



ELSEVIER

Contents lists available at ScienceDirect

## Data in Brief

journal homepage: [www.elsevier.com/locate/dib](http://www.elsevier.com/locate/dib)



### Data Article

# Trajectories of microsecond molecular dynamics simulations of nucleosomes and nucleosome core particles

Alexey K. Shaytan<sup>a,b</sup>, Grigoriy A. Armeev<sup>b</sup>,  
Alexander Goncarenco<sup>a</sup>, Victor B. Zhurkin<sup>c</sup>,  
David Landsman<sup>a</sup>, Anna R. Panchenko<sup>a,\*</sup>

<sup>a</sup> National Center for Biotechnology Information, NLM, NIH, Bethesda, MD 20894, United States

<sup>b</sup> Biology Department, Lomonosov Moscow State University, Moscow 119991, Russia

<sup>c</sup> Laboratory of Cell Biology, National Cancer Institute, NIH, Bethesda, MD 20892, United States

### ARTICLE INFO

#### Article history:

Received 15 March 2016

Received in revised form

22 March 2016

Accepted 29 April 2016

Available online 6 May 2016

#### Keywords:

Molecular dynamics

Nucleosome

Linker DNA

Histone

Histone tails

### ABSTRACT

We present here raw trajectories of molecular dynamics simulations for nucleosome with linker DNA strands as well as minimalistic nucleosome core particle model. The simulations were done in explicit solvent using CHARMM36 force field. We used this data in the research article Shaytan et al., 2016 [1]. The trajectory files are supplemented by TCL scripts providing advanced visualization capabilities.

Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

### Specifications Table

|                            |                   |
|----------------------------|-------------------|
| Subject area               | Biology           |
| More specific subject area | Chromatin biology |

DOI of original article: <http://dx.doi.org/10.1016/j.jmb.2015.12.004>

\* Corresponding author.

E-mail address: [panch@ncbi.nlm.nih.gov](mailto:panch@ncbi.nlm.nih.gov) (A.R. Panchenko).

<http://dx.doi.org/10.1016/j.dib.2016.04.073>

2352-3409/Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

|                       |   |
|-----------------------|---|
| Type of data          | <i>DCD files, PDB files, PSF files, TCL files</i>   |
| How data was acquired | <i>Molecular dynamics simulations</i>   |
| Data format           | <i>raw</i>  |
| Experimental factors  | <i>Starting structure generated from PDB ID 1KX5</i>  |
| Experimental features | <i>Explicit solvent, CHARMM36 force field</i>   |
| Data source location  | –   |
| Data accessibility    | via National Center for Biotechnology Information FTP server at <a href="ftp://ftp.ncbi.nih.gov/pub/panch/Nucleosome/MD">ftp://ftp.ncbi.nih.gov/pub/panch/Nucleosome/MD</a> |

---

### Value of the data

- Provides an atomic level data about nucleosome dynamics at microsecond timescale.
  - Allows understanding the DNA dynamics in nucleosomes and nucleosome core particles.
  - Shows interaction patterns of flexible histone tails with the DNA molecule.
  - Can serve as a benchmark for further molecular dynamics (MD) simulations of nucleosomes.
- 

## 1. Data

This data article presents microsecond molecular dynamics simulation trajectories which were analyzed in detail in Ref. [1]. Table 1 lists the simulated system names, short description and corresponding subfolder names under which the data was archived at the following URL <ftp://ftp.ncbi.nih.gov/pub/panch/Nucleosome/MD>. The trajectory files are presented in DCD file format, a corresponding topology PSF file and PDB file with the coordinates of the starting nucleosome structure are provided, as well as a TCL script to visualize the trajectory in VMD. Table 2 lists files with the short descriptions. Water molecules and ions are not included in the trajectory files. Each trajectory file contains 10,000 frames spanning one microsecond simulation time. Trajectories can be conveniently viewed in VMD by using the following command: `vmd -e view.tcl`.

## 2. Experimental design, materials and methods

### 2.1. Initial model construction

As a starting point we used high resolution (1.9 Å) X-ray crystal structure of nucleosome core particle formed by recombinant variants of *X. laevis* canonical core histones and modified human  $\alpha$ -satellite DNA (PDB ID 1kx5 [2]). To create a full nucleosome (FN) model with the linker DNA segments, a straight 20 bp long B-DNA duplex (AGTC)<sub>5</sub> was constructed using the NAB software [3].

**Table 1**  
Simulated systems and sub-folder names.

| Simulated system code (see Ref. [1] for details) | System description   | Sub-folder name |
|--|--|-----------------|
| FN   | Full nucleosome together with linker DNA and full-length histone tails at physiological ionic strength | FN_model        |
| NCPm   | Minimalistic NCP model: no linker DNA, truncated histone tails   | NCPm_model      |

---

**Table 2**

Data files provided for every simulation.

| File name          | Description   |
|--------------------|---|
| only_nucl_init.pdb | Initial conformation of nucleosome before equilibration and simulations |
| only_nucl_init.psf | psf topology file   |
| md.dcd             | 1 $\mu$ s MD trajectory saved every 100 ps                              |
| view.tcl           | TCL script for VMD that can be used to view the trajectory              |

The linker DNA sequence is balanced in the number of flexible and rigid base pair steps [4]. It was attached to the core DNA at both ends of NCP. One of the H3 histone tails was slightly rotated to avoid steric clashes with the linker DNA ( $\psi$  angle of Lys36 was set to  $-35^\circ$ ). The minimalistic model of NCP (NCPm) was obtained from the same crystal structure by clipping histone tails at the sites specified in Figure 2 of Ref. [1] by triangles. All models were explicitly solvated in a rectangular box with a minimum distance between the solute and the box boundaries of 20 Å. Sodium ions were added to the system for neutralization, and then additional sodium and chloride ions were added at a concentration of 150 mM with respect to the volume of water. Crystallographic water molecules were retained in the system, while all crystallographic ions were removed. Protonation states of amino acids were assigned based on their solution pK values at neutral pH, histidine residues were considered neutral and protonated on  $\epsilon$ -nitrogens.

## 2.2. Simulation protocols

The CHARMM36 force field was used for DNA and protein [5,6], TIP3P parameters for water molecules and adjusted ion parameters from Luo and Roux [7]. See detailed discussion of force field parameters in Ref. [1].

The simulation systems were prepared with the VMD program [8] and MD simulations were performed with the NAMD 2.9 package [9]. Langevin dynamics with 2 fs integration step, damping parameter of  $0.5 \text{ ps}^{-1}$  and  $T=310 \text{ K}$  were used as means to perform constant temperature simulations. Pressure coupling was implemented via Langevin piston method and set to 1 atm. Simulations were performed with the rigid covalent bonds and Van der Waals interactions were gradually switched off over the distance between 10 and 12 Å. Electrostatic calculations employed PME method with grid spacing of 1 Å, cubic interpolation, 12 Å real space cutoff and direct space tolerance of  $10^{-6}$ . Periodic boundary conditions were used. To remove the nucleosome diffusion, slight constraints of  $0.003 \text{ kcal/mol/\AA}^2$  were applied to C- $\alpha$  atoms of H3 histone folds (residue numbers 64–78, 86–114, 121–131). To avoid base pair fraying at DNA ends in NCPm simulation a restricting artificial wall potential was used to keep the distance between the centers of mass of bases in the terminal base pairs within 120% of the initial.

All systems were initially subjected to energy minimization and initial equilibration via the following protocol: (i) 1000 steps of energy minimization with all protein molecules, DNA and crystallographic waters fixed, (ii) 10,000 minimization steps without atom fixation, (iii) four rounds of 200 ps simulations with elastic constraints on C-alpha atoms of protein and N1, N9 atoms of DNA bases which are gradually relaxed as follows: 300- > 150- > 30- >  $0.9 \text{ kcal/mol/\AA}^2$ .

The production simulations were then performed up to the simulation time of 1  $\mu$ s. The trajectory frames were saved every 100 ps. We run simulations in parallel on high performance computer clusters/supercomputers using effective parallelization available in NAMD. The simulation speed varied depending on the simulated system, number of CPUs and machine architecture. As a reference, the FN model system was simulated in parallel on 384 CPU cores for 120 days progressing at a pace of  $\sim 8 \text{ ns/day}$ .

## Acknowledgments

This research was supported by the Intramural Research Programs of the National Library of Medicine and the National Cancer Institute, National Institutes of Health; and Russian Science Foundation [Grant no. 14-24-00031] (development of nucleosome visualization algorithms). Funding for open access charge: Intramural Research Program, NIH. AS was supported by the US-Russia Collaboration in the Biomedical Sciences NIH visiting fellows program. This study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, Md. (<http://biowulf.nih.gov>). The reported study was supported in part by the Supercomputing Center of Lomonosov Moscow State University and with computational resources of Hexagon Cray XE6m-200 supercomputer, operated by the University of Bergen and the Norwegian metacenter for High Performance Computing (NOTUR).

## Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2016.04.073>.

## References

- [1] A.K. Shaytan, G.A. Armeev, A. Goncareenco, V.B. Zhurkin, D. Landsman, A.R. Panchenko, Coupling between histone conformations and DNA geometry in nucleosomes on a microsecond timescale: atomistic insights into nucleosome functions, *J. Mol. Biol.* 428 (2016) 221–237.
- [2] C.A. Davey, D.F. Sargent, K. Luger, A.W. Maeder, T.J. Richmond, Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution, *J. Mol. Biol.* 319 (2002) 1097–1113.
- [3] J. Macke Thomas, A. Case David, Modeling unusual nucleic acid structures, in: N.B. Leontis, J. SantaLucia, Jr. (Eds.), *Molecular Modeling of Nucleic Acids*, American Chemical Society, 1997, pp. 379–393.
- [4] W.K. Olson, A.A. Gorin, X.J. Lu, L.M. Hock, V.B. Zhurkin, DNA sequence-dependent deformability deduced from protein-DNA crystal complexes, *Proc. Natl. Acad. Sci. USA* 95 (1998) 11163–11168.
- [5] R.B. Best, X. Zhu, J. Shim, P.E. Lopes, J. Mittal, M. Feig, A.D. Mackerell Jr., 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, *J. Chem. Theory Comput.* 8 (2012) 3257–3273.
- [6] K. Hart, N. Foloppe, C.M. Baker, E.J. Denning, L. Nilsson, A.D. Mackerell Jr., Optimization of the CHARMM additive force field for DNA: improved treatment of the BI/BII conformational equilibrium, *J. Chem. Theory Comput.* 8 (2012) 348–362.
- [7] Y. Luo, B. Roux, Simulation of osmotic pressure in concentrated aqueous salt solutions, *J. Phys. Chem. Lett.* 1 (2010) 183–189.
- [8] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, *J. Mol. Graph.* 14 (1996) 33–38.
- [9] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, Scalable molecular dynamics with NAMD, *J. Comput. Chem.* 26 (2005) 1781–1802.