

# Protein Explorer: A Petaflops Special-Purpose Computer for Molecular Dynamics Simulations

**Makoto Taiji**<sup>1</sup>

taiji@gsc.riken.go.jp

**Tetsu Narumi**<sup>1</sup>

narumi@gsc.riken.go.jp

**Yousuke Ohno**<sup>1</sup>

ohno@gsc.riken.go.jp

**Noriyuki Futatsugi**<sup>1</sup>

foota@gsc.riken.go.jp

**Atsushi Suenaga**<sup>1</sup>

suenaga@gsc.riken.go.jp

**Naoki Takada**<sup>1</sup>

ntakada@gsc.riken.go.jp

**Akihiko Konagaya**<sup>2</sup>

konagaya@gsc.riken.go.jp

<sup>1</sup> High Performance Biocomputing Research Team, Bioinformatics Group, Genomic Sciences Center, RIKEN, 214 Maedacho, Totsuka, Yokohama, Kanagawa 244-0804, Japan

<sup>2</sup> Bioinformatics Group, Genomic Sciences Center, RIKEN, 1-7-22 Suehiro, Tsurumi, Yokohama, Kanagawa 230-0045, Japan

**Keywords:** molecular dynamics simulation, special-purpose computer

## 1 Introduction

Molecular dynamics (MD) simulation is one of the most exact and powerful method to study protein biophysics and biochemistry. However, its use is still limited by its high requirement of computational power. To solve this problem, several special-purpose computers have been built. The most successful machines are GRAPE family [1], GRAPE-2A, MD-GRAPE, and MDM. Especially, MDM [2] with 78 teraflops is still keeping the fastest computer in the world since 2000. Based on these success, we have started a project to build a special-purpose computer with petaflops performance, ‘Protein Explorer’. It will become the first petaflops machine in the world. The scientific targets are the prediction of affinity of (mutated) proteins and simulations of huge biomolecules to clarify its functions and mechanisms. The petaflops dedicated machine will give great contributions to bioscience.

## 2 Hardware

The system design of the Protein Explorer is similar to those of MD-GRAPE/MDM. The Protein Explorer accelerates calculation of non-bonding forces which dominates MD simulations, i.e. Coulomb and Van der Waals forces. The key ideas to achieve high performance are dedicated pipelines and broadcast memory parallelization. A dedicated pipeline maximize the circuit efficiency since we can choose the most suitable precision for each arithmetic unit and the most efficient scheduling can be realized by pipelined operations. The broadcast memory parallelization is the architecture in which all parallel pipelines share a common memory unit and use the same data for every pipeline. It enables highly-parallel calculations with low bandwidth. This special architecture is not always valid but powerful for applications like molecular dynamics. Figure 1 shows the block diagram of MDGRAPE-3 chip, the pipeline LSI for the Protein Explorer. It has 20 pipelines and each pipeline calculate a pairwise interaction each cycle. Thus, the nominal peak performance of MDGRAPE-3 chip will reach 200 gigaflops in the worst case, 12 times faster than MDGRAPE-2 chip used in MDM. The other improvements are the inclusion of a memory for particle coordinates and communication circuits between chips. They are necessary for efficient operation of highly-parallel systems.

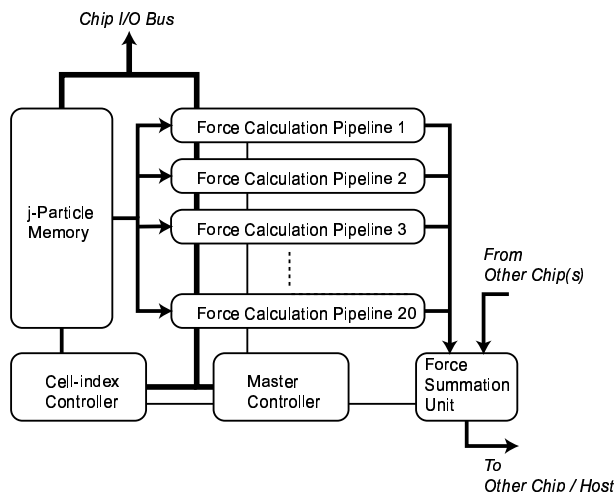


Figure 1: Block diagram of MDGRAPE-3 chip.

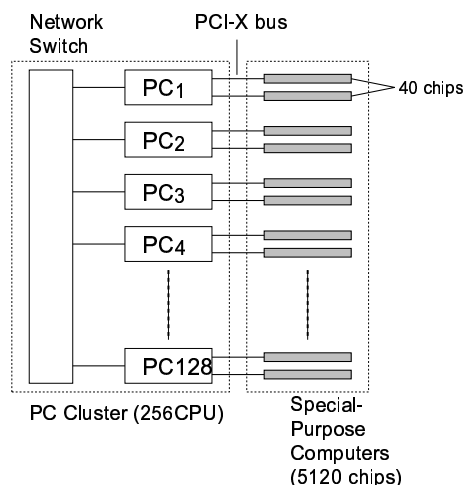


Figure 2: Block diagram of Protein Explorer system.

Figure 2 shows the block diagram of the Protein Explorer system. It consists of the host computer system and the special-purpose engines. We plan to use a PC cluster with 128 nodes for the host computer. The dedicated special-purpose engines are connected with each PC nodes by PCI-X bus. Each node will have 8 teraflops performance with 40 MDGRAPE-3 chips. The entire system will have 5,120 chips in total to achieve the peak performance of 1 petaflops.

### 3 Software

On the MDM system, we have already ported the most of modules in AMBER-6 [3], one of the most popular molecular dynamics package. We are now working on CHARMM [5] and plan to port NAMD [4]. Since the Protein Explorer works as force calculation subroutines and is quite similar to MDM in high-level programming, these codes on MDM are easily modified for it. The other MD packages are also not very difficult to be ported for a single-node system at least.

### Acknowledgments

This work is partially supported by the contracted research ‘Protein 3000 Project’ by Ministry of Education, Culture, Sports, Science and Technology in Japan.

### References

- [1] Makino, J. and Taiji, M., *Scientific Simulations with Special-Purpose Computers*, John Wiley & Sons, 1998.
- [2] Narumi, T., Susukita, R., Ebisuzaki, T., McNiven, G., and Elmegreen, B., Molecular dynamics machine: Special-purpose computer for molecular dynamics simulations, *Molecular Simulation*, 21:401–415, 1999.
- [3] <http://www.amber.ucsf.edu/amber/amber.html>
- [4] <http://www.ks.uiuc.edu/Research/namd/>
- [5] <http://www.scripps.edu/brooks/>