J. Plasma Physics (2006), *vol.* 72, *part* 6, *pp.* 1037–1040.
doi:10.1017/S0022377806005174 Printed in the United Kingdom

Atomic model construction of protein complexes from electron micrographs and visualization of their 3D structure using a virtual reality system

KAZUHIRO NODA[1], MIHO NAKAMURA[1], RYOKO NISHIDA[1], YUKARI YONEDA[1], YOKO YAMAGUCHI[1], YUICHI TAMURA[2], HIROAKI NAKAMURA[2] and TAKUO YASUNAGA[1]

[1]Department of Bioscience and Bioinformatics, Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology, 680-4 Ooaza-kawazu, Iizuka, Fukuoka, Japan

[2]National Institute for Fusion Science, Oroshi-cho 322-6, Toki, Gifu, Japan

(Received 30 August 2005 and accepted 4 January 2006)

Abstract. The methods and applications that integrate electron microscopic density maps and comparative atomic coordinates to generate atomic models of protein complexes and their assemblies have recently advanced. Here, we also report on a newly developed tool for integration between density maps at a spatial resolution of about one nanometer and atomic coordinates by introducing the spatial constraints into a molecular dynamics calculation, which we call spatially restricted molecular dynamics. We, successfully, constructed the atomic models well-fitted to the density maps generated from nine different atomic coordinates of proteins. The method will give us the useful, refined, and presumable atomic structure of macromolecules to elucidate the relationship between their structure and functions.

1. Introduction

Electron microscopy is one of the powerful techniques to elucidate the molecular mechanism of macromolecules such as protein and/or nucleotide complexes in organisms. This is because the technique gives us direct images under physiological conditions and we can reconstruct the three-dimensional structure after image analysis. This direct-imaging ability results from the strong interactions between electrons emitted from the gun and specimens, whereas the interactions also induce damage. Thus, the spatial resolution of the three-dimensional images obtained relatively easily is about a nanometer, but it is too difficult to resolve the atomic coordinates of macromolecules because of low signal-to-noise ratios. Thus, to combine the other biotechnologies such as protein engineering, we need to obtain the three-dimensional structure at the atomic level from low/middle resolutions and visualize their structures with virtual reality (VR). Recently, a number of algorithms and applications have been developed for fitting atomic-resolved structures of assembly components into the lower-resolution density of the whole assembly (Topf and Sali

2005; Topf et al. 2005). Thus, we also have developed a novel algorithm for the construction of atomic models of protein complexes from electron micrographs by integrating steered molecular dynamics simulation: we call the algorithm spatially restricted molecular dynamics (SRMD). Using this SRMD, we succeeded in the atomic model construction of model proteins. Although the simulation may include some artifacts, the structural changes of myosin induced by ATP hydrolysis, could also be calculated using this algorithm. Furthermore, for interpretation of the structure, we have developed a new tool for observation *in silico* using a VR system. The observed results supply us with the newly proposed mechanism of molecular motors. Here, we report on the above progress made for atomic model construction of proteins and their visualization.

2. Materials and methods

The density models ($\rho(x, y, z)$) were prepared from nine different atomic coordinates of proteins from a protein data bank (PDB) as shown in Table 1. The initial models as comparative atomic coordinates were calculated by 3 nm translation and 30° rotation. The external forces ($F_{\text{external}}(x_i, y_i, z_i)$) for the ith atom during molecular dynamics simulation were, then, calculated as a gradient of the density maps ($\rho(x, y, z)$) as follows:

$$F_{\text{external}}(x_i, y_i, z_i) = -\nabla(-\rho(x_i, y_i, z_i)), \tag{2.1}$$

where the contour level of 70 vol.% of proteins is greater than that of ρ. Within the area of 70 vol.% of proteins, F_{external} is zero. This is because $F_{\text{external}}(x_i, y_i, z_i)$ should prevent the shrinkage of proteins. We adopted the combination of molecular dynamics with the external forces defined from electron microscopic density maps (EM maps), whereas most of the reported combination tools adopt rigid body refinement (Topf and Sali 2005) and/or semi-flexible refinements (Tama et al. 2002; Wriggers 2004; Ma 2005).

Using the molecular dynamics calculator, NAMD (Kal et al. 1999), the above external forces were applied to each of atoms during the simulation as a steering force for 1 ps per 3 ps. The time step of calculations was 1 fs. The resolutions of the density maps were changed from 5 nm for weak constraints to 1 nm for strong constraints, next, from 1 nm to 5 nm for 1 ps, which prevented the structure from being trapped in the local energy minima during calculation. In addition to the above external force, we applied simulated annealing method between 300 K and 400 K. The external forces to induce the translational and rotational random walk were also added. The former was introduced to simulate the structural changes for a long time (milliseconds and microseconds), while the latter was added to simulate the translation and rotation over the calculated time ranges. The calculations were performed for about 10 ns using PC(Pentium-IV)/Linux. The fitting procedure was evaluated by the root mean squre of the difference between the model coordinates, which are the answer atomic coordinates, and the final coordinates. During all calculations, water molecules were not added. One reason is that the constraints from EM maps can be effective in preventing the friction due to water molecules and another is shorter calculation time without water. The VMD (Humphrey et al. 1996) and/or our developed tools under Eos (Yasunaga and Wakabayashi 1996) were used for the visualization of the calculated coordinates under UNIX systems. In the developed tools, the VTK libraries (open source, Kitware Inc.) were used.

https://doi.org/10.1017/S0022377806005174 Published online by Cambridge University Press

Table 1. Application to kinds of proteins.

Protein name	ID	M_w (Da)	Initial length (Å)	Final length (Å)
BPTI	1G6X	6,300	7.82	2.35
Myoglobin	102M	16,700	12.84	2.76
Apolipopholin	1AEP	17,700	15.05	3.79
Ras	121P	18,300	8.30	2.28
GFP	1C4F	25,200	10.14	2.97
Actin	1C0F	41,300	16.14	3.82
Arp-2	1K8K	46,000	13.85	3.54
GroEL	1AON	57,800	17.83	3.74
Myosin	1B7T	92,000	16.11	4.52

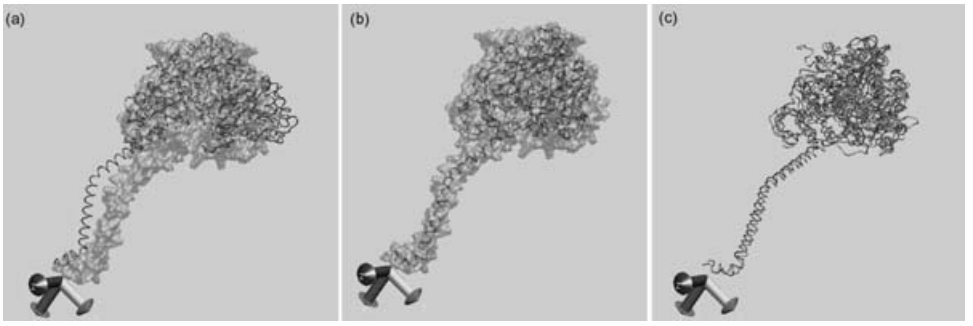

Figure 1. An example of a myosin molecule. The comparison of (a) the initial atomic coordinates or (b) the finally calculated atomic coordinates with the density map shown in contour surfaces. (c) Comparison between the finally calculated atomic coordinates and the answer coordinates of the model protein. The atomic coordinates are shown in light gray tubes and dark gray tubes, respectively. Both tubes are almost superimposed to shaw the good agreement.

The degree of fitting was observed by these applications using the VR system with liquid-crystal shutter glasses.

3. Results and discussion

Here we show one example for a myosin molecule (1B7T) from Table 1. Figure 1(a) shows the contour surfaces of the model density map and the initial model, which was prepared by translation/rotation of the atomic coordinates. The whole of the atomic coordinates stick out of the contour surfaces. After the SRMD we developed, almost all atoms were pushed into the density map as shown in Fig. 1(b). This indicates that the defined external forces should be effective for fitting.

Figure 1(c) shows the comparison between the initial and final atomic coordinates of the protein myosin and so we, in success, obtained the atomic coordinates which are, as a whole, similar to the answer coordinates which produced the density maps. The whole difference between the final model and the answer model was about 4.5 Å, which is smaller than the resolution (1 nm) and indicates good performance of the developed algorithm. The differences are found in the peripheral regions and the long α-helices when observing them with VR software. The former may originate from the flexibility of the protein structure, because the peripheral regions

https://doi.org/10.1017/S0022377806005174 Published online by Cambridge University Press

of proteins are known to be more flexible than the core domain of proteins. The molecular dynamics should simulate such flexiblity.

The latter should come from the shortage of the resolution of the density map, which is difficult to improve. Almost half of the pitch of the helices can be out of those of the answer coordinates as observed in Fig. 1(c). However, we do need not to worry about this type of structural difference because this region has less constraint from the other part of proteins and this structural change also might be possible/true.

Nine different proteins whose molecular weights are from 10 to 200 kDa were tested by the SRMD and we succeeded in creating the final atomic coordinates, which are similar to the answer coordinates within the root mean square of 0.5 nm as a whole (Table 1), which is smaller than the resolution of density maps. This indicates the usefulness of the algorithm.

Interestingly, during 400 K calculations with constraints from EM maps, the protein structure did not break out. This may be because the constraints from EM maps work as a solvent and so prevent the proteins from burn out. The effect of the constraints might be related to a solvent effect such as the excluded volume effect (Irisa 2002).

In the case of myosin molecules, the SRMD could transform a particular structure into another structure (data not shown). This may give us some information such as for the molecular mechanism of motor proteins.

4. Conclusions

We suceeded in developing a new algorithm for the integration of comparative atomic coordinates and molecular dynamics simulation with EM maps at nanometer resolutions. The validity of the obtained structure was less than 0.5 nm, less than the resolution. Observation of the simulation processes from the initial comparative atomic model to the refined model will also give us useful information for elucidating the relationship between the structure and functions.

Acknowledgement

This research was partially supported by the Ministry of Education, Culture, Sports, Science, and Technology, Grant-in-Aid for Exploratory Research, 2003–2004.

References

Humphrey, T., Dalke, A. and Schulten, K. 1996 *J. Mol. Graph.* **14**, 33–38.

Irisa, M. 2002 *J. Biol. Phys.* **28**, 347–357.

Kal, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K. and Schulten, K. 1999 *J. Comput. Phys.* **151**, 283–312.

Ma, J. 2005 *Structure* **13**, 373–380.

Tama, F., Wriggers, W. and Brooks, C. L. III 2002 *J. Mol. Biol.* **321**, 297–305.

Topf, M. and Sali, A. 2005 *Curr. Opinion Struct. Biol.* **15**, 1–8.

Topf, M., Baker, M. L., John, B., Chiu, W. and Sali, A. 2005 *Curr. Opinion Struct. Biol.* **15**, 1–8.

Wriggers, W. 2004 *Structure* **12**, 1–2.

Yasunaga, T. and Wakabayashi, T. 1996 *J. Struct. Biol.* **116**, 155–160.