Nucleosome Binding in Highly Occupied Sequences

Victoria Bamburg1, James Solow[2], Ronald Brown3, James Lineman[1], Thomas C. Bishop4

1Department of Biomedical Engineering, Louisiana Tech University
2Department of Computational Analysis and Modeling, Louisiana Tech University
3Department of Physics, Ratford University
4Department of Physics and Chemistry, Louisiana Tech University

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 bind sites. Our results were compared to previously found experimental and theoretical nucleosome bind sites.

Keywords: Nucleosome, self-energy, molecular dynamics

1. Introduction

Interaction between functional DNA sites with non-histone proteins influence a nucleosome positioning and may play a role in determining nucleosome location [1]. Energy differences as function of DNA sequence should tell us about the nucleosome positioning determination. Thermodynamically the relative population of two locations on a given sequence is determined by the ratio of the energies associated with the two nucleosomes; however, these simple thermodynamic arguments can be can be overridden by cellular machinery or other processes(eg. remodeling proteins) which utilize ATP (i.e. energy) to affect reorganization of chromatin or even nucleosome-nucleosome interactions. In the simulations, sequences of DNA, corresponding to known nucleosome positions are thread, base pair by base pair, around a histone core. From the simulations internal energies as well as external 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.

2. Methods

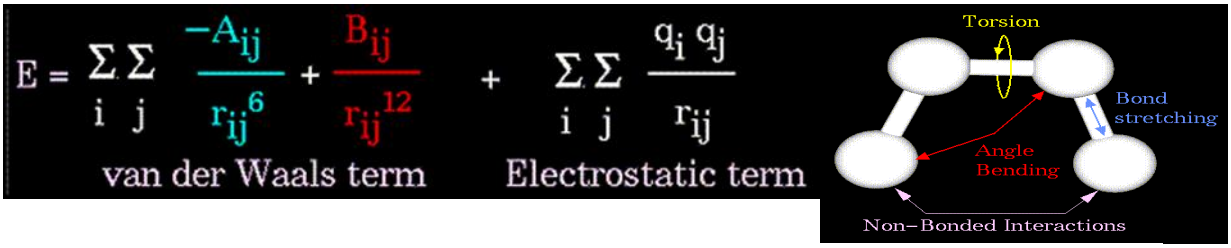
The yeast genome database provides maps of all nucleosomes in the yeast genome based on experimentally determined locations; theoretical predictions are also provided It must be noted that there is variation in the observed nucleosome positions. We selected sequences corresponding to highly occupied and least variable nucleosome locations from each chromosome. This provides 16 sequences (one from each chromosome of *S. cerevisiae*) In the MD simulations a 20 basepair window about these positions with 10 basepair upstream and 10 basepair downstream was scanned. This gave us a 167 basepair sequence and 21 nucleosomes to simulate. In total, there were 16 sequences by 21 simulations giving 336 systems that were simulated. Since there are 147 basepair in a nucleosome, each set of simulations contained a single 127 basepair subsequence kernel. Each system was simulated for 20ns equal to 6.720μs of nucleosome dynamics. Here we look at only the last nanosecond from a subset of the 336 sequences.

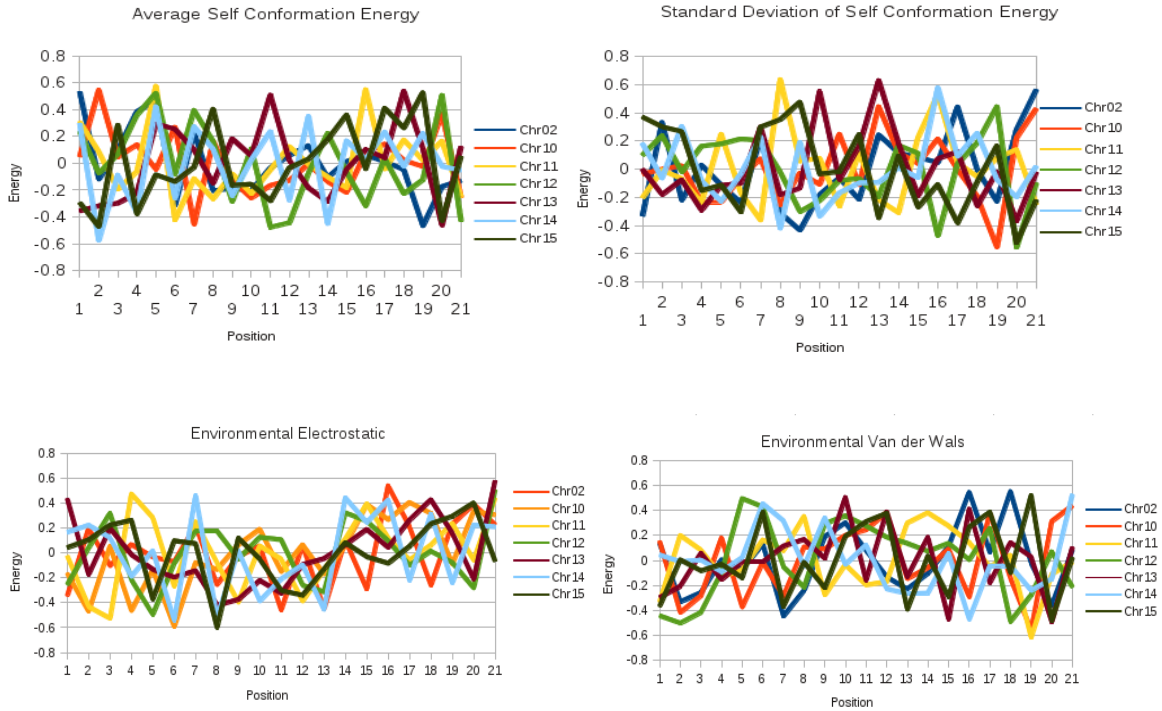
2.1 Simulations

The NAMD energy plug-in for VMD (23) was used to calculate Coulomb and van der Waals interaction energies between the histone core and DNA for 127 basepair A total of 100 equally spaced snapshots from each last nanosecond trajectory were evaluated. For this analysis the water and ions were stripped from the snapshots and a dielectric constant of 80 was used as an ad hoc method of approximating solvent screening. For all energy calculations a cutoff of 400 Å was used with switching that began 300 Å and no periodic boundaries. This provided a complete accounting of the long range Coulomb interactions in our analysis. Note that during the simulations PME with periodic boundaries and a 10-12 Å switching function was utilized to determine long range interactions. Bishop and Mukherjee had 336 nucleosomes each simulated for 20ns. Here we consider only energy from the last nanosecond (time period 19ns to 20ns) for a subset of the 336.Energies in a molecular dynamics experiment are defined by potential energy functions. The DNA self-energy includes bonded and non-bonded, DNA interaction only included the non-bonded terms.

$$E = \sum_{\text{torsions}} A\,[1 + \cos(n\tau - \phi)] \qquad E = \sum_{\text{angles}} k_\theta\,(\theta - \theta_o)^2$$

$$E = \sum_{\text{bonds}} k_b\,(r - r_o)^2$$





3. Results and Discussions

Figure 1. Shows the average self conformation energy that has been 'normalized by applying (x-avg)/ (max-min). The Self conformation energy is the sum of the energy terms.

Figure 2. Shows the normalized standard deviation of the self conformation energy.

Figure 3. Shows the average Environmental Electrostatic energy normalized by the previously stated method (figure 1).

Figure 4. Shows the average van der Wals normalized energy.

Neither the DNA self-energy nor the environmental energy exhibited a pattern consistent with a single well positioned nucleosome between the location of the nucleosome and energy fluctuations. Only seven sets out of the sixteen sequences DNA sequences have been analyzed.

4. Conclusion

Reasons that the data analyzed did not offer any conclusive results are many and include: The physical laws regulating nucleosome occupancy and variability in yeast nucleosome are not governed by DNA sequence. Our simulations may not yet have equilibrated. Our data must be more carefully scrutinized and plots analyzed more extensively to determine if trends do in fact exist. Even if DNA sequence does govern positioning we may not be able to see it with the techniques employed here.

5. Acknowledgments

The current work is funded by the NSF EPSCoR LA-SiGMA project under award #EPS-1003897.

6. References

[1] Miele, V., Vaillant, C., d'Aubenton-Carafa, Y., Thermes, C., Grange, T., DNA physical properties determine nucleosome occupancy from yeast to fly. *Nucleic Acids Research*, 36:3746–3756, 2008.

[2] Fernandez M, Fujii S, Kono H, Sarai A. Evaluation of DNA intramolecular interactions for nucleosome positioning in yeast. Genome Informatics. International Conference On Genome Informatics [serial online]. October 2009;23(1):13-20. Available from: MEDLINE, Ipswich, MA.

[3] Kaplan, N., Hughes, T., Lieb, J., Widom, J., & Segal, E. (2010). Contribution of histone sequence preferences to nucleosome organization: proposed definitions and methodology. Genome Biology, 11(11), 140.

[4] Bishop. T.C. (2009). Geometry of the nucleosomal DNA superhelix. *Biophys J*, 95(3), 1007-1017.

[5] Mukherjee, R., Bishop, T. C. (2011). Nucleosomal DNA: Kinked, Not Kinked, or Self-Healing Material?. *Frontiers Nucleic Acids, 5, 69-92.*