF-GBIS and Rosetta methods in modeling, and considering the popularity and power of these methods, their combined use will open new opportunities in many areas including protein engineering and design. Also, the synergistic use of AAFF-GBIS, Rosetta, and AlphaFold[42-43] for the refinement of NMR structures and the use of NMR restraints for modeling can lead to significant improvements in this field. Our data will be a valuable addition to future studies in this direction.

References

1. K. Wüthrich, *NMR of Proteins and Nucleic Acids.* Wiley, New York (1986)
2. T. J. Ragan, R. H. Fogh, R. Tejero, W. Vranken, G. T. Montelione, A. Rosato and G. W. Vuister, *J. Biomol. NMR*, **62**, 413 (2015)
3. A. Rosato, J. M. Aramini, C. Arrowsmith, A. Bagaria, D. Baker, A. Cavalli, J. F. Doreleijers, A. Eletsky, A. Giachetti, P. Guerry, A. Gutmanas, P. Güntert, Y. He, T. Herrmann, Y. J. Huang, V. Jaravine, H. R. Jonker, M. A. Kennedy, O. F. Lange, G. Liu, T. E. Malliavin, R. Mani, B. Mao, G. T. Montelione, M. Nilges, P. Rossi, G. van der Schot, H. Schwalbe, T. A. Szyperski, M. Vendruscolo, R. Vernon, W. F. Vranken, S. Vries, G. W. Vuister, B. Wu, Y. Yang and A. M. Bonvin, *Structure*, **20**, 227 (2012)
4. C. D. Schwieters, J. J. Kuszewski, N. Tjandra and G. M. Clore, J. Magn. Reson., **160**, 65 (2003)
5. A. T. Brunger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson and G. L. Warren, *Acta Crystallogr. D Biol. Crystallogr.*, **54**, 905 (1998)

6. P. Güntert, C. Mumenthaler and K. Wüthrich, *J. Mol. Biol.*, **273**, 283 (1997)
7. J.-G. Jee, *Bull. Korean Chem. Soc.*, **35**, 1944 (2014)
8. N. Sekiyama, J.-G. Jee, S. Isogai, K. Akagi, T. H. Huang, M. Ariyoshi, H. Tochio and M. Shirakawa, *J. Biomol. NMR*, **52**, 339 (2012)
9. J.-G. Jee, T. Mizuno, K. Kamada, H. Tochio, Y. Chiba, K. Yanagi, G. Yasuda, H. Hiroaki, F. Hanaoka and M. Shirakawa, *J. Biol. Chem.*, **285**, 15931 (2010)
10. J.-G. Jee, *Bull. Korean Chem. Soc.*, **31**, 2717 (2010)
11. J.-G. Jee and H.-C. Ahn, *Bull Korean Chem. Soc.*, **30**, 1139 (2009)
12. A. Ohno, J.-G. Jee, K. Fujiwara, T. Tenno, N. Goda, H. Tochio, H. Kobayashi, H. Hiroaki and M. Shirakawa, *Structure*, **13**, 521 (2005)
13. K. Fujiwara, T. Tenno, K. Sugasawa, J.-G. Jee, I. Ohki, C. Kojima, H. Tochio, H. Hiroaki, F. Hanaoka and M. Shirakawa, *J. Biol. Chem.*, **279**, 4760 (2004)
14. K. Joo, I. Joung, J. Lee, J. Lee, W. Lee, B. Brooks, S. J. Lee and J. Lee, *Proteins*, 83, 2251 (2015)
15. R. Das and D. Baker, *Annu. Rev. Biochem.*, **77**, 363 (2008)
16. Y. Shen and A. Bax, *Nat. Methods*, **12**, 747 (2015)
17. O. F. Lange, P. Rossi, N. G. Sgourakis, Y. Song, H. W. Lee, J. M. Aramini, A. Ertekin, R. Xiao, T. B. Acton, G. T. Montelione and D. Baker, *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 10873 (2012)
18. S. Raman, O. F. Lange, P. Rossi, M. Tyka, X. Wang, J. Aramini, G. Liu, T. A. Ramelot, A. Eletsky, T. Szyperski, M. A. Kennedy, J. Prestegard, G. T. Montelione, D. Baker, *Science*, **327**, 1014 (2010)
19. Y. Shen, O. Lange, F. Delaglio, P. Rossi, J. M. Aramini, G. Liu, A. Eletsky, Y. Wu, K. K. Singarapu, A. Lemak, A. Ignatchenko, C. H. Arrowsmith, T. Szyperski, G. T. Montelione, D. Baker and A. Bax, *Proc. Natl. Acad. Sci. U. S. A.*, **105**, 4685 (2008)
20. B. Mao, R. Tejero, D. Baker, G. T. and Montelione, J. Am. Chem. Soc., **136**, 1893 (2014)
21. H. Park, F. DiMaio, and D. Baker, *Structure*, **23**, 1123 (2015)
22. S. Lindert, J. A. and McCammon, *J. Chem. Theory Comput.*, **11**, 1337 (2015)
23. V. Mirjalili, K. Noyes and M. Feig, *Proteins*, **82** Suppl. 2, 196 (2014)
24. V. Mirjalili and M. Feig, *J. Chem. Theory Comput.*, **9**, 1294 (2013)
25. S. Lindert, J. Meiler and J. A. McCammon, *J. Chem. Theory Comput.*, **9**, 3843 (2013)
26. J. Chen and C. L. Brooks, 3rd, *Proteins*, **67**, 922 (2007)
27. R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, R. C. Walker, *J Chem Theory Comput*, **9**, 3878 (2013)
28. A. W. Gotz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, **8**, 1542 (2012)
29. J.-G. Jee, *J. Kor. Mag. Res. Soc.*, **18**, 69 (2014)
30. D. A. Case, T. A. Darden, T. E. Cheatham, I. Simmerling, C.L., J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Goetz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko and P. A. Kollman, *University of California, San Francisco* (2012)
31. S. Le Grand, A. W. Götz and R. C. Walker, *Computer Physics Communications*, **184**, 374 (2013)
32. D. K. Kirchner and P. Guntert, *BMC Bioinformatics*, **12**, 170 (2011)
33. J. Xu and Y. Zhang, *Bioinformatics*, **26**, 889 (2010)
34. Y. Zhang and J. Skolnick, *Proteins*, **57**, 702 (2004)
35. R. J. Read and G. Chavali, *Proteins*, **69** Suppl 8, 27 (2007)
36. A. Zemla, C. Venclovas, J. Moult and K. Fidelis, *Proteins*, Suppl 5, 13 (2001)
37. A. Zemla, C. Venclovas, J. Moult and K. Fidelis, *Proteins*, Suppl 3, 22 (1999)
38. A. Bhattacharya, R. Tejero and G. T. Montelione, *Proteins*, **66**, 778 (2007)
39. J.-G. Jee, *J. Kor. Mag. Res. Soc.*, **18**, 24 (2014)
40. G. T. Montelione, M. Nilges, A. Bax, P. Guntert, T. Herrmann, J. S. Richardson, C. D. Schwieters, W. F. Vranken, G. W. Vuister, D. S. Wishart, H. M. Berman, G. J. Kleywegt and J. L. Markley, *Structure*, **21**, 1563 (2013)

41. T. Ikeya, J.-G. Jee, Y. Shigemitsu, J. Hamatsu, M. Mishima, Y. Ito, M. Kainosho and P. Güntert, *J. Biomol. NMR*, **50**, 137 (2011)
42. K. Tunyasuvunakool, J. Adler, Z. Wu, T. Green, M. Zielinski, A. Zidek, A. Bridgland, A. Cowie, C. Meyer, A. Laydon, S. Velankar, G. J. Kleywegt, A. Bateman, R. Evans, A. Pritzel, M. Figurnov, O. Ronneberger, R. Bates, S. A. A. Kohl, A. Potapenko, A. J. Ballard, B. Romera-Paredes, S. Nikolov, R. Jain, E. Clancy, D. Reiman, S. Petersen, A. W. Senior, K. Kavukcuoglu, E. Birney, P. Kohli, J. Jumper and D. Hassabis, *Nature*, **596**, 590 (2021)
43. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli and D. Hassabis, *Nature*, **596**, 583 (2021)