Hybrid approaches to characterize structure and dynamics of biomolecular systems from single molecule experiments

Osamu Miyashita, Yuji Sugita and Florence Tama*
[1]*RIKEN, Advanced Institute for Computational Sciences*
7-1-26, Minatojima-minami-machi, Chuo-ku
Kobe, Hyogo, 650-0047, Japan

*florence.tama@gmail.com

1. INTRODUCTION

In recent years, the number of structures solved by cryo-EM methods has considerably increased and more structures are becoming available through the EMDB database [1]. In parallel, numerous computational tools have been developed to interpret cryo-EM data. In particular, multiple approaches have been developed to characterize conformational changes, which provide an opportunity to evaluate accuracy of pseudo atomic models obtained from computational tools. Using multiple flexible fitting methods can be a useful approach to evaluate the accuracy of models by comparing difference in their conformation. While flexible fittings methods can produce accurate atomic models, there are not always successful. In some particular cases, several of these methods consistently fail [2]. In light of such results, to obtain a better understanding of the behaviors observed in these fittings and evaluate the limitation of the flexible methods, a more comprehensive survey of the fitting performance is performed using multiple initial fitting conditions. In addition, we discuss a new fitting method that utilizes enhanced sampling, which enables better conformational sampling, and therefore increase fitting accuracy.

2. RESULTATS

2.1 Computational methods

Several proteins known to undergo a significant conformational change were considered. In addition, this study specifically focuses on a few proteins for which multiple flexible fitting methods have failed to consistently provide accurate fitting. For each of these proteins, a low-resolution map generated from one of the known conformation of the protein (target). The other known conformation (initial) was deformed to fit the simulated EM map. We used a flexible fitting method based on MD simulation where an additional biasing potential is introduced to guide proteins motions toward the target map, the goal being to maximize the correlation coefficient between the deformed atomic model and the cryo-EM map. The method was implemented within GENESIS [3] and can be run using standard all-atom models or coarse-grained model. For efficiency reasons, all simulations were run using a CA based Go-model. Go model were built using the MMSTSB tool set [4]. For the additional potential, a force constant parameter needs to be set. Simulations were performed using multiple force constants that determine the biasing strength. 16 force constants equally distributed between 100 and 5000 were used in the simulations. In addition, we implemented a new scheme to perform flexible fitting, which relies on replica exchange simulation, a method that enables better conformational sampling [5]. In such approach, 16 simulations are run concurrently and conformations are allowed to exchange between simulations, therefore experiencing different force constant and enhancing sampling (REUS). In addition 20 simulations with random initial velocities were performed. In total, 120 simulations with or without REUS were run and analyzed for each protein.

2.2 Results

To illustrate the results, the specific case of the Ca ATPase is discussed. Ca ATPase exists in two distinct conformations, with an overall conformational difference of 14Å. In previous studies, using one of the conformation as a target low-resolution map (with resolution ranging from 5 to 15), the initial conformation could be only deformed to a model with the minimum of ~4Å RMSD to the target (6Å using MD simulation fitting).

Our simulations with 20 different initial velocities (NOREUS in Figure 1) show that in more than half of the cases, the RMSD indeed, cannot be reduced below 6Å. On the other hand, a few models are very close to the target structure (~2.5Å). This result indicates that MD simulations can provide accurate atomic models but,

because of limited sampling, the optimal model is not always obtained. In order to increase sampling efficiency, we implemented REUS. RMSD obtained with REUS are also shown in Figure 1. With this new method, conformation around ~2.5Å RMSD to the target are consistently obtained. Similar results were obtained on other proteins that were studied.

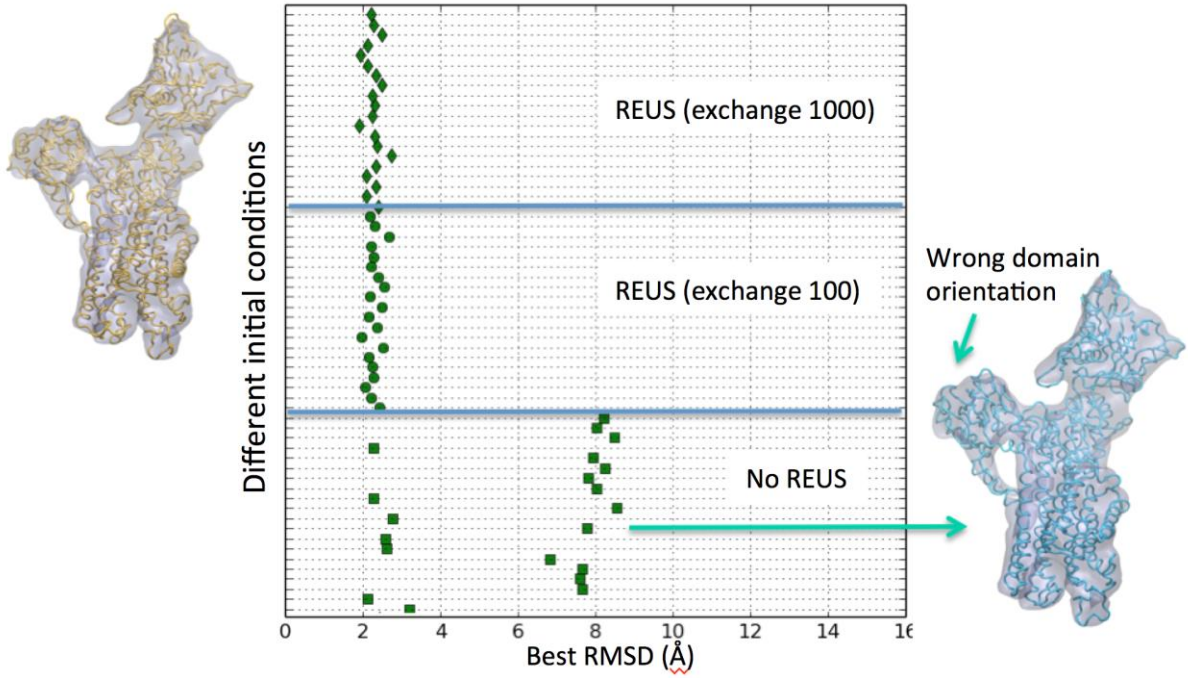

Figure 1. REUS flexible fittings provide consistently atomic models closer to the target structure

3. CONCLUSION

This work shows that flexible fitting methods are indeed reliable and that appropriate conformational sampling is critical to obtain accurate model. The new approach we introduced in this work can enhance conformational sampling and therefore improve accuracy of the models.

REFERENCES

[1] Lawson CL. et al. *Nucleic Acids Research* **39**, D456-464 (2011)

[2] Ahmed A, Tama F. *Journal of Structural Biology* **182**, 67-77 (2013)

[3] http://www.riken.jp/TMS2012/cbp/en/research/software/genesis/index.html

[4] Feig M, Karanicolas J. and Brooks III CL. *Journal of Molecular Graphics and Modeling* **22**, 377-395 (2004)

[5] Sugita Y and Okamoto Y. *Chemical Physics Letters* **314**, 141-151 (1999)