A HIGH PERFORMANCE SYSTEM FOR MOLECULAR DYNAMICS SIMULATION OF BIOMOLECULES USING A SPECIAL-PURPOSE COMPUTER

Y. KOMEIJI, H. YOKOYAMA
Electrotechnical Laboratory, 1-1-4 Umezono, Tsukuba-shi, Ibaraki 305, Japan
M. UEBAYASI
National Institute of Bioscience and Human Technology 1-1-3 Higashi, Tsukuba-shi, Ibaraki 305, Japan
M. TAIJI, T. FUKUSHIGE, D. SUGIMOTO
College of Arts and Science, The University of Tokyo, Komaba, Meguro-ku, Tokyo, 153, Japan
R. TAKATA, A. SHIMIZU, K. ITSUKASHI
ITL Corp., 1-8-6 Higashitsutsujigaoka, Chofu-shi, Tokyo 182, Japan

GRAPE (GRavity PipE) processors are special purpose computers for simulation of classical particles. The performance of MD-GRAPE, one of the GRAPEs developed for molecular dynamics, was investigated. The effective speed of MD-GRAPE was equivalent to ~6 Gflops. The precision of MD-GRAPE was good judging from the acceptable fluctuation of the total energy. Then a software named PEACH (Program for Energetic Analysis of bioCHemical molecules) was developed for molecular dynamics of biomolecules in combination with MD-GRAPE. Molecular dynamics simulation was performed for several protein-solvent systems with different sizes. Simulation of the largest system investigated (27,000 atoms) took only 5 sec/step. Thus, the PEACH-GRAPE system is expected to be useful in accurate and reliable simulation of large biomolecules.

1 Introduction

The physico-chemical nature of structure-dynamics-function relationships of biomolecules such as proteins, nucleotides, and lipids draws the attention of scientists from many disciplines including biology, chemistry, and physics. The molecular complexity of irregular structures obscures theoretical analysis, but recent advancement of computation has enabled us to simulate biomolecules based on theoretical physics. Molecular dynamics simulation (MD) has been recognized as a promising approach [1], which gives both dynamic and thermodynamic information on the molecules.

Computational demand for MD of biomolecules is enormous, however. Almost all the computation time is consumed in calculation of Coulomb interactions. Hence, in most MD studies of biomolecules, the cut-off scheme is applied to Coulomb interactions; the interaction between two atoms separated farther than a specified cut-off radius (usually ~10 Å) is ignored. Though the cut-off method was unavoidable for simulation of large biomolecular systems within limited computer resource, several serious artifacts caused by the method have been reported in the recent literature [2-6]. Therefore, it is desirable if we can perform MD without cut-off while not increasing the computation time. Special purpose computers have made this possible.

The machine we used in this report, MD-GRAPE [7], belongs to a series of special purpose computers for N-body simulations of classical particles named GRAPE (GRAvity PipE) [8, 9]. GRAPE was originally developed for gravitational problems, but so far its application has been expanded to other kinds of simulations such as fluid dynamics[10], MD [11], and so on. Generally, GRAPE is connected to a host such as a workstation or a personal computer. An application program written in a high level language is implemented in the host, and GRAPE is called from the host through an interface software library. In N-body simulations, GRAPE computes tasks proportional to N^2, while the host computes those proportional to N, such as integration of the equation of motions or calculation of bonded interactions. MD-GRAPE is a high precision type GRAPE which calculates any central force or potential. The architecture of MD-GRAPE is based on GRAPE2A, the preceding machine for MD [11].

The advantages of the GRAPE type machines over several other special purpose computers for MD have been discussed elsewhere [7, 11], and we give only the main points. First of all, GRAPE calculates only simple but time-consuming part of the simulation. Hence, the architecture of GRAPE is simple and easy to develop. Also, it is easy to make GRAPE faster by massive parallelization. Secondly, GRAPE retains some flexilibily, for any kind of central force could be treated. This feature will be discussed in detail in Sec. 2. Thirdly, the software for GRAPE is fairly easy to develop because GRAPE is equivalent to a subroutine to calculate nonbonded interactions [12].

Below we give a typical example of potential function U of biomolecules used for MD (AMBER/OPLS) [13, 14]:

$$\begin{aligned} U &= \sum_{\text{bonds}} K_R (R - R_0)^2 && \text{(Bond)} \\ &+ \sum_{\text{angles}} K_\theta (\theta - \theta_0)^2 && \text{(Angle)} \\ &+ \sum_{\text{torsions}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] && \text{(Torsion)} \\ &+ \sum_{i>j} [\frac{a_{ij}}{R_{ij}^{12}} - \frac{b_{ij}}{R_{ij}^{6}}] && \text{(VDW)} \\ &+ \sum_{i>j} \frac{Q_i Q_j}{R_{ij}} && \text{(Coulomb)} \qquad (1) \end{aligned}$$

The term VDW (Van Der Waals) indicates Lennard-Jones 6-12 potential throughout this paper. VDW and Coulomb interactions, called nonbonded interactions, are calculated for all of the atom pairs except those having the bond, angle, or torsion interactions. We try to reduce the computational demand of MD by calculating the nonbonded interactions by MD-GRAPE.

In this report, we give the hardware architecture, usage, and performance of MD-GRAPE. Then we describe PEACH (Program for Energetic Analysis of bioCHemical molecules), a newly developed software for MD of proteins and other biomolecules using MD-GRAPE. We finally present the performance of the PEACH-GRAPE system when applied to simulation of solvated proteins.

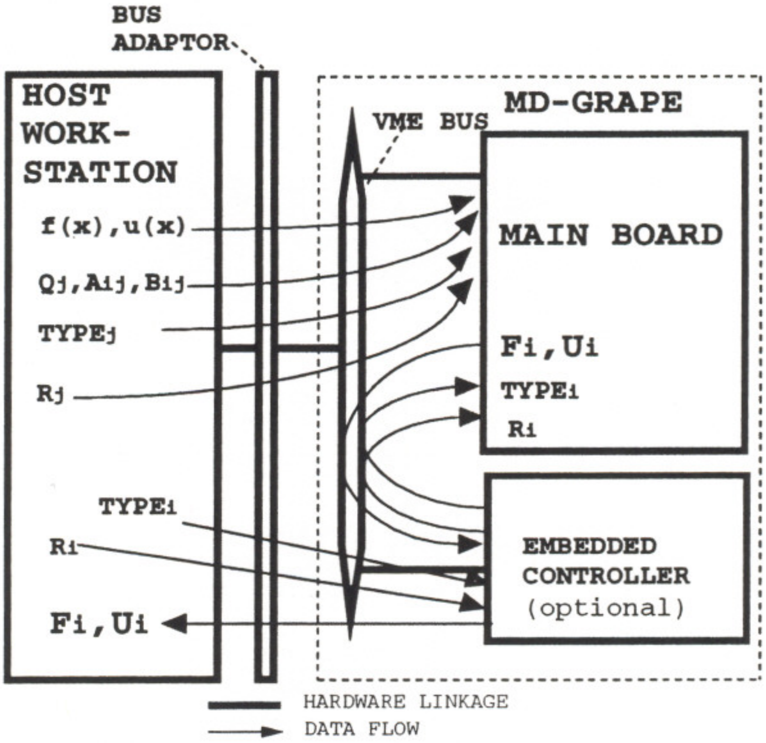

Figure 1: The hardware architecture of the MD-GRAPE system and the data flow between the machines. See Secs. 2.1 and 2.2 for details

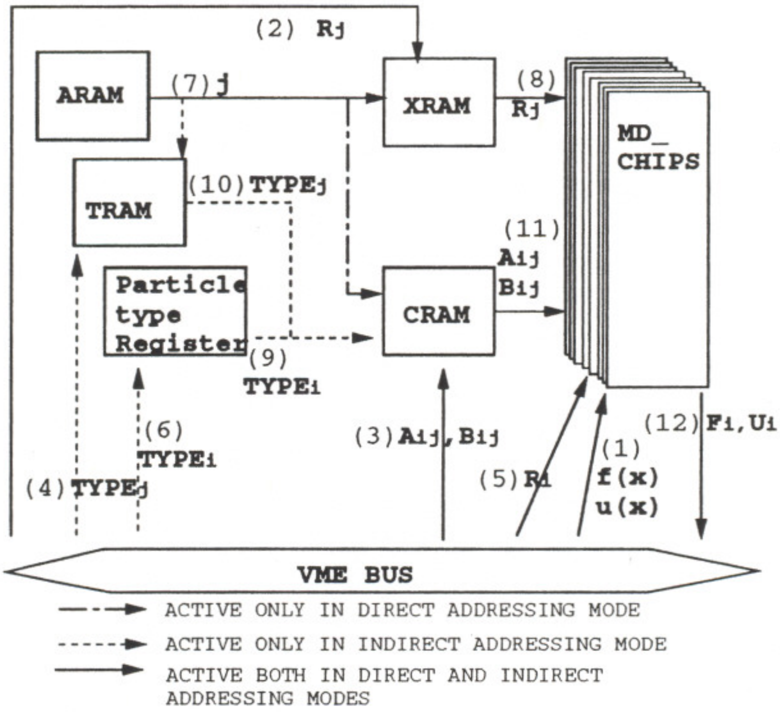

Figure 2: Data flow within MD-GRAPE. The numbers (1)~(12) roughly correspond to sequential order. See Secs. 2.1 and 2.2 for details.

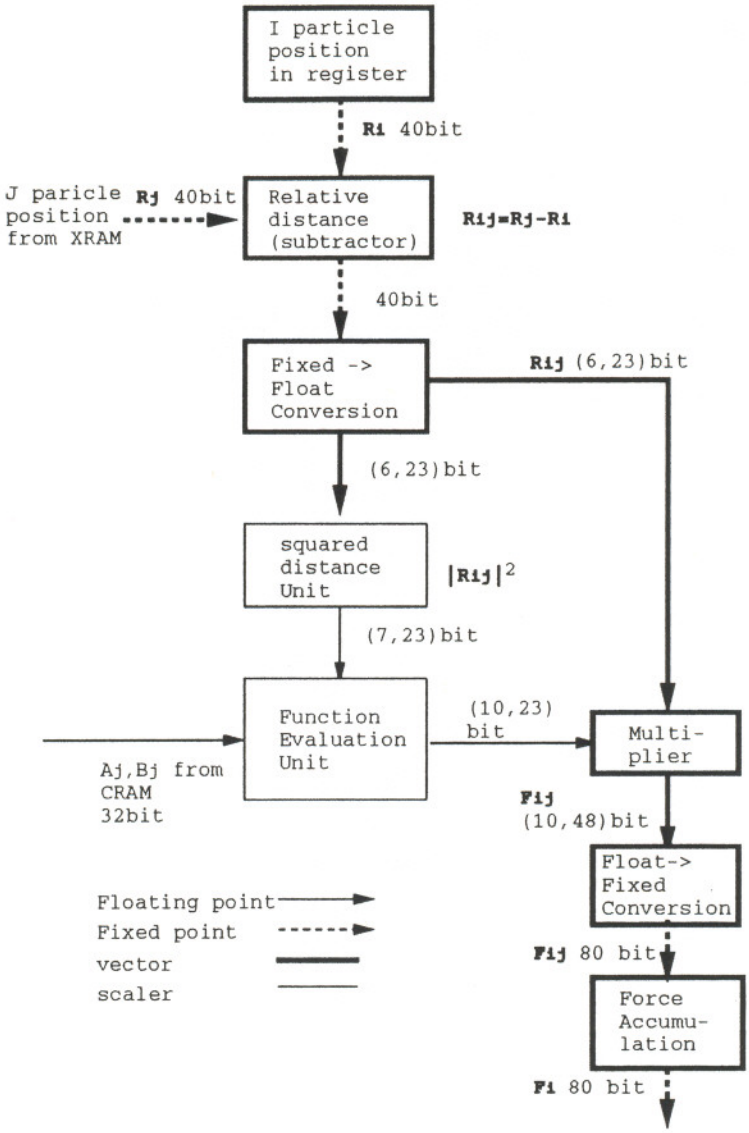

Figure 3: Data flow within MD chip. See Secs. 2.1 and 2.2 for details.

2 Methods

2.1 MD-GRAPE

The hardware architecture of the MD-GRAPE system is schematically presented in Figure 1. MD-GRAPE consists of the main board, embedded controller (optional, MVME162, Motorola), and VME bus (Versa Module Europe). MD-GRAPE is connected to the current host workstation (DEC 3000/700, Digital Equipment Corp.) through a bus adapter (DEC Turbo-VME bus adapter, Model 497, BIT3). MD-GRAPE has several Random Access Memories (RAM), which hold the coordinates of the particles (XRAM), force parameters (CRAM), and particle type indexes (TRAM) sent from the host (Fig. 2). The embedded controller sends data to or receives data from the board by random access, since this function is not achieved effectively by the current host and adapter.

The LSI (Large Scale Integrated circuit) used in this study (MD chip) will be described elsewhere [7]. Here we give only an outline. MD chip handles 6 particles simultaneously by using a virtual multiple pipeline method. Currently, eight MD chips are implemented in MD-GRAPE; hence, force or potential interacting on 6 x 8 (=48) particles are treated by parallel computation. The peak speed is 0.75 GFLOPS/chip at a clock of 25 MHz. MD-GRAPE is capable of handling up to 43,690 interacting particles at once. The data flow within MD chip is given in Figure 3. The precision of MD chip differs depending on the computational steps (Fig. 3), but most of the computation is performed in 32-bit single precision format. The final accumulation of force is performed with 80 bit fixed-point format (Fig. 3). In most MD's the precision is better than single precision [7].

The MD chip computes any kind of central force ($\mathbf{Fi}$) and potential (Ui) that has the following forms:

$$\mathbf{F}_{i} = \sum_{j} A_{ij} f(B_{ij} R_{ij}^{2}) \mathbf{R}_{ij} \qquad (2)$$

$$Ui = \sum_{j} C_{ij} u(D_{ij} R_{ij}^{2}) \qquad (3)$$

where $\mathbf{R_{ij}}$ is the distance displacement vector between particles i and j, R_{ij} is the absolute value of $\mathbf{R_{ij}}$, and A_{ij}, B_{ij}, C_{ij}, D_{ij} are coefficients derived from the force field parameters. Hereafter, particle i corresponds to the particle experiencing the force to be calculated, and j is the interacting particle. Functions f(x) and u(x) determine the kind of interaction. See Table 1 for Aij ~ Dij, f(x), and u(x) of Coulomb and VDW interactions. The tables of these functions are stored in the RAM of the MD chip (Figs. 1, 2 (1)). In the function evaluator unit (Fig. 3), the value of the function is evaluated using a segmented polynomial of degree four by table look-up [7]. The MD chip is also capable of computation of Ewald potential and force in a periodic boundary condition [7], though we do not report this capability here.

Table 1: Interactions computed by MD_GRAPE

	Coulomb	VDW
force(i,j)[a]	$Q_j/R_{ij}^3\mathbf{R_{ij}}$	$(12a_{ij}/R_{ij}^{14}-6b_{ij}/R_{ij}^{8})\mathbf{R_{ij}}$
f(x)	$x^{-3/2}$	$x^{-7}-x^{-4}$
A_{ij}	Q_j	$3(b_{ij}^7/2a_{ij}^4)^{1/3}$
B_{ij}	1	$(b_{ij}/2a_{ij})^{1/3}$
potential(i,j)[a]	Q_j/R_{ij}	$a_{ij}/R_{ij}^{12}-b_{ij}/R_{ij}^{6}$
u(x)	$x^{-1/2}$	$x^{-6}-x^{-3}$
C_{ij}	Q_j	b_{ij}^2/a_{ij}
D_{ij}	1	$(b_{ij}/a_{ij})^{1/3}$

[a] These should be multiplied by Q_i later.

2.2 Usage of MD-GRAPE

In this subsection, usage of the MD-GRAPE system is described along with the data flow within the hardware (Figs. 1~3). The application program uses GRAPE via subroutine call of the interface library. See reference [12] for a general description of the GRAPE interface.

Below we give a Fortran90-like program to calculate Coulomb force of N interacting particles using MD-GRAPE. Coulomb interaction is calculated via direct addressing mode. The names of the interface library subroutines are headed by M1. Comments are headed by !.

```
Program 1
integer:: N            ! the number of interacting particles
real:: R(3, N)         ! coordinates of the particles
                       ! (R(1, i), R(2, i), R(3, i)) = (x, y, z) of particle (i)
real:: F(3, N)         ! force acting on the particles
real:: Q(N)            ! charges of the particles
integer:: mode         ! mode of GRAPE
real:: Rmax, Fmax      ! described later

mode = 1                          ! Coulomb force mode
call M1presetmode(mode)     ! set mode of GRAPE
call M1setautoscale(Rmax, Fmax)
                                  ! set input & output decimal points
call M1setxjn(1, N, R(1,1))   ! send coordinates of particles 1~N
                                  ! to GRAPE
call M1setg0n(1, N, Q(1))     ! send charges of particles 1~N
                                  ! to GRAPE
call M1execn(R(1,1), N, N, 0, F(1,1))! calculate force
do i = 1, N
        F(:, i) = Q(i) * F(:, i)  ! multiply force by -Q(i)
end do
```

M1presetmode determines the mode of MD-GRAPE. Currently, Coulomb force mode is set to 1. When M1presetmode is called, the host sends the function table, f(x) or u(x), to the MD chips (Figs. 1 and 2 (1)). The input and output of the MD chip are not floating- but fixed-point decimals (Fig. 3), and it is necessary to set the decimal points by M1setautoscale. Rmax is the maximum value of |r(m,i)| and |R(m,i)-r(m,j)| (m = 1, 3; i = 1, N; j = 1, N), and Fmax is the presumable maximum of |F(m,i)| (m = 1, 3; i = 1, N). M1setxjn and M1setg0n send coordinates and charges of the interacting particles (j = 1, N in this case) to XRAM and CRAM of MD-GRAPE, respectively (Figs. 1 and 2 (2), (3)).

Now MD-GRAPE is ready to calculate force by M1execn. The second argument (N) of M1execn is the number of the computed particles and the third (N) is that of the interacting ones. In program 1 the particles to be computed and those interacting are the same, but they could be different. Self-interaction is avoided automatically by the MD chip. Upon calling M1execn the coordinates ($\mathbf{R}_i$) are temporarily sent to the embedded controller (Fig. 1), and the controller chooses 48 particles subject to parallel computation and sends their coordinates to the MD chips (Fig. 2 (5)). Then ARAM generates the counter j (j = 1, N; Fig 2 (7)); accordingly, XRAM and CRAM send to the MD chips the coordinate $\mathbf{R_j}$ and force parameters, A_{ij} and B_{ij}, respectively (Fig. 2 (8) (11)). The fourth argument (0) of M1execn is the particle type index, not used in the direct addressing mode. Finally, the force acting on particle i is calculated in MD chip (Fig. 3). At first $\mathbf{R_{ij}}$ ($= \mathbf{R_j} - \mathbf{R_i}$) is calculated and converted to a floating point decimal. Next, $|\mathbf{R_{ij}}|^2$ is obtained in the square distance unit and sent to the function evaluation unit, where $A_{ij}f(B_{ij}|\mathbf{R_{ij}}|^2)$ ($=Q_j/|\mathbf{R_{ij}}|^3$) is calculated by using A_{ij} ($=Q_j$) and B_{ij} ($=1.0$) sent from CRAM. Then the value is multiplied by $\mathbf{R_{ij}}$ (this procedure is omitted if potential is calculated). The force between particles i and j thus obtained, $\mathbf{F_{ij}}$, is converted to a fixed-point decimal and accumulated (Fig. 3). The above procedure (Fig. 2 (7)~(11)) is repeated until j = N, and the force acting on particle i, $\mathbf{F_i}$, is finally obtained when the summation over j is ended (Fig. 3). Then $\mathbf{F_i}$ is sent to the controller (Fig. 2 (12)) and eventually to the host (Fig. 1). The last argument of M1execn is the computed force. In Coulomb mode, the returned force should be multiplied by Q_i in the host (the last three lines of program 1).

VDW interaction is calculated via indirect addressing mode. Below is an example of the program for calculation of VDW force.

```
Program 2
integer:: ntype             ! number of particle types
integer:: point(ntype)  ! pointer to the first particle of the type
                            ! in the coordinate array (R)
real:: Aij(ntype, ntype), Bij(ntype, ntype)
                            ! force parameters of types
integer:: N                 ! the number of interacting particles
real:: R(3, N)              ! coordinates of the particles that have been
                            ! sorted according to their particle types
real:: F(3, N)              ! force acting on the particles
integer:: type(N)           ! particle type index of each particle
real:: Rmax, Fmax           ! see the description of program 1
integer:: mode              ! mode of GRAPE

mode = 3                            ! VDW force mode
call M1presetmode(mode)     ! set GRAPE to VDW force mode (=3)
call M1setautoscale(Rmax, Fmax)
                                    ! set input & output decimal points
do k = 1, ntype
        call M1setgrscalen(k, 1, ntype, Aij(1, k), Bij(1, k))
        ! send Aij & Bij to GRAPE
end do
call M1setxjn(1, N, R(1,1))  ! send coordinates of particles 1~N
                                    ! to GRAPE
call M1setcjn(1, N, type(1)))
                                    ! send particle types of particles 1~N
                                    ! to GRAPE
do k = 1, ntype                     ! loop over particle types
        i1 = point(k)   ! the first particle of type(k)
        i2 = point(k+1)-1 ! the last particle of type(k)
        m = i2-i1+1    ! number of particles whose type is k
        call M1execn(R(1, i1), m, N, k, F(1, i1))    ! calculate force
end do
```

In the indirect addressing mode exemplified above, the particles should have been sorted according to their particle types (O, C, H, S, P and so on), because the particles computed by the MD chips simultaneously should have the same force coefficients, Aij and Bij, or Cij and Dij (Eqs. 2 and 3). In program 2, it is postulated that there are ntype types of particles and the coordinates R were already sorted according their types. The coordinates of the particles whose type is k are stored in R(:, point(k)) ~ R(:, point(k+1)-1). In calculation of VDW force, M1presetmode sets the mode of MD-GRAPE to 3. Read the description of program 1 for usage of M1setautoscale. M1setgrscalen sends the coefficients defined by Eqs. 2 and 3 to CRAM of MD-GRAPE (Figs. 1 and 2 (3)). Coordinates of the interacting particles are sent to XRAM by M1setxjn as in example 1, and particle type indexes (type_j) are sent to TRAM of MD-GRAPE (Fig. 2 (4)). In the indirect addressing mode, M1execn is called once in each particle type. The arguments are the same as in the Coulomb mode described above, but the particle type index k (= type_i) of the computed particles is used in this mode; namely, type_i is sent to the particle type register (Fig. 2 (6)). The counter j generated by ARAM is not sent to CRAM directly but sent to TRAM to find type_j (Fig. 2 (7)). Then A_{ij} and Bij are chosen by using typei and type_j (Fig. 2 (9, 10)) and sent to MD chips (Fig. 2 (11)). The force acting on the particles is calculated similarly to the direct addressing mode, then returned.

We have written of a set of fortran subroutines (emulator) that work in the same way as MD-GRAPE except for the speed and precision, and used them to compare the performance and precision of MD-GRAPE to general-purpose computers.

2.3 Test of performance and precision of MD-GRAPE

The performance and precision of MD-GRAPE was tested by MD of a dimensionless N-particle system interacting via either Coulomb or VDW interaction. Suppose the particles are interacting under potential U

$$U = \sum_{i>} \sum_{j} \varepsilon \phi\left(\frac{\sigma}{R_{ij}}\right) \qquad (4)$$

where ε and σ are constants whose dimensions are energy and length, respectively, and R indicates the distance. If we make the conversions $R = r\sigma$ and $\tau = t\gamma$, Newton's equation of motion for particle i

$$m\frac{d^2\mathbf{R}_i}{dt^2} = -\nabla_i U \qquad (5)$$

is converted to a dimensionless equation as follows,

$$\frac{d^2\mathbf{r}_i}{d\tau^2} = \sum_{j} \phi'\left(\frac{1}{\mathbf{r}_{ij}}\right)\frac{\mathbf{r}_{ij}}{\mathbf{r}_{ij}{}^3} \qquad (6)$$

where m is the mass of each particle and γ is defined as below.

$$\gamma = \sigma\sqrt{\frac{m}{\varepsilon}} \qquad (7)$$

In MD of Eq. 6, energy is reported in ε unit.

In simulation of identical particles interacting via VDW interaction, the potential is transformed as follows.

$$U = \sum_{i>}\sum_{j}\varepsilon\left[(\frac{\sigma}{R_{ij}})^{12} - (\frac{\sigma}{R_{ij}})^{6}\right] \qquad (8)$$

This system is made dimensionless if we define $\phi(x)=x^{12}-x^{6}$.

Next, consider a system consisting of particles each of which has a charge of either +Q or -Q. The potential function of this system is

$$U = \sum_{i>}\sum_{j} Q^2 \frac{\delta_{ij}}{R_{ij}} \qquad (9)$$

where δ_{ij} =-1 if the charges of the particles i and j differ and 1 if not. The equation of motion under such a potential is also made dimensionless by replacing Q^2 with ε. In this case $\sigma = 1$ and $\phi(x)=x$.

Then particles were generated so as to form a cubic lattice, the distance between the nearest neighbors being 1.5σ for VDW particles and 5σ for Coulomb particles, and 1088 particles were scooped to form a sphere. Each system was equilibrated for 1 γ at a time step of 0.001 γ by Cray C90. The tests were started from the configuration thus obtained. MD simulations were performed both by Cray (64-bit precision) and MD-GRAPE connected to the host for a period of 5 γ with various time-steps in a free boundary condition. A simple fortran program was used with the velocity-Verlet integration scheme [15], which was also adapted to PEACH.

Root Mean Square Deviation ($RMSD_{init}$) from the initial energy (E_{init}) was used to check the fluctuation in energy. $RMSD_{init}$ is defined below:

$$RMSD_{init} = \sqrt{\frac{1}{N_{step}}\sum_{i=1}^{Nstep}(E_i - E_{init})^2} \qquad (10)$$

where E_i is the energy at the ith step and N_{step} is the number of time steps.

2.4 MD Software PEACH

PEACH is a newly developed software package for MD of biomolecules. The overall structure of PEACH follows conventional MD software packages such as AMBER [16]; namely, the software is divided into several modules and the data are transmitted among the modules via files (Fig. 4). The module MKDBAS makes a database of amino acid residues, nucleosides, and any kind of molecules the user defines. MKMOL constructs the topology of the molecule, and MKCOR makes the coordinates, adds solvents, and so on. MKPARA assigns the force field parameters. Currently, PEACH uses the residue types and force field parameters of AMBER [13, 17, 18] or AMBER/OPLS [14]. RUNMD performs

MD by reading the initial configuration from the MKCOR coordinate file and the force field parameters from the MKPARA topology file. The integration scheme is the velocity-Verlet algorithm [17]. In RUNMD, Coulomb and VDW (and H bond if AMBER force field) interactions are calculated by MD-GRAPE, but they could be calculated by the workstation too if the emulator is linked instead of the interface library. PEACH is written in Fortran90, which offers many new features such as dynamic memory allocation and elegant treatment of arrays.

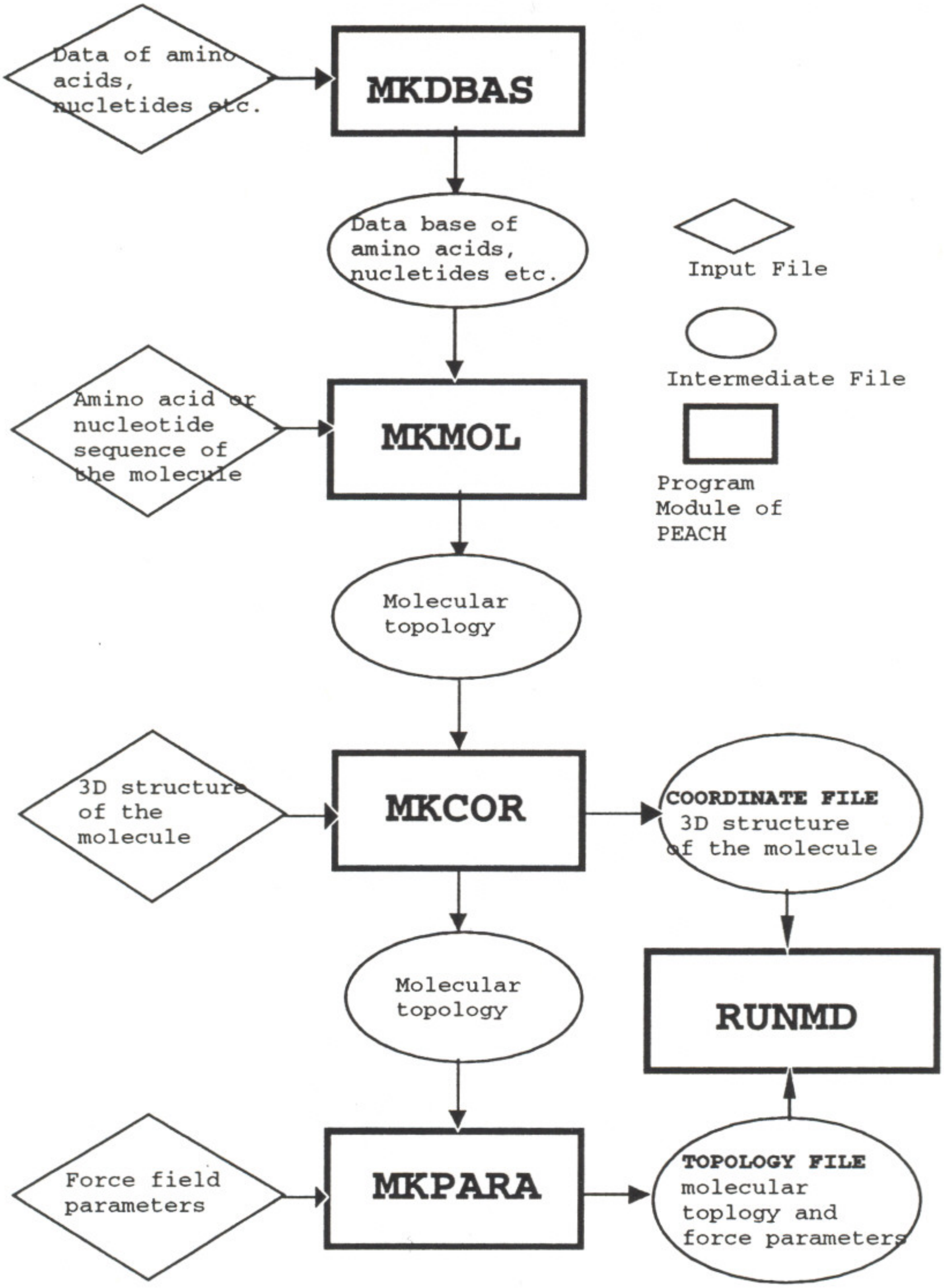
Data of amino acids, nucletides etc.
MKDBAS
Data base of amino acids, nucletides etc.
Input File
Intermediate File
Program Module of PEACH
Amino acid or nucleotide sequence of the molecule
MKMOL
Molecular topology
3D structure of the molecule
MKCOR
COORDINATE FILE 3D structure of the molecule
Molecular topology
RUNMD
Force field parameters
MKPARA
TOPOLOGY FILE molecular toplogy and force parameters

Figure 4: Schematics of the structure of PEACH. See Sec. 2.4 for description.

2.5 Test of the PEACH-GRAPE System

To examine the performance of the PEACH-GRAPE system, we performed MD of three different protein-solvent systems. Systems A and B used Histidine-containing phosphocarrier protein (HPR) of *Streptococcus faecalis* [19], a small protein with 88 amino acid residues. In system A, a shell water (8 Å thick) was generated around HPR. In system B, a CAP water (30.3 Å radius) was generated around the center of mass of the protein. System C consists of Calmodulin - M13 peptide complex (CaM/M13) [20] and a surrounding Cap water (42.1 Å radius). N~ 3,200 for system A, ~10,000 for system B, and ~27,000 for system C (Table 2). Computation time per each time step was evaluated. The AMBER/OPLS parameter set was used with TIP3P water [21] for the solvent.

System A was also used for examination of the precision of the PEACH-GRAPE system. MD was performed in a free boundary condition. For the first 3 ps the system was heated up to 300 K by velocity scaling, followed by 2 ps MD at 300 K. The software emulator was used during the heating and equilibration periods. Then 5 ps classical MD was performed both by the software emulator and by MD-GRAPE. $\mathrm{RMSD_{init}}$ (Eq. 10) of kinetic, potential, and total energies were calculated.

3 Results and Discussion

3.1 MD of Dimensionless N-body Systems

3.1.1 Performance

We compared the computation time/step of electrostatic interactions of N-body systems with various number of particles. The computation time was then normalized per interaction; namely, the time was divided by N^2. The performance was compared for two GRAPEs, MD-GRAPE and GRAPE2A (the prototype machine of MD-GRAPE) [11], and for two general-purpose computers, DEC3000/700 (the host of MD-GRAPE) and CRAY C90 supercomputer. The peak speeds of the machines are: 6 (MD_GRAPE), 0.18 (GRAPE2A), 0.23 (DEC), and 0.9 Gflops (CRAY, 1 CPU). The emulator software of MD-GRAPE was used to test the general purpose computers.

The results of the comparison are shown in Figure 5. When N>10,000, MD-GRAPE: GRAPE2A: DEC: CRAY ≅ 1: 30: 160: 13. At this stage, Cray C90 had an effective speed of ~0.5 Gflops. Hence, the effective speed of MD-GRAPE is equivalent to ~ 6 Gflops in comparison with Cray. Though the value of flops of GRAPE does not directly correspond to that of the general purpose computers, MD-GRAPE is suggested to have attained its peak speed. MD-GRAPE had ~30 times faster performance than GRAPE2A, the preceeding machine, indicating the power of LSI. The performance of MD-GRAPE decreased when N < 10,000, mainly due to the overhead of communication with the host. Most of MD studies of solvated biolmolecules require more than several thousands of particles, however; hence, MD-GRAPE should be effective in MD of such molecules. Also, use of another host with a better random access capability could improve the performance when N <10,000.

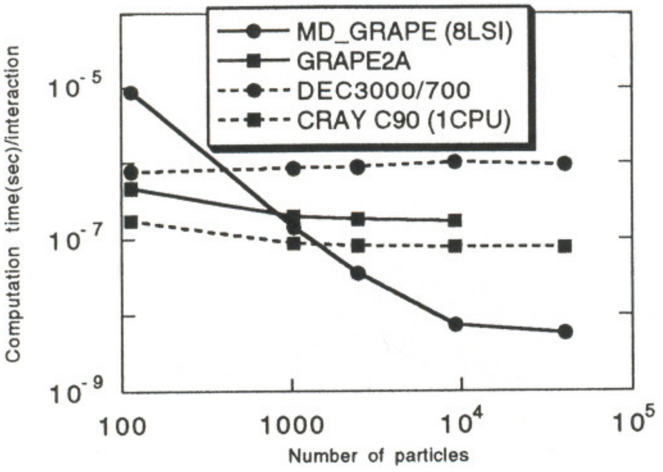

Figure 5: Result of benchmark test of various machines. Computation time/step/(number of particles)^2 of Coulomb interaction is shown for several machines in MD of various numbers of particles.

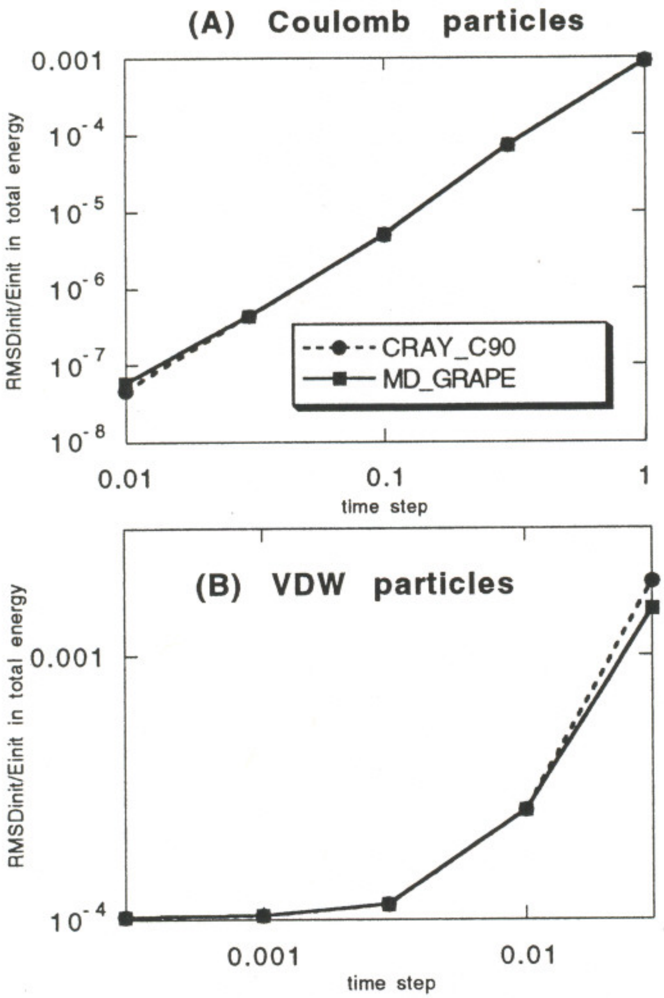

Figure 6: RMSD$_{\mathrm{init}}$ of total energy in MD of a dimensionless system (N = 1088). See Sec. 2.3 for a description.

3.1.2 Precision

In a Verlet integration method, $\mathrm{RMSD_{init}}$ (Eq. 10) of the total energy should be proportional to (Δt; time step)^2; hence, the double logarithmic plot of $\mathrm{RMSD_{init}}$ and Δt should be a straight line with a slope of 2 unless the round-off error is dominant [22]. By using the plot (Fig. 6), we compared the precision of MD-GRAPE against Cray C90 (64 bit precision, using the software emulator).

The plot is almost straight in MD of the particles interacting via Coulomb interaction (Fig. 6 A). MD-GRAPE gave nearly the same precision as Cray C90 when $\mathrm{RMSD_{init}/E_{init}}$ is 10^{-4}~10^{-5}. In the VDW system, the plot is not straight, however. This is due to the fact that the functions f(x) and u(x) for VDW interaction (Table 1) are only defined for the range of [1/4, 64), and the emulator mimics this behavior. Unlike Coulomb interaction, the cut-off scheme may be applied without serious effects on the dynamics of the system, and the error revealed by Fig. 6 (B) should be negligible.

Thus, in MD of the systems we considered here, MD_GRAPE gave results comparable to that obtained under 64-bit precision. The precision of MD-GRAPE when applied to a protein is presented in the next subsection.

3.2 MD of Solvated Proteins (test of the PEACH-GRAPE system)

3.2.1 Performance

We performed MD of three protein-solvent systems with different sizes to examine the performance of the PEACH-GRAPE system. The results are presented in Table 2.

The performance for each system was 0.8 (system A), 1.6 (system B), and 5.5 s/step (system C), and acceleration of the host by MD-GRAPE was 24, 91, and 213 fold, respectively. These values are nearly equivalent to those computed by current supercomputers using the cut-off approximation; for instance, 0.6 s/step by CHARMM on a CRAY C90 for solvated helical peptides (4000 atoms, ~10 Å cut-off) [23], 2.5 s/step by CHARMM on a CRAY YMP-3 for a solvated anti-freeze peptide (4500 atoms, ~12 Å cut-off) [24], 3.5 s/step by AMBER on a Cray C90 for solvated *trp*-repressor (19,000 atoms, 7~18 Å twin-range cut-off, Komeiji unpublished), and so on. Thus, by using the PEACH-GRAPE system, 100 ps MD of systems A~C would take only 2, 4, and 13 days (when $\Delta t = 0.5$ fs) without the nonbonded cut-off scheme.

Naturally, acceleration by MD-GRAPE increased according to the system size N, because the ratio of nonbonded interactions within the computation time increases in proportion to N, and also because the performance of MD-GRAPE itself is better in larger systems while that of the host is almost invariant (Fig. 5).

Coulomb interaction consumed more computation time than VDW interaction in systems B and C (Table 2). This fact is attributable to the use of the OPLS force field with TIP3P water. In this set of force fields, all of the hydrogen atoms have no VDW radius or well depth, and PEACH avoids calculation of VDW interaction of such atoms to save computation time. In system A, the computation time of VDW was comparable to Coulomb (Table 2), presumably because the overhead of communication with the host by calling M1execn (Program 2) so often though the number of atoms was small.

Table 2: Performance of the PEACH-GRAPE system

Simulated system [a]	System A (HPR in Shell)		System B (HPR in Cap)		System C (CaM/M13 in Cap)	
Number of atoms						
protein	793		793		1685	
ion	0		0		4	
water	2448		9111		25464	
Total	3241		9904		27153	
Machine [b]	WS	GR	WS	GR	WS	GR
Computation time/step [c]						
VDW	7.83	0.32	46.39	0.56	301.2	1.59
	(7.79)	(0.28)	(46.29)	(0.48)	(300.9)	(1.35)
Coulomb	9.98	0.30	99.56	0.71	863.3	2.97
	(9.95)	(0.27)	(99.49)	(0.64)	(863.0)	(2.76)
others	0.12	0.13	0.33	0.34	0.94	1.36
Total	17.93	0.75	146.3	1.61	1165.4	5.47
	(17.73)	(0.55)	(145.8)	(1.12)	(1163.9)	(4.11)
Acceleration by MD-GRAPE						
	24		91		213	

[a] See Sec. 2.5 for description of the simulated systems.

[b] WS means that all the calculations were performed by the workstation (DEC3000/700) using the software emulator, while GR means that GRAPE computed Coulomb and VDW interactions and the workstation computed other interactions.

[c] Computation time per each time step of MD is shown in second. The values within parentheses indicate the calculations performed by MD-GRAPE (GR) or by the software emulator (WS). Averages of 10 steps are shown.

[d] Ratio between total computation time/step by MD-GRAPE and that by the emulator.

Table 3: Energy flucuations during MD by the PEACH-GRAPE system[a].

	E_{init}	$RMSD_{init}$	
		Δt=0.5 fs	1.0 fs
Potential			
WS [b]	-1.023×10^4	39.35	38.26
GR [b]	"	36.33	38.90
Kinetic			
WS	2.896×10^3	39.35	38.15
GR	"	36.33	39.07
Total			
WS	-7.330×10^3	0.117	0.462
GR	"	0.125	0.446

[a] System A (Table 2) was simulated and the energies are reported in kcal/mol.

[b] See legend b to Table 2 for definitions of WS and GR.

3.2.2 Precision

We compared $\mathrm{RMSD_{init}}$ (Eq. 10) of kinetic, potential, and total energies obtained by classical MD of system A between MD-GRAPE and the host (64-bit). The $\mathrm{RMSD_{init}}$ agreed between the two machines within acceptable errors at $\Delta t = 0.5$ fs and 1.0 fs (Table 3). Thus, the PEACH-GRAPE system seems precise enough at least in application to MD of solvated proteins under normal conditions.

3.3 Conclusion and Perspective

In this article, we reported the construction of a MD system for biomolecules. MD of biomolecules without the truncation of the nonbonded interactions has become possible within reasonable computation time by combination of a special-purpose computer, MD-GRAPE, and a newly developed software package, PEACH. We are planning to apply this system to analyses of dynamics, stability, domain communication, and molecular interactions of proteins and nucleotides.

The potential of special purpose computers to N-body simulations is enormous. A Tflops machine for cosmological problems has been recently constructed (Makino, personal communication). Similarly, a 100 Gflops~1 Tflops MD machine can be developed in the near future by massive parallelization of the MD chips if the communication between the host and MD-GRAPE is made faster by using another bus bridge. Advent of such a high-performance machine will enable us to analyze behavior of biomolecules within longer time domains such as allosteric transition and folding.

MD-GRAPE described in this report is commercially available as ITL-MD-ONE from ITL Corp. (1-8-6 Higashitsutsujigaoka, Chofu-shi, Tokyo 182, Japan, Tel +81 3 3308 2431, Fax +81 3 3308 4406). The cost of ITL-MD-ONE (8 MD chips) is ¥7,800,000 (~$78,000).

Acknowledgments

The Research Processing Station at the Agency of the Industrial Science and Technology of Japan is acknowledged for providing time on the Cray C90 computer. This work was performed through the special coordination funds of the Science and Technology Agency of the Japanese Government.

References

1. M. Karplus and G. A. Petsko, *Nature* 347, 631 (1990).
2. J. Guenot and P. A. Kollman, *Prot. Sci.* 1, 1185 (1992).
3. H. Schreiber and O. Steinhauser, *Biochemistry* 31, 5856 (1992).
4. M. Saito, *Mol. Simul.* 8, 321 (1992).
5. M. Saito, *J. Chem. Phys.* 101, 4055 (1994).
6. F. Zhau and K. Schulten, *J. Phys. Chem.* 99, 2194 (1995).
7. M. Taiji *et al.,* in preparation.
8. D. Sugimoto *et al, Nature* 345, 33 (1990).
9. T. Ebisuzaki *et al.* , *Publ. Astron. Soc. Jpn.* 45, 269 (1993).
10. Y. Ohno *et al.* , *Publ. Astron. Soc. Jpn.* 45, 377 (1993).
11. T. Ito *et al.*, *Proteins* 20, 139 (1994).
12. J. Makino and Y. Funato, *Publ. Astron. Soc. Jpn.* 45, 279 (1993).

3.2.2 Precision

We compared $\mathrm{RMSD_{init}}$ (Eq. 10) of kinetic, potential, and total energies obtained by classical MD of system A between MD-GRAPE and the host (64-bit). The $\mathrm{RMSD_{init}}$ agreed between the two machines within acceptable errors at $\Delta t = 0.5$ fs and 1.0 fs (Table 3). Thus, the PEACH-GRAPE system seems precise enough at least in application to MD of solvated proteins under normal conditions.

3.3 Conclusion and Perspective

In this article, we reported the construction of a MD system for biomolecules. MD of biomolecules without the truncation of the nonbonded interactions has become possible within reasonable computation time by combination of a special-purpose computer, MD-GRAPE, and a newly developed software package, PEACH. We are planning to apply this system to analyses of dynamics, stability, domain communication, and molecular interactions of proteins and nucleotides.

The potential of special purpose computers to N-body simulations is enormous. A Tflops machine for cosmological problems has been recently constructed (Makino, personal communication). Similarly, a 100 Gflops~1 Tflops MD machine can be developed in the near future by massive parallelization of the MD chips if the communication between the host and MD-GRAPE is made faster by using another bus bridge. Advent of such a high-performance machine will enable us to analyze behavior of biomolecules within longer time domains such as allosteric transition and folding.

MD-GRAPE described in this report is commercially available as ITL-MD-ONE from ITL Corp. (1-8-6 Higashitsutsujigaoka, Chofu-shi, Tokyo 182, Japan, Tel +81 3 3308 2431, Fax +81 3 3308 4406). The cost of ITL-MD-ONE (8 MD chips) is ¥7,800,000 (~$78,000).

Acknowledgments

The Research Processing Station at the Agency of the Industrial Science and Technology of Japan is acknowledged for providing time on the Cray C90 computer. This work was performed through the special coordination funds of the Science and Technology Agency of the Japanese Government.

References

1. M. Karplus and G. A. Petsko, *Nature* 347, 631 (1990).
2. J. Guenot and P. A. Kollman, *Prot. Sci.* 1, 1185 (1992).
3. H. Schreiber and O. Steinhauser, *Biochemistry* 31, 5856 (1992).
4. M. Saito, *Mol. Simul.* 8, 321 (1992).
5. M. Saito, *J. Chem. Phys.* 101, 4055 (1994).
6. F. Zhau and K. Schulten, *J. Phys. Chem.* 99, 2194 (1995).
7. M. Taiji *et al.,* in preparation.
8. D. Sugimoto *et al*, *Nature* 345, 33 (1990).
9. T. Ebisuzaki *et al.* , *Publ. Astron. Soc. Jpn.* 45, 269 (1993).
10. Y. Ohno *et al.* , *Publ. Astron. Soc. Jpn.* 45, 377 (1993).
11. T. Ito *et al.*, *Proteins* 20, 139 (1994).
12. J. Makino and Y. Funato, *Publ. Astron. Soc. Jpn.* 45, 279 (1993).

13. S. J. Weiner *et al.*, *J. Am. Chem. Soc.* 106, 765 (1984).
14. W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.* 110, 1657 (1988).
15. W. C. Swope *et al.*, *J. Chem. Phys.* 76, 637 (1982).
16. P. K. Weiner and P. A. Kollman, *J. Comput. Chem.* 2, 287 (1981).
17. S. J. Weiner *et al.*, *J. Comput. Chem.* 7, 230 (1986).
18. W. D. Cornell *et al.*, *J. Am. Chem. Soc.* 117, 5179 (1995).
19. A. Jia *et al.*, *Nature* 361, 94 (1993).
20. M. Ikura *et al.*, *Science* 256, 632 (1992).
21. W. L. Jorgensen *et al.*, *J. Chem. Phys.* 79, 926 (1983).
22. J. M. Haile, *Molecular Dynamics Simulation-Elementary methods* (Wiley & Sons, N. Y., 1992)
23. J. D. Hirst and C. L. Brooks III, *Biochemistry* 34, 7614 (1995).
24. Y. Yang *et al.*, *J. Am. Chem. Soc.* 115, 7439 (1993).