John von Neumann Institute for Computing

RedMD – A New Package for Reduced Molecular Dynamics

M. Szypowski, A. Górecki, J. Trylska

published in

From Computational Biophysics to Systems Biology (CBSB08),
Proceedings of the NIC Workshop 2008,
Ulrich H. E. Hansmann, Jan H. Meinke, Sandipan Mohanty,
Walter Nadler, Olav Zimmermann (Editors),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **40**, ISBN 978-3-9810843-6-8, pp. 389-392, 2008.

© 2008 by John von Neumann Institute for Computing
Permission to make digital or hard copies of portions of this work for personal or classroom use is granted provided that the copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise requires prior specific permission by the publisher mentioned above.

http://www.fz-juelich.de/nic-series/volume40

RedMD – A New Package for Reduced Molecular Dynamics

Marcin Szypowski[1]**, Adam Górecki**[1,2]**, and Joanna Trylska**[1]

[1] Interdisciplinary Centre for Mathematical and Computational Modelling,
University of Warsaw, Warsaw, Poland
E-mail: {M.Szypowski, joanna}@icm.edu.pl

[2] Department of Biophysics, Faculty of Physics, University of Warsaw, Warsaw, Poland
E-mail: A.Gorecki@icm.edu.pl

We developed a RedMD package to perform molecular dynamics simulations for coarse-grained models of proteins, nucleic acids and its complexes. Simulations can be carried out in the microcanonical ensemble, as well as with Berendsen and Langevin thermostats. We provide tools to generate initial configuration and topology which are based on the elastic network approach and its extensions. Topology generators can be modified by users to add for example a new potential type. The code is written in C/C++ languages and the structure/topology of a molecule is based on an XML format. The code is distributed under GNU public licence and will be available at `http://bionano.icm.edu.pl/`.

1 Introduction

We created an open-source, scalable package for reduced (coarse-grained) molecular dynamics (MD) simulations of biomolecules[1] on a micro- to mili-second time scales. It is written in C/C++ and parallelized with an OpenMP technology. To generate the topology and force field of a molecule we use an XML-based format. Currently implemented force field generators include the Elastic Network Model[2] and its anharmonic extensions for the ribosome[3], nucleosome[4] and HIV-1 protease[5].

2 Molecular Dynamics

MD is a widely used technique to investigate the dynamical properties of molecules. It numerically solves in finite time steps the Newton's equations of motion and provides trajectories i.e., the coordinates and momenta of particles as a function of time. Our MD package generates coarse-grained representations of biopolymers in which entire groups of atoms are represented by single interacting centers (pseudo-atoms), (see Figure 1). In one bead models, one pseudo-atom can represent the whole amino acid or nucleotide. The coarse-graining procedure is applied to reduce the number of degrees of freedom, increase the integration time step, and achieve at least a microsecond MD simulation time scale. RedMD currently supports MD simulations in the microcanonical ensemble, in the canonical ensemble with the Berendsen thermostat, and with the Langevin bath thermal coupling. To enlarge the integration time step we implemented the SHAKE algorithm[6,7] which was developed to satisfy the bond geometry constraints in MD simulations.

Figure 1. The coarse-graining procedure. Left: all-atom representation, right: reduced one-bead model.

2.1 Microcanonical Ensemble

In the microcanonical ensemble (NVE) we consider the Newton's equation:

$$\frac{d\vec{p}_i}{dt} = -\frac{\partial V}{\partial \vec{r}_i} \qquad , \; i = 1, \ldots, N$$

where V is a potential energy function which depends on nucleic coordinates r_i, p_i is the momentum of a particle i and N is the number of particles in the system. Numerical solution of this equation generates the trajectory of motion. We implemented two popular algorithms: velocity Verlet and Leap-Frog[8] which are based on Taylor expansion.

2.2 Thermostats

To maintain constant temperature RedMD provides various thermostats. A common method of both thermal coupling and reproducing the contact with solvent molecules is Langevin Dynamics. We consider the equation:

$$\frac{d\vec{p}_i}{dt} = -\frac{\partial V}{\partial \vec{r}_i} - \gamma_i \vec{p}_i + \vec{R}(t) \qquad , \; i = 1, \ldots, N$$

where γ_i is the collision parameter and $\vec{R}(t)$ is a random force vector satisfying:

$$< \vec{R}(t) >= 0, \;\; < \vec{R}(t)\vec{R}(t') >= 2\gamma_i k_B T m_i \delta(t - t')$$

where k_B is the Boltzmann constant and T is the bath temperature. We solve this equation with Brünger-Brooks-Karplus integrator (BBK)[9]. Through the Langevin equation the system couples to a heat bath globally, but is also locally subjected to random noise. If we are interested in imposing the global coupling with minimal local disturbance the Langevin equation is modified to:

$$\frac{d\vec{p}_i}{dt} = -\frac{\partial V}{\partial \vec{r}_i} + m_i \gamma \left(\frac{T_0}{T} - 1 \right) \vec{p}_i \qquad , \; i = 1, \ldots, N$$

where T_0 is the bath temperature, T the current temperature and γ (in ps^{-1}) determines the strength of coupling with the thermal bath. This is known as the Berendsen thermostat[10]. The solution is based on scaling the momenta in each step by a factor

$$\lambda = \sqrt{1 + \frac{\Delta t}{\tau_T} \left(\frac{T_0}{T} - 1 \right)}$$

where τ_T is equal to $(2\gamma)^{-1}$.

2.3 Brownian Dynamics (BD)

RedMD provides a BD simulation based on the Ermak-McCammon algorithm[11]. The trajectory is generated according to the equation:

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \frac{D}{k_b T}\vec{F}_i(t)\Delta t + \vec{R}_i(t) \qquad , \ i = 1, \ldots, N$$

where D is a diffusion coefficient of a molecule and $\vec{R}(t)$ is a random force vector satisfying:

$$< \vec{R}(t) >= 0, \ \ < \vec{R}(t)\vec{R}(t') >= 6D\Delta t\delta(t - t').$$

3 Parallelization

Because the calculation of the non-bonded interactions and forces is the most time consuming step in MD, its optimization is very important. To parallelize this calculation we applied the OpenMP technology which uses a shared memory architecture. For the nucleosome model test force calculations for 1000 NVE simulation steps show that the speedup (defined as the ratio of time using one thread to the time of using N threads) for 7 cores is over 6 so it is almost linear.

4 The XML Force Field Format

The initial configuration and force field of the molecule are generated in an XML-based format. We provide utilities to produce the input XML file from the PDB or PDBML/XML files (`http://www.rcsb.org/pdb/`). Our XML format is a flexible way of representing the topology and force field. It contains information on atomic ids, masses, names, coordinates, and properties such as bonded and non-bonded potentials. It is possible to save any calculated values like temperature factors, forces or energies (see example below).

```xml
<?xml version="1.0"?>
<STRUCTURE>
<!-- Generated with: ./RedMD_genModel_Rib 1A36.txml -->
  <NONBONDED>
    <NBMORSE cutoff="35.000000">
      <PAIR type1="CA" type2="CA" alpha="0.707000" E0="0.055055" l0="9.500000"/>
      <PAIR type1="P" type2="P" alpha="0.707000" E0="0.071348" l0="17.600000"/>
      <PAIR type1="CA" type2="P" alpha="0.707000" E0="0.062674" l0="12.930584"/>
    </NBMORSE>
  </NONBONDED>
  <MOLECULE molId="mol1">
   <GROUP label="basePairs" k="0.600000"/>
   <ATOM id="1" name="P" x="-50.7" y="76.7" z="327.1" resName="U" chainID="A" resSeq="2" m="305"/>
   <ATOM id="2" name="P" x="-50.8" y="73.7" z="332.1" resName="U" chainID="A" resSeq="3" m="305"/>
   <ATOM id="3" name="P" x="-48.5" y="75.0" z="336.9" resName="G" chainID="A" resSeq="4" m="344"/>
   <BOND idAtom1="1" idAtom2="2" k="3.0" l0="5.80"/>
   <BOND idAtom1="1" idAtom2="3" k="2.5" l0="10.15"/>
   <BOND idAtom1="1" idAtom2="4" k="0.5" l0="9.40"/>
   <MORSE idAtom1="1" idAtom2="5" alpha="0.707" l0="6.19" E0="0.71" mark="o:P:P"/>
   <MORSE idAtom1="1" idAtom2="6" alpha="0.707" l0="6.87" E0="0.63" mark="o:P:P"/>
   <MORSE idAtom1="1" idAtom2="7" alpha="0.707" l0="11.90" E0="0.27" mark="o:P:P"/>
 </MOLECULE>
</STRUCTURE>
```

5 From PDB to Trajectory

First, one needs to convert a PDB or PDBML file to our XML file format with one of the programs distributed in the RedMD package. The user can choose from a few coarse-grained force fields. Second, the user needs to specify MD simulation parameters and then can generate the trajectory. Currently supported trajectory output formats are XYZ, PDB, DCD and VEL (analogous to XYZ but saves velocities).

Acknowledgments

We acknowledge support from University of Warsaw (115/30/E-343/S/2007/ICM BST1255 and 115/30/E-343/BST1345/ICM2008), Polish Ministry of Science and Higher Education (3 T11F 005 30), Fogarty International Center (NIH Research Grant R03 TW07318) and Foundation for Polish Science. Simulations were performed at ICM University of Warsaw grant no G31-4.

References

1. A. Gorecki, M. Szypowski, B. Domagala, and J. Trylska, *RedMD - a coarse-grained molecular dynamics package*, in preparation, 2008.
2. M. Tirion, *Large amplitude elastic motions in proteins from a single-parameter, atomic analysis*, Phys. Rev. Lett., **77**, 1905–1908, 1996.
3. J. Trylska, V. Tozzini, and J. A. McCammon, *Exploring Global Motions and Correlations in the Ribosome*, Biophys. J., **89**, 1455–1463, 2005.
4. K. Voltz, J. Trylska, V. Tozzini, V. Kurkal-Siebert, J. Langovski, and J. Smith, *Coarse-grained Force Field for the Nucleosome from Self-consistent Multiscaling*, J. Comp. Chem., **in press**, 2008.
5. V. Tozzini, J. Trylska, C. E. Chang, and J. A. McCammon, *Flap Opening Dynamics in HIV-1 Protease Explored with a Coarse-Grained Model*, J. Struct. Biol., **157**, 606–615, 2007.
6. Jean-Paul Ryckaert, Giovanni Ciccotti, , and Herman J.C. Berendsen, *Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkenes*, Journal of Computational Physics, **23**, 327–341, 1977.
7. Timothy R. Forester and William Smith, *SHAKE, Rattle, and Roll: Effcient Constraint Algorithms for Linked Rigid Bodies*, Journal of Computational Chemistry, **19**, 102–111, 1998.
8. Andrew R. Leach, *Molecular Modelling. Principles and Applications*, vol. second edition, Pearson Education Ltd., England, 2001.
9. Jesus A. Izaguirre, Daniel P. Castarello, Justin M. Wozniak, and Robert D. Skeel, *Langevin stabilization of molecular dynamics*, J. chem. Phys., **114**, 2090–2098, 2001.
10. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Dinola, and J. R. Haak, *Molecular dynamics with coupling to an external bath*, J. chem. Phys., **81**, 3684–3690, 1984.
11. D.L. Ermak and J.A. McCammon, *Brownian dynamics with hydrodynamics interactions*, J. Chem. Phys., **69**, 1352–1360, 1978.