Molecular Simulation of DNA Microarrays – an Application of 2D particle mesh Ewald algorithm –

T. Watanabe[*], M. Kawata[**], and U. Nagashima[*,**]

[*]Japan Science and Technology Corporation (JST), 4-1-8 Honcho, Kawaguchi 332-0012, Japan, donjo@ni.aist.go.jp
[**]Grid Technology Research Center, National Institute of Advanced Industrial Science and Technology, 1-1-1 Umezono, Tsukuba, 305-8568 Japan

ABSTRACT

We have developed the two-dimensional particle mesh Ewald (2D-PME) algorithm for molecular dynamics (MD) simulation, and applied it to have knowledge of the behavior of DNA and its interaction on microarrays, which crucial in the design of DNA microarrays. The 2D-PME algorithm can rapidly calculate the Coulomb interaction for a 3D system with 2D periodicity, which is the most time-consuming process in the MD simulation.

The MD simulation of DNA microarrays that contain a single DNA molecule in a unit-cell was carried out under the two types of the boundary conditions, 3D systems with the 2D and 3D periodicities, in which the Coulomb interactions were calculated by using 2D-PME and 3D-PME methods, respectively. Effect of the infinite replicas in the vertical direction to the microarray surface in 3D periodic system was examined based on the results of the 3D system with 2D periodicity. Comparison of the computational efficiency and accuracy between the 2D-PME and 3D-PME methods revealed the availability of the 2D-PME method.

***Keywords*:** molecular dynamics, DNA microarrays, two-dimensional particle mesh ewald, 2D-PME

1 INTRODUCTION

DNA microarrays have attracted widespread attention in various areas such as drug design and genetic screening/sequencing. Although knowledge of the behavior of DNA and its interaction on microarrays is crucial in the design of DNA microarrays, such knowledge is still lacking.

Molecular dynamics (MD) simulation has been applied to have such knowledge at the molecular level. Wong and Pettitt, for example, executed MD simulation of DNA microarrays by using an all-atom model.[1] They applied novel periodic boundary condition, the glide-plane boundary condition.[2] Their system has the 3D periodicity, thus also had the lamellar structure. The effect of the infinite replicas in the direction vertical to the microarray surface is not negligiable. One of the solutions for this lamellar structure is the use of the 3D system with 2D periodicity. On the other hand, there were sophisticated methods to treat the Coulomb interaction in the 3D periodic system,[3,4] in contrast to the 3D system with the 2D periodicity.

Recently, Kawata et al[5,6] were developed computationally efficient method, two-dimensional particle mesh Ewald method (2D-PME), to calculate the Coulomb interaction in the 3D system with 2D periodicity. The 2D-PME method was not only computationally efficient but also gave the accrate energy and force[5,6]. There was only preliminary result but no application to the practical MD simulation.

The 2D-PME algorithm was applied to the MD simulation of DNA microarray systems to verify that the method can simulate efficiently the 3D systems with 2D periodicity. Effect of the infinite replicas in the vertical direction to the microarray surface in 3D periodic system was also examined based on the results of the 3D system with 2D periodicity.

2 METHOD

2.1 Two Dimensional Particle Mesh Ewald Algorithm

In 2D-PME method, the Coulomb potential energy, V, for a 3D system of N charged particles, $\mathbf{r}_1$, $\mathbf{r}_2$,…, $\mathbf{r}_N$, satisfying the neutrality condition, $q_1+q_2+\cdots+q_N=0$, with 2D periodicity can be expressed as

$$V = \frac{1}{2}\sum_{\mathbf{h}_x,\mathbf{h}_y}{}'\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_i q_j}{\left|\mathbf{r}_i-\mathbf{r}_j+\mathbf{h}_x+\mathbf{h}_y\right|} \cong V^r + V^k_{\mathbf{k}\neq 0} + V^k_{\mathbf{k}=0} + V^s, \quad (1)$$

where $\mathbf{h}_x$ and $\mathbf{h}_y$ are the 2D lattice vectors, the prime indicates that terms with $i=j$ are omitted when $\mathbf{h}_x+\mathbf{h}_y=\mathbf{0}$, and $\mathbf{k}$ is the reciprocal vector. V^r is the contribution from the real space sum, $V^k_{\mathbf{k}\neq 0}$ and $V^k_{\mathbf{k}=0}$ are the contributions from the reciprocal space sums with and without the $\mathbf{k}=\mathbf{0}$, respectively, and V^s is the self-correction term. Here, a Gaussian charge

distribution, such as $\rho(\mathbf{r}_i) = q_i \alpha^3 \pi^{-3/2} \exp(-\alpha^2 |\mathbf{r}_i|^2)$, is used in the charge decomposition, α is a screening parameter. These terms can be written as

$$V^r = \frac{1}{2} \sum_{\mathbf{h}_x, \mathbf{h}_y}{}' \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{q_i q_j}{|\mathbf{r}_{ij} + \mathbf{h}_x + \mathbf{h}_y|} \times erfc\left(\alpha |\mathbf{r}_{ij} + \mathbf{h}_x + \mathbf{h}_y|\right) \quad (2)$$

where *erfc* represents the complementary error function.

$$V^k_{\mathbf{k}\neq 0} = \frac{1}{|\mathbf{h}_x \times \mathbf{h}_y|} \sum_{m_x=-G_x/2}^{G_x/2} \sum_{m_y=-G_y/2}^{G_y/2} \int_{-\infty}^{\infty} \mathrm{d}h \left| C\left(\frac{2\pi m_x}{G_x}, n_x\right) \right|^2 \times \left| C\left(\frac{2\pi m_y}{G_y}, n_y\right) \right|^2 |C(h, n_z)|^2 \frac{\left|\tilde{Q}(m_x, m_y, h)\right|^2}{4\pi^2\left(m_x^2 + m_y^2\right) + h^2} \times \exp\left\{ -\frac{4\pi^2\left(m_x^2 + m_y^2\right) + h^2}{4\alpha^2} \right\} \quad (3)$$

G_x and G_y are the number of grid points for the FFTs in the *x*- and *y*-direction, respectively, and n_λ is the order of the B-splines in the λ-direction. $C(v,w)$ is defined as

$$C(v, w) \equiv \frac{\exp[\mathrm{i}\,(w-1)v]}{\sum_{t=0}^{w-2} M_w(t+1) \exp(\mathrm{i}vt)} \quad (4)$$

where M_w is cardinal B-splines. $\tilde{Q}(m_x, m_y, h)$ is the Fourier integral of $Q'(m_x, m_y, t_z)$ with respect to t_z and $Q'(m_x, m_y, t_z)$ is the 2D FFTs of $Q(t_x, t_y, t_z)$ with respect to t_x and t_y. $Q(t_x, t_y, t_z)$ is a charge distribution calculated by using the B-splines.

$$V^k_{\mathbf{k}=0} = \sum_{i=1}^{N} q_i B\left[V^k_{\mathbf{k}=0\,g}, z_i\right] \quad (5)$$

where $B[V^k_{\mathbf{k}=0\,g}, z_i]$ is the Coulomb potential energy at position z_i, interpolated with a B-spline interpolation by using the values of $V^k_{\mathbf{k}=0\,g} V^k_{\mathbf{k}=0\,g}, \ldots, V^k_{\mathbf{k}=0\,Ng}$, which the Coulomb potential energy on the *g*th grid point (g=1,…,N_g) in the *z* direction.

$$V^S = -\frac{\alpha}{\sqrt{\pi}} \sum_{i=1}^{N} {q_i}^2 \quad (6)$$

These equations are same with the two-dimensional Ewald method (2D-EW) [7,8] except Eq (3). The overall computational efficiency of the 2D-EW method strongly depends on the computational efficiency of the reciprocal-space sum. In the 2D-PME method, the calculation of the reciprocal-space sum without the **k=0** term can be accelerated by using the B-spline interpolations and the FFT.

2.2 Molecular Dynamics Simulation

The 2D-PME algorithm was implemented into the NAMD Version 2.5 program package, [9] which is widely used. The NAMD have implemented both the 3D-PME method and the lattice for 3D system with 2D periodicity. In the implementation of the 2D-PME method, the classes and data structure for the 3D-PME in the NAMD program package were fully reused.

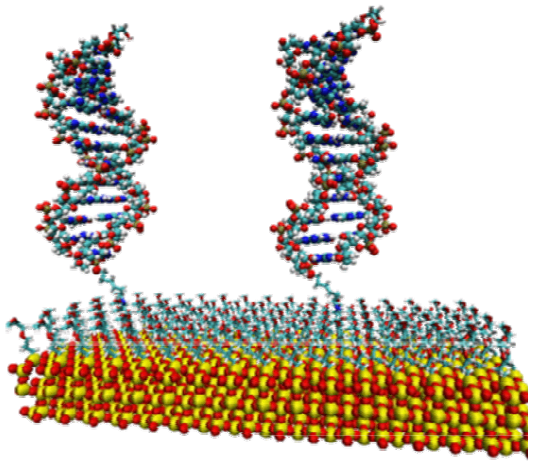

Figure 1. The system used in the MD simulation. This figure contains two unit-cells. Water molecules and the sodium and chloride ions were removed for visibility.

All calculation was carried out on the KOUME and UME PC-clusters at Grid Technology Research Center, National Institute of Advanced Industrial Science and Technology.

3 RESULTS AND DISCUSSION

The model system was essentially same with Ref.1 but used the pseudo-2D boundary condition instead of the glide-plane boundary condition. The system constructed on silica layer of β-cristobalite plate 1.2nm thick, of which a (1,1,1) surface was covered with sixty-four epoxides. A DNA strand d(CGTGTCCCTCTC) hybridized with its complement was tethered to the terminal of an epoxide by the amine linker. The empty space of the simulation box was filled with the water molecules and ions. This system is free from the infinite replicas in the vertical direction to the microarray surface, and has a gas-liquid interface in the non-periodic direction, *z*-direction. The water molecules which vapor into gas phase were inverted sign of z component of their velocity at the distance from a gas-liquid interface by 1.5 nm. The initial system without water molecules and ions was displayed in Figure 1. The potential parameters were used in the CHARMM22 proteins and CHARMM27 nucleic acids.

For purpose of comparison, the 3D periodic system of the DNA microarray was also constructed. The difference with the 3D system with 2D periodicity was presence of a silica-liquid interface instead of the

gas-liquid interface and a lamellar structure. The total number of the molecules in the 3D periodic system was same with the 3D system with 2D periodicity. The calculation of the Coulomb interaction was carried out by the 3D-PME method implemented in the NAMD program package.

The main difference of the computational efficiency between the 3D system with 2D periodicity and the 3D periodic system was due to the structural difference described above, such as the lamellar structure and the silica-liquid interface, however, the total computational efficiencies of the 2D-PME and 3D-PME methods were about the same.

3 CONCLUSION

The two-dimensional particle mesh Ewald (2D-PME) algorithm was implemented into the NAMD program package for carrying out the practical MD simulation. The 2D-PME method implemented into the NAMD program package was applied to the MD simulation of the DNA microarrays. Two systems for the DNA microarrays were constructed for the 3D system with 2D periodicity and for the 3D periodic system. As these results, the 2D-PME method was free from the problem caused by the lamellar structure and the computational efficiency was comparable with the 3D-PME method. These results indicated the 2D-PME method had large possibility to be standard method to calculate the Coulomb interaction in the simulation for the interface, membrane, and surface.

Acknowledgements

This work was supported in part by the grant from ACT-JST ("Research and Development for Applying Advanced Computational Science and Technology" of Japan Science and Technology Corporation).

References

[1] K.-Y. Wong and B. M. Pettitt, *Theor. Chem. Acc.*, **106**, 233 (2001).
[2] K.-Y. Wong, B. M. Pettitt, *Chem. Phys. Lett.*, **326**, 193 (2000).
[3] T. A. Darden, D. M. York, L. G. Pedersen, *J. Chem. Phys.*, **98**, 10089 (1993).
[4] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.*, **103**, 8577 (1995).
[5] M. Kawata and U. Nagashima, *Chem. Phys. Lett.*, **340**, 165 (2001).
[6] M. Kawata and U. Nagashima, *J. Chem. Phys.*, **116**, 3430 (2002).
[7] M. Kawata and M. Mikami, *Chem. Phys. Lett.*, **340**, 157 (2001).
[8] M. Kawata, M. Mikami, and U. Nagashima, *J. Chem. Phys.*, **115** 4457 (2001).
[9] L. Kale, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten, *J. Comp. Chem.*, **151**, 283 (1999). (NAMD)
[10] D. MacKerel Jr., D. Bashford, M. Bellott, R. L. Dumbrack Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B*, **102**, 3586 (1998).
[11] N. Foloppe, A. D. MacKerell Jr., *J. Comput. Chem.*, **21**, 86 (2000).