Analysis of the Function of a Large-scale Supra-biomolecule System by Molecular Dynamics Simulation System, SCUBA (Simulation Codes for hUge Biomolecular Assembly)

Project Representative
Atsushi Matsumoto Japan Atomic Energy Agency

Authors
Hisashi Ishida *1, Kei Yura *1, Takuma Kano *1 and Atsushi Matsumoto *1
*1 Japan Atomic Energy Agency

The Earth Simulator has the highest power ever achieved to perform molecular dynamics simulation of large-scale supra-molecular systems. We are developing a molecular dynamics simulation system, called SCUBA, which is designed to run a system composed of more than a million particles efficiently on parallel computers. In order to understand how the exit tunnel of ribosome regulates the passage of the polypeptide at an atomic level, molecular dynamics simulation of the 70S ribosome in water with and without a nascent polypeptide inside the tunnel. Simulations on this approximately 2,000,000 atoms system were performed using SCUBA. The likely path of the nascent polypeptide through the L4-L22 constriction was determined, and the relationship between global motions of the whole 70S ribosome molecule and the conformation of the exit tunnel were analyzed.

Keywords: large-scale supra-biomolecular MD simulation, 70S ribosome, nascent polypeptide, ribosomal proteins L4 and L22

1. Introduction

Molecular dynamics (MD) simulation not only provides dynamic descriptions of molecules on the atomic scale, but also provides valuable information for the interpretation of experimental data. The rapid development of computer power and the elucidation of the structures of biological macromolecules by X-ray crystallography and other experiments have increased the need for large-scale MD simulations in the field of biology.

We are developing an integrated molecular simulation system for biological macromolecules, called SCUBA (Simulation Codes for hUge Biomolecular Assembly), which is designed to run a system composed of more than a million particles efficiently on parallel computers.

SCUBA has several special features:

1. Topology of biomolecules

The structure of SCUBA's program code is optimized for a system in which the topology of a biomolecular structure is considered.

2. A variety of force field parameters

At present SCUBA can use the AMBER, CHARMM and GROMOS force field. The parameter file for the topology of biomolecules can be obtained from the PDB file by using the input module in SCUBA.

3. A variety of simulation methods

A variety of simulation methods, such as energy minimization, molecular dynamics, free energy calculations, normal mode analysis, principal component analysis and so on, are included.

4. Algorithm for non-cutoff electrostatic interactions

SCUBA utilizes the Particle-Particle Particle-Mesh (PPPM) algorithm, which efficiently calculates all the Coulomb electrostatic interactions [4]. This algorithm reduces the computational time required to calculate the electrostatic forces from the conventional $O(N^2)$ to $O(N\log N)$.

5. Input and output compatibility

SCUBA's input and output file format is currently compatible with those used by AMBER [1].

6. Portability

Written in Fortran90, SCUBA is designed to be easy to read, modify and extend. Users can easily maintain the existing code, and develop the current algorithms and integrate new ones.

7. Control files

The control files for running SCUBA are described in a user-friendly manner.

8. Time-integral algorithms for long time steps

SHAKE, and RATTLE which allow the time step taken to be larger by fixing the bond lengths and angles in the system are available.

9. Time-integral algorithms with high accuracy

The Martyna-Klein-Tuckerman (MKT) algorithm which produces the correct ensemble thermodynamically is available. The MKT algorithm was extended to utilize the multiple time step (MTS) method, which increases the time step length significantly [5, 6].

10. Parallelization

SCUBA employs the domain decomposition (DD) method, which divides the volume of the physical system into rectangular subcells with a length longer than the potential cutoff radius. The processor assigned to a subcell needs to evaluate the interactions between the atoms in the subcell and between the atoms in 26 neighboring subcells. SCUBA employs the method for minimizing communication between processors proposed by D. Brown [7], which enables the number of processors between which data must be transferred to be reduced to only 7 of the neighboring subcells.

11. Vectorization

In order to improve the performance of SCUBA on the Earth Simulator, the algorithm to calculate the interactions among the atoms is intensively vectorized.

12. Dynamic load balance

To overcome the load imbalance associated with irregular atomic distribution, a dynamic load-balancing algorithm is implemented. Moreover, the number of processors used to calculate the PM part of PPPM can be optimized to minimize the computation by allowing the number to change.

13. High performance

By intensive parallelization and vectorization, and by using the dynamic load balance mentioned in 10, 11, and 12, SCUBA has achieved both a high parallelization efficiency ratio and a high vectorization ratio. SCUBA has achieved a parallelization efficiency ratio of 75.8%, and a vectorization ratio of 96.2% even 45 nodes (360 processors) were used to perform an MD simulation for a system of RuvAB-Holliday junction complex which consisted of 546,725 atoms. The performance of SCUBA on the Earth Simulator is shown in Fig. 1.

14. Optimization of memory use

The arrays used in the program of SCUBA are intensively optimized to reduce the amount of memory use. This optimization enables SCUBA to perform molecular dynamics simulations of large-scale supra-molecular systems comprised of more than a million atoms on the Earth Simulator.

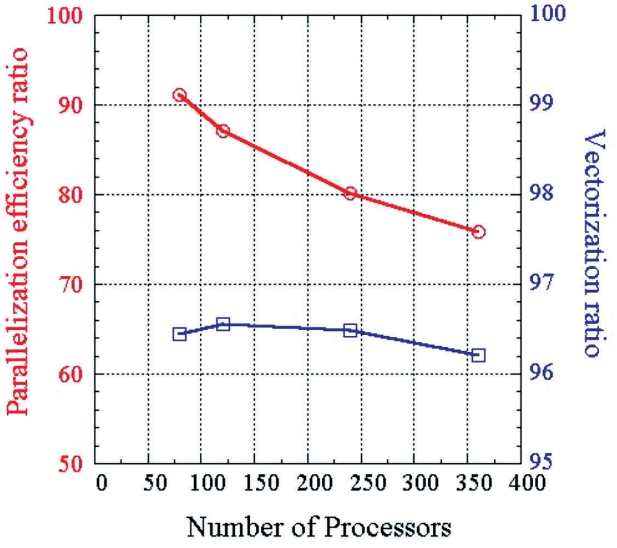

Fig. 1 Present performance of SCUBA on the Earth Simulator. The parallelization efficiency ratio and vectorization ratio are drawn in red and blue, respectively.

2. Molecular dynamics simulation of the 70S ribosome

2-1. Introduction

Ribosome is one of the supra-biomolecules used in the process of translating genetic information for the synthesis of polypeptides. The 70S ribosome from eubacteria is composed of a small (30S) and a large (50S) subunits. The 30S subunit decodes genetic information, and the 50S subunit is considered to be responsible for the formation of peptide bonds, and the elongation of the nascent polypeptide. The nascent peptide is generated at the peptidyl transferase center (PTC) of the 50S subunit, and passes through a tunnel, which starts at the PTC and continues through the 50S subunit.

Modeling of a nascent polypeptide in the tunnel of the 70S ribosome in the crystal form has shown that the extended loops of L4 and L22 partially hinder the passage of the polypeptide. Figure 2 shows that with 50S on top and 30S underneath, modeling revealed two possible paths for the polypeptide: one over Arg92 of L22 and one under Arg92 [8].

In our study, to understand how the exit tunnel of ribosome regulates the passage of the polypeptide at an atomic level and the role of global motions in the conformation of the exit tunnel, all-atom molecular dynamics (MD) simulations were performed on a *Thermus thermophilus* 70S ribosome in water with and without a nascent polypeptide inside the tunnel. The likely path of the nascent polypeptide through the L4-L22 constriction was determined, and the relationship between global motions of the whole 70S ribosome molecule and the conformation of the exit tunnel were analyzed.

We used the atomic structure of *Thermus thermophilus* 70S ribosome, which was determined by X-ray crystallography (PDB code: 1YL3 and 1YL4, resolution: 5.5 Å) [9]. Using SCUBA, we have carried out molecular dynamic simulations

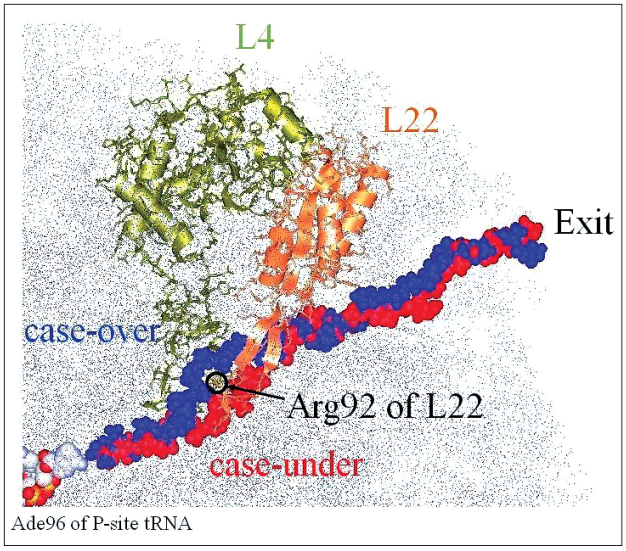

Fig. 2 Two of the modeled polyalanines inside the exit tunnel. L4 and L22 (wire-ribbon models) are shown in dark khaki and coral, respectively. From the PTC, two nascent polyalanines (space-filling models) pass through the exit tunnel in the large subunit, 50S (dot model), passing over and under Arg92 of L22. They are shown in blue and red, respectively.

of the 70S ribosome with and without the nascent polypeptide inside the exit tunnel. Each system of the 70S ribosome comprised 1,878,425 atoms including water molecules. The MD simulations were carried out for several nanoseconds at a constant pressure of one bar and a temperature of 350 K.

2-2. Results

We analyzed the conformation of a large tunnel around the L4-L22 gate in three cases; one where the polypeptide was positioned over the L4-L22 gate (case-over), one where the polypeptide was positioned under the L4-L22 gate (case-under), and one where no polypeptide was inside the tunnel (case-no). Figure 3 shows a snap shot of the structure around the L4-L22 constriction in case-under and case-over. (The structure around the L4-L22 constriction in case-no was similar to that in case-over.) In case-no and case-over, strong interactions among Phe61, Gln67 and Asp74 of L24 and Arg92 of L22 were observed. In case-under, however, these interactions between them were not observed, and Arg92 repositioned to interact with Ade2059 of 23S rRNA. As Ade2059 and the neighboring residue Ade2058 are well known to be important for interacting with some antibiotics such as erythromycin. This result suggests that the tunnel is closed by the Arg92-L4 interaction before elongation of the polypeptide and the tunnel leads the entering polypeptide from the PTC to the passage under Arg92, causing Arg92 to switch to an open position. It is possible, therefore, that Arg92 plays the role of a gate, opening and closing the tunnel at L4-L22.

The average atomic fluctuations of the whole of the 70S ribosome in the thermal equilibrium in case-under are shown in Fig. 4. Figure 4 shows that the region near the exit of the large tunnel had small atomic fluctuations, while the L7/L12 stalk and L1 stalk were found to have especially large atomic fluctuations. It has been suggested that the conformational change of the 70S ribosome may be accompanied by changes in tunnel conformation [10], such as would be the case for a peristaltic pump facilitating the movement of the polyalanine along the tunnel. However, Voss et al. [11,12] have concluded that the tunnel is rigid. We performed principal component analysis (PCA) to understand the role of

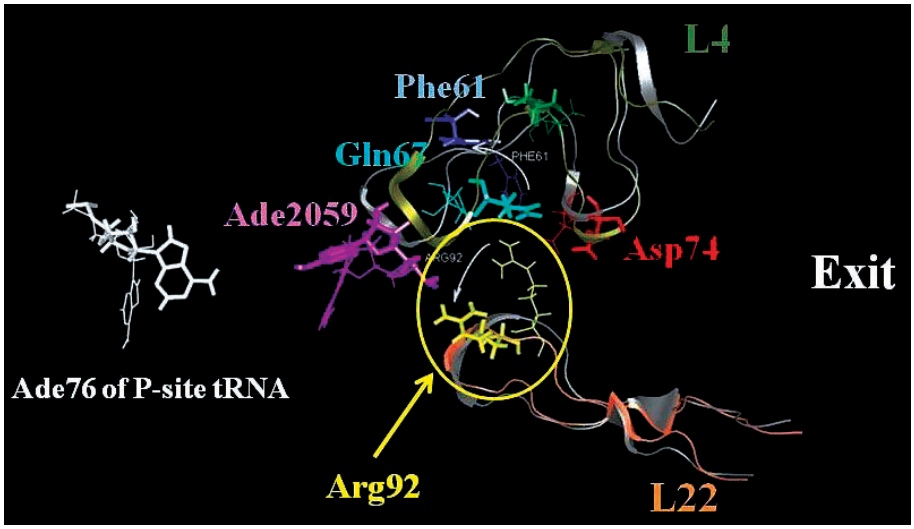

Fig. 3 Superposition of the conformations around the L4-L22 constriction at 1ns of the free MD simulation in case-over and case-under. In case-over, L4 and L22 are shown in white, and residues are depicted as thin wire models. In order to keep the image clear, the polypeptides have not been included.

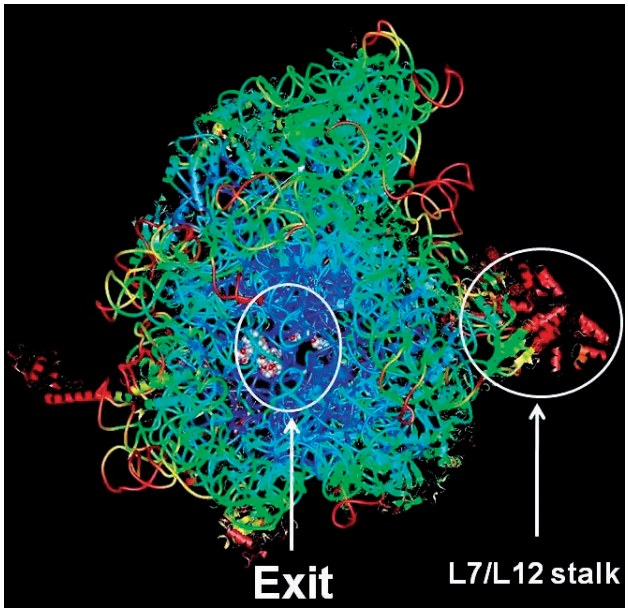

Fig. 4 The average RMSF of the 70S ribosome are color coded. High amplitude thermal fluctuations are shown in red (more than 3.5 Å), and low amplitude fluctuations are shown in blue (less than 0.8 Å). A top view of the 50S ribosome is shown.

global motions in the conformation of the exit tunnel. PCA showed that the main modes of motion are global motions mainly involving the relative movement of the 50S and 30S subunits [8]. These motions may be important for the translocation of tRNA molecules between 50S and 30S, but did not reveal any conformational changes within the tunnel such as a peristaltic pump to facilitate the movement of the polypeptide.

3. Conclusion

We developed an MD simulation system, SCUBA, which achieves a high parallelization efficiency ratio and a high vectorization ratio on the earth simulator. Our simulations of the 70S ribosome revealed the likely path of nascent polypeptide in the exit tunnel to be under Arg92 of L22 [8].

4. Acknowledgements

We are grateful to Dr. Steven Hayward and Dr. Guru Poornam at the University of East Anglia for helpful discussion regarding the conformational analysis of the ribosome.

References

[1] D.A. Pearlman, D.A. Case, J.W. Caldwell, W.R. Ross, T.E. Cheatham, III, S. DeBolt, D. Ferguson, G. Seibel and P. Kollman. AMBER, a computer program for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to elucidate the structures and energies of molecules. *Comp. Phys. Commun*. **91**, 1–41 (1995).

[2] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations, *J. Comp. Chem*. **4**, 187–217 (1983).

[3] W. F. van Gunsteren and H. J. C. Berendsen. GROMOS: GROningen MOlecular Simulation software. Technical report, Laboratory of Physical Chemistry, University of Groningen, Nijenborgh, The Netherlands, (1988).

[4] R.W. Hockney and J.W. Eastwood. Computer Simulation Using Particles. McGraw-Hill, NY, (1981).

[5] G. J. Martyna, M. E. Tuckerman, D. J. Tobias and M. L. Klein. Explicit reversible integrators for extended systems dynamics, *Mol. Phys*. **87**, 1117–1157 (1996).

[6] R. Zhou, E. Harder, H. Xu and B. J. Berne. Efficient multiple time step method for use with Ewald and particle mesh Ewald for large biomolecular systems, *J. Chem. Phys*. **87**, 1117–1157 (1996).

[7] D. Brown, J.H.R. Clarke, M. Okuda and T. Yamazaki. A domain decomposition parallelization strategy for molecular dynamics simulations on distributed memory machines. *Comp. Phys. Commun*. **74**, 67–80 (1993).

[8] H. Ishida and S. Hayward. Likely path of nascent polypeptide in exit tunnel revealed by molecular dynamics simulation of ribosome, submitted to *Biophysical Journal*.

[9] L. Jenner, P. Romby, B. Rees, C. Schulze-Briese, M. Springer, C. Ehresmann, B. Ehresmann, D. Moras, G. Yusupova. and M. Yusupov. Translational Operator of mRNA on the Ribosome: How Repressor Proteins Exclude Ribosome Binding. *Science* **308**, 120–123 (2005).

[10] I. S. Gabashvili, S. T. Gregory, M. Valle, R. Grassucci, M. Worbs, M. C. Wahl, A. E. Dahlberg and J. Frank. The polypeptide tunnel system in the ribosome and its gating in erythromycin resistance mutants of L4 and L22. *Mol. Cell* **8**, 181–188 (2001).

[11] N. R. Voss, M. Gerstein, T. A. Steitz and P. B. Moore. The geometry of the ribosomal polypeptide exit tunnel. *J. Mol. Biol*. **360**, 893–906 (2006).

[12] T. A. Steitz. A structural understanding of the dynamic ribosome machine. *Nat. Rev. Mol. Cell Biol*. **9**, 242–253 (2008).

分子動力学シミュレーションシステムSCUBA（Simulation Codes of hUge Biomolecule Assembly）を用いた大規模生体超分子系の機能解析

プロジェクト責任者
松本　淳　日本原子力研究開発機構

著者
石田　恒*1，由良　敬*1，叶野　琢磨*1，松本　淳*1
*1　日本原子力研究開発機構

地球シミュレータは従来にはない大規模生体超分子系の分子動力学シミュレーションを可能とする計算能力をもつ。我々は生体超分子系を扱う大規模な分子動力学シミュレーションシステムSCUBAを開発している。SCUBAは長距離相互作用を高速かつ高精度に計算するPPPM計算法、系のエネルギー、温度、圧力を一定に保つ様々な時間積分アルゴリズムなど、最新のアルゴリズムを採用した計算性能に優れたシミュレーションシステムである。SCUBAは、地球シミュレータ360プロセッサ使用時でベクトル化率95％以上、並列化効率50％以上の優れた性能を達成している。さらに、SCUBAが用いる配列を機能別にモジュール化し、メモリ使用量を最適化することで、約300万原子からなる系でも分子動力学シミュレーションが実行可能である。

本年度は、約200万原子からなるリボソーム（遺伝情報を翻訳する生体超分子）と新生ポリペプチドの複合体の系について分子動力学シミュレーションを地球シミュレータ上で実行した。結果、翻訳されたタンパク質（新生ペプチド）が通るリボソームトンネルの一部を構成しているL4タンパク質とL22タンパク質の相互作用がトンネルの開閉状態を制御していることがわかった。

キーワード：大規模生体超分子動力学シミュレーション，リボソーム，新生ポリペプチド，リボソームタンパク質L4とL22