# Molecular Dynamics Simulation of RNA

YIXIONG SUN, McGill University

**Code: https://github.com/yixiongsun/MDSimulationRNA**
**Video: https://www.youtube.com/watch?v=kaTLkjoZ7P0**

Additional Key Words and Phrases: molecular dynamics, numerical integration, structural biology

**ACM Reference Format:**
Yixiong Sun. 2021. Molecular Dynamics Simulation of RNA. 1, 1 (April 2021), 3 pages. https://doi.org/0000001.0000001

## 1 INTRODUCTION

Molecular Dynamics (MD) is the physical simulation of the motion of microscopic particles, specifically atoms and molecules to describe the dynamics of a system. While multiple approaches to MD simulations exist, the most common approach to simulating the trajectory of the system involves numerically solving classical Newtonian mechanics where forces are derived from the potential energy of the system. MD simulations are used in a variety of fields, especially in structural biology to examine atomic level properties that cannot be easily observed such as local motions or large scale interactions. Here, we implement a simple MD simulation software to analyze the dynamics of a specific type of biomolecule, RNA.

## 2 RELATED WORK

The concept of using MD simulations to explore microscopic properties dates back to the 1950s to the simulation of liquids [Alder and Wainwright 1957]. Since then, there has been multiple algorithms and approaches to MD in several fields. Several different types of force fields have been parameterized from experimental data, including CHARMM [MacKerell et al. 1998], AMBER [Cornell et al. 1995], GROMOS [Schmid et al. 2011] and UFF [Rappe et al. 1992]. Although the field of MD is well established and is mainly used as a technique to investigate microscopic dynamics, recent advances in MD simulations take advantage of better computing hardware including scalability with GPU acceleration [Phillips et al. 2020].

## 3 METHODS

In the following sections, we describe a brief overview of the methods used to implement the MD simulation. The described approach is a simplified version of the classical MD simulation approach, using Newtonian mechanics to describe the equations of motion:

$$m_i \frac{d^2\mathbf{r_i}}{dt^2} = F_i \tag{1}$$

### 3.1 Force Fields

A force field is defined as a vector field that describes non contact forces acting on a particle in specific areas in space. In the case of MD, the force field represents the inter-atomic potential energy, including bonded intramolecular interactions and non-bonded intermolecular interactions. Here in our implementation, we use a simplified version of the CHARMM36 force field with parameters designed for RNA molecules [Best et al. 2012].

We define the force field as $U(r_1, r_2, ..., r_n)$, where $r_i$ is the 3D position of the particle in space and the function $U$ is the potential energy of the system. The equation for $U$ in our simulation is defined as:

$$\begin{aligned} U = &\sum_{bonds} \frac{1}{2}k_b(r - r_0)^2 + \sum_{angles} \frac{1}{2}k_\theta(\theta - \theta_0)^2 \\ &+ \sum_{LJ} 4\epsilon_{ij}((\frac{\sigma_{ij}}{r_{ij}})^{12} - (\frac{\sigma_{ij}}{r_{ij}})^6) + \sum_{elec} \frac{q_i q_j}{\epsilon r_{ij}} \end{aligned} \tag{2}$$

where the first two terms describe the intramolecular interactions and the last two terms describe the intermolecular interactions that contribute to the potential energy.

The model for intramolecular interactions is defined by simple elastic forces (spring forces) between atoms. The first term,

$$U_{bond} = \sum_{bonds} \frac{1}{2}k_b(r - r_0)^2 \tag{3}$$

can be broken down into an elastic potential between particles $r_i$ and $r_j$ only if they are connected with a bond. Thus the potential of an individual bond is

$$U_{bond}(r_{ij}) = \frac{1}{2}k_b^{ij}(r_{ij} - r_0^{ij})^2 \tag{4}$$

and the total potential energy from the bonds is the sum over all bonds. Similarly for the second term,

$$U_{angles} = \sum_{angles} \frac{1}{2}k_\theta(\theta - \theta_0)^2 \tag{5}$$

the potential energy can be broken down into an elastic potential but with respect to the angle between the bonds. Let $r_i, r_j, r_k$ be consecutive bonded particles, then the angle is

$$\theta_{ijk} = \arccos \frac{\mathbf{r_{ij}} \cdot \mathbf{r_{kj}}}{r_{ij}r_{kj}} \tag{6}$$

and the potential of an individual angle is

$$U_{angles}(r_{ijk}) = \frac{1}{2}k_\theta^{ijk}(\theta_{ijk} - \theta_0^{ijk})^2 \tag{7}$$

Note, the bolded $\mathbf{r}$ represents the vector between particles. An additional Urey-Bradley term is also added to some of the particles $r_i$ and $r_k$ of the angles, using the same formulation as the bond potential as a correction term based on experimental observations.

For the intermolecular interactions or non-bonded interactions, they describe the Van der Waals forces and the electrostatic interactions. The Van der Waals force is described using a 12-6 Lennard Jones potential as shown in Figure 1 and the electrostatic interaction is

Author's address: Yixiong Sun, McGill University, yixiong.sun@mail.mcgill.ca.

2021. XXXX-XXXX/2021/4-ART $15.00
https://doi.org/0000001.0000001

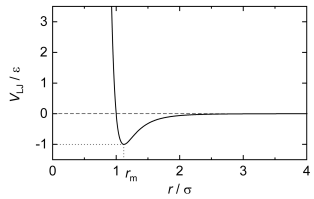


Fig. 1. Graph of the 12-6 Lennard Jones potential, where the intermolecular potential energy is graphed as a function of distance. Particles at closer distances than the potential minimum experience repulsion from overlapping orbitals while particles are farther distances experience attraction from dispersion forces.

represented with the Coulomb's Law. Similar to the intramolecular interactions, these potentials can be broken down into individual $r_i$ and $r_j$ components, however there must be at least 3 particles in between $r_i$ and $r_j$. This is because we want to avoid numerical issues with particles that are too close to each other (notice the LJ potential blows up) and that these local interactions are already captured by the intramolecular terms.

### 3.2 Evaluating forces

To calculate the forces on each atom, the force is defined as

$$F = -\nabla U \tag{8}$$

The explicit derivative is easily calculated for the LJ potential and electrostatic potential as shown below.

$$F_{ij} = -\frac{\partial U_{LJ}}{\partial_{\mathbf{r_{ij}}}} = 4\epsilon_{ij}(13\frac{\sigma_{ij}^{12}}{r_{ij}^{13}} - 6\frac{\sigma_{ij}^{6}}{r_{ij}^{7}})\frac{\mathbf{r_{ij}}}{r_{ij}} \tag{9}$$

$$F_{ij} = -\frac{\partial U_{elec}}{\partial_{\mathbf{r_{ij}}}} = \frac{q_i q_j}{\epsilon r_{ij}^2}\frac{\mathbf{r_{ij}}}{r_{ij}} \tag{10}$$

where $F_{ij}$ is the force exerted by particle $r_j$ on particle $r_i$. Thus by Newton's third law

$$F_{ji} = -F_{ij} \tag{11}$$

and summation of the forces, we get the total force acting on particle $r_i$

$$F_i = \sum_{j \neq i} F_{ij} \tag{12}$$

For the intramolecular terms, explicit differentiation is more difficult, especially for the angle terms so numerical differentiation is used to obtain the force.

### 3.3 Initial conditions

To perform a simulation of the system, initial positions and velocities of the system are required. Initial positions of particles in the system are obtain from experimental crystal structures. Velocities are randomly initialized with a uniform distribution in $[-0.5, 0.5]$. Then an adjustment is made to zero the linear momentum of the system by subtracting the center of mass velocity using the equation

$$v_i := v_i - \frac{\sum m_i v_i}{\sum m_i} \tag{13}$$

and the velocities are scaled such that the kinetic energy represents the temperature of the system, where $k_b$ is the Boltzmann constant, $N$ is the number of particles and $T$ is the temperature.

$$v_i := v_i \sqrt{\frac{3k_bT}{\frac{1}{N}\sum m_i v_i^2}} \tag{14}$$

### 3.4 Integrator

To solve the differential equation numerically, an integration scheme is required. We use the velocity-Verlet algorithm [Swope et al. 1982] as shown below

$$r_i(t + \Delta t) = r_i(t) + v_i(t)\Delta t + \frac{1}{2}a_i(t)\Delta t \tag{15}$$

$$v_i(t + \Delta t) = v_i(t) + \frac{1}{2}(a_i(t) + a_i(t + \Delta t))\Delta t \tag{16}$$

where $a$ is the acceleration which can be calculated as $\frac{F_i}{m_i}$.

The velocity-Verlet integrator, which is a symplectic integrator, is chosen because it provides good stability, accuracy (with error $O(t^4)$) and is computationally easy to compute.

### 3.5 Implementation details

The MD simulation code is written in Python in an Anaconda environment. RNA crystal structure are obtained from the Protein Data Bank [Berman et al. 2000]. Parameters used for the force field calculations are from CHARMM36 [Best et al. 2012].Numerical differentiation and GPU acceleration for computing force fields are implemented with the Taichi framework [Hu et al. 2020, 2019]. The trajectories of the system are rendered using PyMOL.

## 4 RESULTS

### 4.1 Simulation of RNA

We simulated several RNA crystal structures retrieved from the PDB using our MD simulation implementation. We selected a time step of 0.5 fs for 2000 steps (1 ps), saving a snapshot of the RNA trajectory every 10 steps. Overall, we observed dynamically evolving RNA structures in the simulation with good stability. We observe dynamic fluctuations of the backbone and nucleic acid residues while preserving the stability of the overall system as shown in Figure 2.

### 4.2 Computation time analysis

To analyze the computation time for each time step of the MD simulation, we used systems of increasing number of particles to evaluate the time taken to run. The expected time complexity of the simulation is $O(N^2)$ where $N$ is the number of particles, which is from the calculation of the LJ and electrostatic potential. For both of these, the force acting on each particle is the contribution from every other particle. However, as shown in Figure 3, we notice the time complexity is closer to linear rather than quadratic. This is likely due to the GPU acceleration, performing these calculations of forces in parallel and drastically speeding up computation time.

(a) t = 0 fs (b) t = 1000 fs

Fig. 2. Images of the RNA structure (pdb code 1IVS) at two different time frames in the simulation.

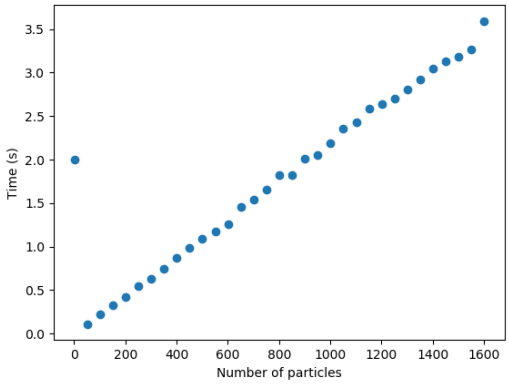


Fig. 3. Plot of simulation time for one time step of force calculation and integration against increasing number of particles in the system.

## 5 CONCLUSIONS

In conclusion, we implemented a simple MD simulation based on classical Newtonian mechanics for RNA structures. Using a simplified version of the CHARMM36 force field, we simulated several RNA structures taken from the Protein Data Bank. We show that the stability of the system is good at small timesteps and that by using GPU acceleration, the computation time for the simulation was drastically faster than expected.

There are many improvements that could be made to the current implementation. Biomolecules are rarely found in isolation in a vacuum and to capture interactions with the environment, solvent molecules can be explicitly used or an implicit model where the solvent is modelled as an external potential could be used. Furthermore, one notable limitation of this implementation is that we do not take into account hydrogen atoms. Hydrogen atoms cannot be determined from crystal structures and thus must be added by the software and since we did not introduce hydrogen atoms, some attractive and repulsive forces may be missing.

## REFERENCES

B. J. Alder and T. E. Wainwright. 1957. Phase Transition for a Hard Sphere System. *The Journal of Chemical Physics* 27, 5 (1957), 1208–1209. https://doi.org/10.1063/1.1743957 arXiv:https://doi.org/10.1063/1.1743957

H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. 2000. The Protein Data Bank. *Nucleic Acids Research* 28, 1 (01 2000), 235–242. https://doi.org/10.1093/nar/28.1.235 arXiv:https://academic.oup.com/nar/article-pdf/28/1/235/9895144/280235.pdf

R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, and A. D. Mackerell Jr. 2012. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone $\phi$, $\psi$ and side-chain $\chi(1)$ and $\chi(2)$ dihedral angles. *Journal of chemical theory and computation* 8, 9 (11 Sep 2012), 3257–3273. https://doi.org/10.1021/ct300400x 23341755[pmid].

W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society* 117, 19 (01 May 1995), 5179–5197. https://doi.org/10.1021/ja00124a002

Y. Hu, L. Anderson, T.-M. Li, Q. Sun, N. Carr, J. Ragan-Kelley, and F. Durand. 2020. DiffTaichi: Differentiable Programming for Physical Simulation. *ICLR* (2020).

Y. Hu, T.-M. Li, L. Anderson, J. Ragan-Kelley, and F. Durand. 2019. Taichi: a language for high-performance computation on spatially sparse data structures. *ACM Transactions on Graphics (TOG)* 38, 6 (2019), 201.

A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, 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, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *The Journal of Physical Chemistry B* 102, 18 (01 Apr 1998), 3586–3616. https://doi.org/10.1021/jp973084f

J. C. Phillips, D. J. Hardy, J. D. C. Maia, J. E. Stone, J. V. Ribeiro, R. C. Bernardi, R. Buch, G. Fiorin, J. Hénin, W. Jiang, R. McGreevy, M. C. R. Melo, B. K. Radak, R. D. Skeel, A. Singharoy, Y. Wang, B. Roux, A. Aksimentiev, Z. Luthey-Schulten, L. V. Kalé, K. Schulten, C. Chipot, and E. Tajkhorshid. 2020. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *The Journal of Chemical Physics* 153, 4 (2020), 044130. https://doi.org/10.1063/5.0014475 arXiv:https://doi.org/10.1063/5.0014475

A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard, and W. M. Skiff. 1992. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American Chemical Society* 114, 25 (01 Dec 1992), 10024–10035. https://doi.org/10.1021/ja00051a040

N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark, and W. F. van Gunsteren. 2011. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *European Biophysics Journal* 40, 7 (30 Apr 2011), 843. https://doi.org/10.1007/s00249-011-0700-9

W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson. 1982. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics* 76, 1 (1982), 637–649. https://doi.org/10.1063/1.442716 arXiv:https://doi.org/10.1063/1.442716