Report: One microsecond nucleosome core particle MD simulations without tails

Alexey Shaytan

August 9, 2014

Abstract

We describe and analyze the results of MD simulations of nucleosome core particle without histone tails. The length of the simulations is 1 μs. We discuss the applicability of MD simulations at this time scale to get interesting biophysical insights, analyze the equilibration and stability of simulations. The analysis plan for this simulations includes following points and question:

a. system equilibration (ions and waters, RMSD - different components).

b. how does the simulations structure differs from crystal ?

c. global movements: DNA rmsd, PCA, several images

d. ion positions

e. waters - diffusion, positions, polarization

f. DNA (12 param profiles, sugars, broken line approximation, grooves and dimensions)

g. contact maps - general characterization

h. local contacts - Arginines.

i. mean electrostatic potential

j. MMPBSA - energy, entropy estimations.

k. make conservational and mutational profiles - compare with contact maps or MMPBSA mutation analysis

Later we will compare these results with simulation results in another force field, and with simulation with linkers and tails.

Contents

1 System setup

The initial structure for simulations is PDB ID 1kx5: 147 bp of human α-satellite DNA with *Xenopus laevis* histones, resolution 1.94 Å, R-value 0.209, R-Free 0.275 [1]. The 1kx5 structure has 3130 water molecules, 4 chloride anions, 14 manganese (Mn) bivalent cations.

To prepare a structure for simulations following steps where performed:

1. Histone tails truncated at following positions (including the declared position): H3 [K37-], H4 [K16-], H2A[A12-K118], H2B[K21-]. Neutral patches applied (acetylated N-terminus, N-methylamide on C-terminus).
2. All ions removed from the system, crystallographic waters retained in the system
3. DNA patches applied - no phosphate groups on 3' or 5' end, hydrogen atoms instead of phosphorus.
4. The system was solvated in a rectangular box with a minimum distance between the NCP and box borders of 20 Å
5. Sodium ions were added to the system for neutralization, and then additional sodium and chloride ions at a concentration of 0.150 mM with respect to the volume of water.

The system was equilibrated using following protocol:

1. System minimized for 1000 steps with all protein, DNA and crystallographic waters fixed
2. System minimized for 10 000 steps with all atoms unfixed
3. System simulated for 200 ps with constraints on CA and P atoms of 300 kcal/mol/A^2
4. System simulated for 200 ps with constraints on CA and P atoms of 150 kcal/mol/A^2
5. System simulated for 200 ps with constraints on CA and P atoms of 30 kcal/mol/A^2
6. System simulated for 200 ps with constraints on CA and P atoms of 0.9 kcal/mol/A^2

Simulation protocol and parameters:

a. Force field CHARMM36

b. Constraints on DNA ends: distance between the centers of mass of (N1 C2 C5 C4 C6 N3) atoms of complimentary bases were constrained with 20 kcal/mol /A^2 potential if the distance increased more than 1.2 times from initial value.

c. Langevin dynamics, 2 fs integration step, temperature 310K, Langevin damping parameter 0.5 ps^{-1}

[1] Some concerns about the structure include: (1) with respect to UNIPROT PDB structure has substitutions G102A in H3 (UNP ID P84233), G99R in H2A (UNP ID P06897), S32T (UNP ID P02281), (2) the quality of atomic positions in regions with high B-factor might be questionable, including DNA (further research is needed).

d. Pressure coupling via Langevin piston method at 1 atm (PistonPeriod= 200.0,PistonDecay =100.0)

e. Bond length constrained for hydrogen atoms (? or all), SHAKE algorithm, SETTLES for water molecules

f. VdW interaction cut-off 12 Åwith switching starting at 10 Å.

g. Electrostatics using PME method, with grid spacing of 1 Å, cubic interpolation, direct space tolerance of 10^{-6}, real space cut-off 12Å.

h. Constraints to remove system diffusion applied to CA atoms of H3 histone fold (residue numbers 64-78,86-114,121-131), constraint constant 0.003 kcal/mol/A^2.

i. Total number of steps: 500 million - 1 micro second.

System summary:

a. 210 331 atoms: 22422 (w/o H 12353) NCP atoms [from them 13076 (w/o H 6332) protein atoms and 9346 (w/o H 6021) DNA atoms], 177 chloride anions, 379 sodium ions, 62451 water molecules (187 353 atoms).

b. Equilibrated system size: 145x141x101 $Å^3$ = 2065 nm^3

c. Water volume: 1867 nm^3

d. Ion concentration by water volume: Na+ 337 mM, Cl- 157 mM.

e. Ion concentration by system volume: Na+ 305 mM, Cl- 142 mM.

2 System equilibration, relaxation properties: time scales, stationarity

In MD simulations it is common to perform an equilibration run before the production run, since the system upon initial construction is in most cases in a conformation far from thermodynamics equilibrium. This refers to both biomolecule, solvent and ions. The equilibration run is usually preceded by energy minimization, and often constraints are imposed on backbone atoms during relaxation to facilitate the process.

In reality the equilibration-relaxation is a complex multi time scale process that spans time scale from rather short (equilibration of water, ions around the protein, equilibration of exposed side chain rotamers) to those that might be well at or over the limits of MD simulations (global protein conformational rearrangements, buried side chain rotations, allosteric transitions).

In this section we only analyze some general aspects of equilibration (we can say, necessary but not sufficient condition that must be met) and the equilibration run itself, while stating that such analysis should be extended to relevant time scales later for each parameter under study.

The scale of different conformational rearrangements and relaxation processes in our system range from very local (eg. water and ions equilibration) through intermediate (e.g. protein side

chain rotations, DNA conformation relaxation) to system wide (eg. DNA wrapping/unwrapping). Analysis of standard parameters (energy, temperature, pressure) shows that 100 ps is enough for them to achieve their stationary values.

2.1 Water and ions relaxation

Normally local and global relaxation of water should not be a problem, the TIP3P diffusion constant is 5.19*10^{-5} cm^2/s, which is 5.19 nm^2/ns - this shows that given the nucleosome dimension of 10 nm, 20 ns is enough for waters to diffuse this distance. From radial distribution plots between DNA phosphorus and different ions we see that 1 ns is enough for the distributions to become stationary.

2.2 Nucleosome conformation equilibration

The RMSD plot for the initial relaxation procedure (minimization + 4*200 ps consecutive relaxation runs with gradual release of constraints, see earlier) is presented in Fig. 1. We can not draw much conclusions about the global drift from the crystal structure, since even at the end of the last relaxation run the harmonic constraints on CA and P positions were still 0.9 kcal/mol/A - which account for the small RMSD of CAs and DNA backbone. On the other hand since the side chains were not restrained, we can estimate from this plot the value of RMSD for side chains contributed solely by their rearrangements and thermal noise, which is around 1 Å.

We leave the discussion of conformational equilibration in the production run until section **??**.

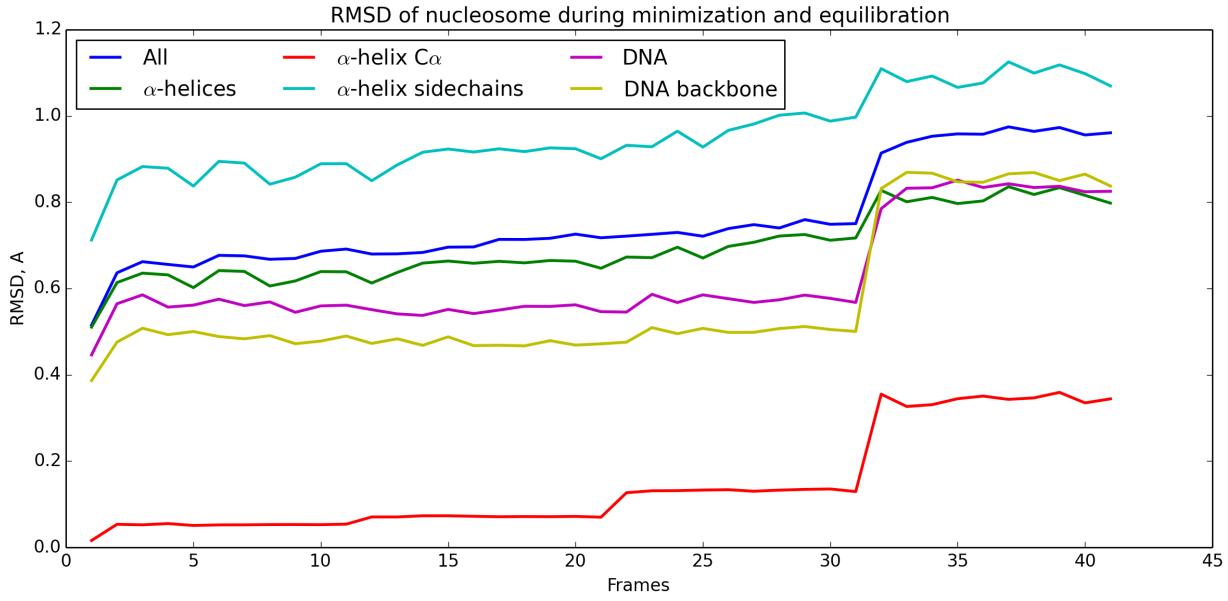

Figure 1: RMSD for atomic positions during initial period of relaxation: frame 1 - after 1000 minimization steps, next 40 frames represent 4*200 ps consecutive relaxation runs with gradual release of constraints on CA and P (see equilibration protocol described earlier). α-helices as defined by STRIDE program. All RMSD calculated using crystal structure reference.