Nucleosome Energetics of Highly Occupied Sequences

Victoria Bamburg, Ronald Brown, James Solow, James, Lineman, Thomas C. Bishop
Louisiana Tech University

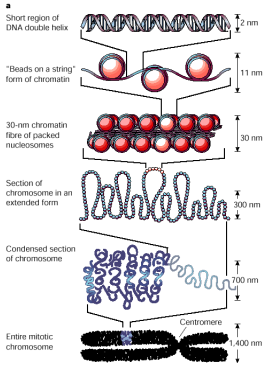
LA-SiGMA
Louisiana Alliance for Simulation-Guided Materials Applications

Abstract

In higher organisms, DNA is packed in the nucleus by histones. Eight histones wrap 147bp long segments of DNA into a left-handed superhelix forming a nucleosome. The location of nucleosomes within the genome plays a crucial role in all genomic processes because it regulates the accessibility of DNA. DNA sequence can influence the location of nucleosomes within the genome by one of two criteria; the DNA internal energy or the interaction energy between DNA and the environment. Internal energy is the associated energy with DNA and the conformation. Interaction of DNA with the environment is between the solvent or histone and the DNA. It is unclear which energy dominates and how differences in these energies compare to other energy considerations within the cell. We analyzed molecular dynamics simulations of nucleosomes containing DNA sequences from *S. cerevisiae* corresponding to highly occupied sequences to determine the behavior of selected DNA. DNA self interaction and interaction between DNA and environment (histones & solvent) were calculated to see if these energies as observed in Molecular Dynamics (MD) correspond with known positions in *S. cerevisiae*. Lower energy areas or areas with high environmental factors were noted as potential nucleosome positioning sites. Our results were compared to previously found experimental and theoretical nucleosome bind sites.

Introduction

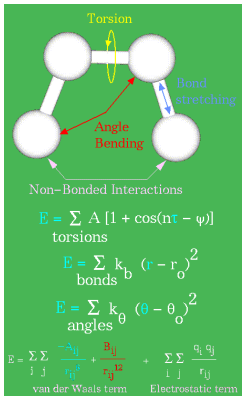
Interactions between DNA and non-histone proteins may influence nucleosome positioning or play a role in determining nucleosome location [1]. Differences in interaction energies as a function of DNA sequence should correlate with nucleosome positioning. Thermodynamically the relative population of two locations on a given sequence of DNA is determined by the ratio of the energies associated with the two nucleosomes; however, these simple thermodynamic arguments can be overridden by cellular machinery or other processes which utilize ATP to affect reorganization of chromatin or by nucleosome-nucleosome interactions. In the simulations studies here, sequences of DNA, corresponding to known nucleosome positions are thread, base pair by base pair, around a histone core. From the simulations internal energies and external interaction energies are calculated to determine if these molecular dynamics energies correspond to experimentally determined positions on S. cerevisiae. Any discrepancies in location can be examined more closely so that we can determine why they are located elsewhere and how this affects the structure and expression of the sequence.



Acknowledgments

The current work is funded by the NSF EPSCor LA-SIGMA project under award #EPS-1003897

Molecular Dynamics

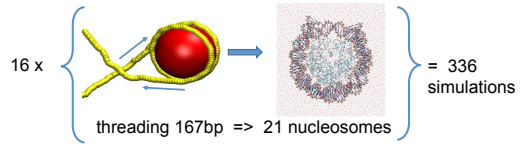


Molecular dynamics can simulate changes in a molecule as a function of time after or be used to sample the range of conformations accessible to a molecule at equilibrium. To make the molecule's environment realistic, the structure is placed in a "bath" of thousands of water molecules. As described below, if the energy landscape in a large molecule is known, the forces acting on those atoms can be deduced.

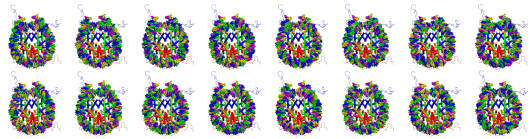
The bonded interaction energy terms describe the geometrical arrangement of atoms that are held together by chemical bonds. The torsion energy in MD is primarily used to correct the other energy terms rather than to represent a physical process. The torsional energy represents the amount of energy that must be added to or subtracted from the Stretching Energy + Bending Energy + Non-Bonded Interaction Energy term to make the total energy agree with experiment.

The non-bonded energy accounts for steric repulsion, van der Waals attraction and electrostatic interactions. van der Waals attraction occurs at short range, and rapidly dies off as the interacting atoms move apart by a few angstroms.[1] http://employees.csbsju.edu/hjakubowski/classes/ch331/protstructure/mechdynam2.html

Methods



The 21 simulations represent the threading of a single 167 base pair long sequence onto the histone core, thus all nucleosomes have 127 base pairs is common. The primary difference between nucleosomes is the relative position of the 127 base pair kernel with respect to the histone core.



Models:
We identified 16 sequences corresponding to the 16 most well-positioned nucleosomes of yeast, one sequence for each chromosome [2]. We expanded the sequences to include 20 additional base pairs, 10 upstream and 10 down stream from the ideal position. All atom nucleosome models were created by threading each 147 base pair subsequence of the 167 base pair parent sequence onto protein databank entry pdbid 1kx5. The 21 nucleosomes allow us to assess positioning over a full turn of the DNA helix on each side of the experimentally determined positioning sequence. For each model nucleosome, the 147bp oligomer is extended by adding two G:C basepairs on each end. This helps stabilize the ends during simulation. An explicit TIP3 solvent shell and sufficient NaCl to both neutralize the system charge and provide a bulk ion concentration of 150mM are added to each nucleosome model. Each fully solvated system contains approximately 160,000 atoms.

Simulations

The NAMD energy plug-in for VMD (23) was used to calculate the energies displayed in Figures 1 through 6 displayed below. The energies were calculated from saved simulation trajectories using only the last nanosecond (19-20ns) of any given simulation. A total of 100 snapshots were evaluated for each simulation. The DNA self-energy represents the energetics associated with the 127 sub-sequence kernel from each simulation. The DNA self-energy calculations used for analysis differed from the calculation used during the simulations. Specifically, the nonbonded terms used a cutoff at 9A instead of including complete long range interactions. Since we are using the DNA self-energy as a metric for localized stress or deformations of the DNA, the 9A cutoff is deemed acceptable. To assess the interaction between DNA and the environment, we determined the nonbonded (vdw and electrostatic) interactions between the 127 sub-sequence kernel and all other atoms in the simulation. For this energy calculation, we utilized the same periodic boundary conditions and PME long range calculation for the electrostatic and van der waals energies as were employed during the simulations. (In this manner, our analysis accurately represents the interactions between DNA and the environment that governed the behavior of DNA during the simulations.)

Results

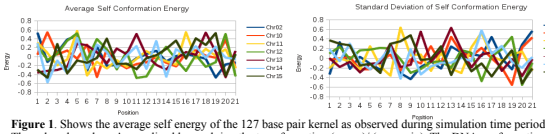


Figure 1. Shows the average self energy of the 127 base pair kernel as observed during simulation time period 19-20ns. The values have been 'normalized by applying the transformation (x-avg)/ (max-min). The DNA conformation energy is a sum of all energy terms.
Figure 2. Shows the standard deviations obtained for the DNA conformation energy. The values have been normalized as in Figure 1.

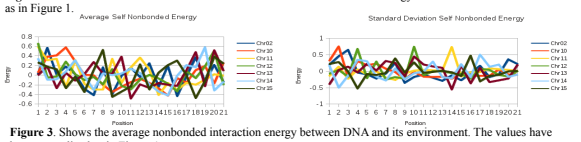


Figure 3. Shows the average nonbonded interaction energy between DNA and its environment. The values have been normalized as in Figure 1.
Figure 4. Shows the standard deviations obtained for the nonbonded interaction energies displayed in Figure 3. The values have been normalized as in Figure 1.

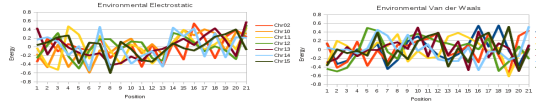


Figure 5. Shows the average Environmental Electrostatic energy normalized by the previously stated method (figure 1).
Figure 6. Shows the average van der Waals normalized energy.

To date only seven sets out of the sixteen set of simulations have been analyzed. Neither the DNA self-energy nor its interaction with the environmental exhibited a clear pattern consistent with a single well positioned nucleosome. However both the DNA self energy and the electrostatic interactions between DNA and the environment show some tendency to "curve up" on each plot, suggesting that a shallow minimum may exist.

Conclusions

There are a number of reasons why the data analyzed did not offer any conclusive results including: the physics regulating nucleosome occupancy and variability in yeast are not governed simply by DNA sequence, our simulations may not yet have equilibrated, or our data must be more carefully analyzed to determine if trends do in fact exist. And finally, even if DNA sequence does govern positioning we may not be able to see it with the techniques employed here. This is consistent with the biologic fact that nucleosomes must fold any sequence of DNA in order to achieve their primary function of compacting DNA into the cell nucleus.