Modeling and simulation of nucleosomes

Yang, Ye

2012

Yang, Y. (2012). Modeling and simulation of nucleosomes. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/50806>

<https://doi.org/10.32657/10356/50806>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

MODELING AND SIMULATION OF NUCLEOSOMES

YANG YE

School of Biological Science

2012

Modeling and Simulation of Nucleosomes

Yang Ye

School of Biological Science

A thesis submitted to the Nanyang Technological University
in fulfillment of the requirement for the degree of
Doctor of Philosophy

2012

Acknowledgement

The author would like to first thank his supervisor Prof. Lars Nordenskiöld for the guidance, gracious help and patience. The author would like to thank Dr. Nikolay Korolev for his work in modeling and experiments with nucleosome, and Dr. Alexander P. Lyubartsev for his expertise in molecular modeling and simulation. The relevant experiments are much indebted to the excellent team of co-workers, Dr. Abdollah Allahverdi, Ms. Liu Ying, and Ms. Lu Cheng Ning.

The author is grateful for Dr. Mu Yuguang for introducing me to molecular dynamics simulation, and also grateful for the following people: Dr. Lin Xin, Dr. Jia Hui, Ms. Cheng Yuhua, Mr. Hong Tian, Mr. Li Weifeng, Mr. Zhang Tong, in computational biology lab; and Dr. Liu Ying, Dr. Yan Jiang, and Mr. Nikolay V. Berezhnoy in the laboratory of Prof. Nordenskiöld.

The author would like to thank School of Biological Sciences, Nanyang Technological University for the opportunity to pursuing graduate studies and generous support of computational facility.

Abstract

Nucleosome is composed of highly charged DNA and histones with unstructured termini exposed to the outside. In nucleus, huge number of them are aggregated and highly compacted. In order to elucidate the roles of tails and cations of various valences in NCP aggregation and single nucleosome array folding, we model NCP as a coarse-grained model of a spherical core with flexible tails, surrounded by explicit mobile ions. By inclusion of such details, we extend the conventional polyelectrolyte model of NCP based on mean-field Debye-Hückel approximation.

Exhaustive simulations under various conditions have been able to reproduce experiment results quantitatively and qualitatively, especially in the effects of multivalent cations and tails modification. We are also able to understand the mechanisms of NCP aggregation in relation to cations and tails more clearly. Our models could be improved with inclusion of more details to describe other inter-nucleosome interactions.

Content

1	Introduction	1
1.1	Nucleosome	1
1.2	NCP.....	2
1.3	Histone	4
1.4	Histone Tails.....	5
1.4.1	Tail Interactions	6
1.4.2	Tail Modification	7
1.5	Compact Structure of Nucleosomes	8
1.5.1	NCP Aggregation	9
1.5.2	Nucleosome Array Folding	11
1.5.3	Compaction in Biopolyelectrolyte and Colloidal System	13
1.6	Theoretical Modeling.....	13
1.6.1	Previous Models	14
1.6.2	Our Approach	18
2	The Sphere Model	20
2.1	Model Setup.....	21
2.1.1	Geometry	22
2.1.2	Variants of Tail.....	23
2.2	Dynamics and Simulation	25
2.2.1	Interaction	25
2.2.2	Dynamics	27
2.2.3	Systems.....	29
2.3	Analysis	32
2.3.1	RDF	32
2.3.2	Cumulative Number.....	33
2.3.3	Contacts per Core.....	33
2.3.4	Trajectory Output	34
3	The Sphere-Bead Model	35
3.1	Model Setup.....	35
3.1.1	Modeling Single NCP: Charges, Sizes and Geometry	35
3.1.2	Modeling Chromatin Array	37

3.1.3	Adoption of Neutrally Charged Core.....	37
3.2	Dynamics and Simulation	40
3.2.1	Non-bonded Interaction.....	40
3.2.2	Bond Interaction	42
3.2.3	Tailless NCP	43
3.2.4	Chromatin Array Modeling in Debye-Hückel Approximation.....	43
3.2.5	Systems.....	44
3.2.6	Dynamics	45
3.3	Analysis	46
3.3.1	RDF, Cumulative Number and Core-Core Contact Number.....	46
3.3.2	Sedimentation Coefficient	46
3.3.3	Radius of Gyration	47
3.3.4	Fiber Length	47
3.3.5	Bulk Concentration	47
4	NCP Simulation with The Sphere Model	49
4.1	Salt-dependent Aggregation.....	49
4.2	Tail Modification	51
4.2.1	Modification of Histone Tails Weakens NCP-NCP Interaction	51
4.2.2	Tail-Core Interaction in Single NCP	53
4.2.3	Tail Modification Positions	55
4.2.4	Titration Simulations of NCP	56
4.3	Discussion	58
4.3.1	Mechanism of Tail Bridging and NCP-NCP attraction.....	58
4.3.2	Relevance of the NCP Simulation to Experiment	59
4.3.3	Approximations in the Sphere Model	61
4.3.4	The Sphere Model and Beyond	62
5	NCP Simulation with The Sphere-Bead Model	63
6	Array Simulation with The Sphere-Bead Model	65
6.1	Preliminary Analysis	65
6.1.1	Size of the Simulation Cell.....	65
6.1.2	Quality of Simulation	66
6.1.3	Cations Distribution.....	67
6.1.4	Convergence of Simulations	68
6.2	Salt-dependent Folding of the Array.....	69
6.3	“Titration” Simulation	71

6.3.1	Folding of wt-Nucleosome Array by Mg^{2+}	73
6.3.2	Folding of wt-Nucleosome Array by $CoHex^{3+}$	74
6.3.3	Folding of wt-Nucleosome Array by Spd^{3+} and Spm^{4+}	75
6.3.4	Folding of Tailless Array	75
6.3.5	Cation-DNA RDF Cumulative Number	76
6.3.6	Estimation of Nucleosome Fiber Dimension	77
6.4	Discussion	78
6.4.1	Mechanism of Array Folding	78
6.4.2	Comparison with Experimental Sedimentation Velocity Studies of the 12-177-601 Array in the Presence of Various Cations.....	79
6.4.3	Comparison with Other Models	81
6.4.4	Chromatin Array Modeling in Debye-Hückel Approximation.....	82
6.4.5	Shortcomings of Debye-Hückel Approximation	85
6.5	Conclusion	86
7	Tailless NCP Simulation with The Sphere-Bead Model	92
7.1	Simulation Results.....	92
7.2	Discussion	95
7.2.1	Experimental Results	95
7.2.2	Comparison of Experiment and Simulation.....	97
7.2.3	Comparison of Simulations: Tailless Nucleosome Array versus g-NCP.....	98
7.3	Conclusion	99
8	Summary and Conclusion	100
	Appendix A: A Primer to Molecular Dynamics Simulation	102
	Reference	107

List of Figures

Figure 1-1 Proposed Hierarchical Structures of Chromatin (adopted from Figure 1 in (Horn and Peterson 2002))	2
Figure 1-2 Structure of the NCP (plot with PyMol (Schrodinger 2010) based on PDB ID: 1KX5 (Davey, Sargent et al. 2002) (Left) Side (Right) Front. DNA is shown with “surface” style and colored by surface potential (both are estimated based on van der Waals radius and simple charge calculation provided by PyMol). Histone proteins are shown with “cartoon” style and colored by different chains. “Tails” outside the DNA are not determined from the crystal but added arbitrarily.	4
Figure 1-3 H3-H4 Dimer, based on PDB ID: 1KX5 (Davey, Sargent et al. 2002). Two histone protein’s helical regions are associated in <i>hand-shaking</i> manner. The dash line indicates the pseudo-dyad axis of the dimer.	4
Figure 1-4 Tail and core domains in four histone (Adopted from (Zheng and Hayes 2003b)) Red T indicates the peptide bond closest to the histone fold domain susceptible to trypsin proteolysis in the nucleosome core. The vertical blue arrows indicate the approximate point where the tails exit either through (H2B/H3) or over/under (H2A/H4) the DNA superhelical gyres.....	5
Figure 1-5 NCP model in (Korolev, Lyubartsev et al. 2006) Left) Atomic structure of NCP with tails extended, adopted from (Wolffe and Hayes 1999). Right) Top view of NCP model (1+88 beads).	17
Figure 1-6 Simulation system built from NCP models (Korolev, Lyubartsev et al. 2006)	17
Figure 2-1 Core histones terminal tails as signaling platforms. Red lollipops represent known phosphorylation sites; green flags represent known acetylation sites (Thorne et al., 1990); red Serine and Histidine residues represent potential phosphorylation sites. Blue Ac-NHs represent N-terminal acetylation of H2A and H4 tails. (Reproduced from Figure 1 in (Cheung, Allis et al. 2000)).....	20
Figure 2-2 Scheme of the Sphere Model.....	21
Figure 2-3 Schematic drawing for tail start bead location (distance in Å).	23
Figure 2-4 Charges on the tails in variants of the Sphere model	25
Figure 2-5 Simulation box with the Sphere model, so-called “ten spiders in a can”.	29
Figure 3-1 The Sphere-Bead model of the NCP. Top view and front view of the Sphere-Bead NCP model illustrated for variant model aNCP, with two acetylation sites shown in larger sized green beads on each of the histone tails. The non-acetylated (recombinant) model shares the same geometry. The histone tails are modeled in the same way as the Sphere model with updated coordinates listed in Table 3-1..	36
Figure 3-2 The Sphere-Bead model of the NCP Array. (A), (B) Partially neutralized DNA beads are in red; fully charged beads modeling the linker DNA are in orange. (C) Model of the “12-177” nucleosome array, in initial configuration in the simulation cell.	37
Figure 3-3 Electrostatic potential molecule surface computed by APBS (Baker, Sept et al. 2001) and rendered in PyMol for protein model 1KX5 at pH=7.0. (Left) Histone protein alone (Right) Whole NCP.	38
Figure 3-4 Core-Core RDF between model with neutral core and model with positive core. The scheme of two NCPs shown near the peak in RDF represents how NCPs stacks locally to be shown as a peak in RDF.....	39
Figure 3-5 Extracting bulk cation density from density across simulation box. The arrows indicate the flat parts.	48
Figure 4-1 RDFs from four rNCP simulation system in different salt conditions. (a) Core-Core RDF. (b) RDFs of the counterions (K⁺ or Mg²⁺) relative to the center of NCPs. In the inset graph showing the far-end of the curve, there are drops in cation concentration in the Normal Salt and Mg Salt systems, implying a boundary between phase with condensed NCP particles and the bulk solution phase. (c) External tail-Core RDF. Appearance of maximum below 6nm distance indicates NCP-NCP aggregation. (d) Internal tail-Core RDF. Normal Salt and Mg Salt systems are similar while Low Salt and Salt Free are similar. The crossover between two groups indicates the detachment of the histone tails from the negative core caused by interaction with other NCP particles observed in the Normal Salt and in the Mg Salt systems.....	49
Figure 4-2 Peak values of (from top to bottom) External tail-Core, Internal tail-Core and Core-Core RDFs (unscaled RDF value) versus the time steps.	50

- Figure 4-3 Results of the simulations for rNCP, aNCP, and paNCP models for the systems with KCl and MgCl₂ salts.** (a) Core-Core RDF and (b) External tail-Core RDF collected for the converged region of the trajectories (last 1.6×10^7 simulation steps). The insert in (b) shows the change in the maximum intensity of External tail-Core RDF during the course of the simulation run. The numerical value of the RDF corresponds to the concentration of particles relative to the value expected for a uniform distribution. (c) Histograms of the averaged peak values of Core-Core RDF and (d) of External tail-Core RDFs under two salt conditions, 4.15 mM MgCl₂ (grid) and 65.4 mM KCl (shade). (e) Close-up view of a snapshot of three NCPs in aggregation, each nucleosome with its core and tails are in one color, showing the presence of tails participating in inter-core interactions..... 51
- Figure 4-4** (a) Internal tail-Core RDFs of rNCP, aNCP, and paNCP in the presence of K⁺ and Mg²⁺ salts for the systems with 10 NCPs in the simulation cell. (b) The intensity of the first maximum of Internal tail-Core RDF of a single-NCP system versus the salt concentration (KCl or MgCl₂). The maxima in (a) are labeled in (b) as filled symbols of the same shape as the corresponding systems in (a). The single unfilled red square with cross and unfilled blue triangle with cross, represent data from the systems corresponding to **Figure 4-6** below at zero percentage of Mg²⁺ in “titration” group of simulations..... 53
- Figure 4-5 Simulation results for aNCP model with two acetylation sites in each tail but with different positions of acetylation.** Values of the first maximum in (a) Core-Core RDF (b) External tail-Core RDF and (c) Internal tail-Core RDF are shown for the four aNCP models, under two salt conditions, 4.15 mM MgCl₂ (grid) 65.4 mM KCl (shade). Data for the “root”, “head”, and “even” aNCP models shown as columns; results of the native aNCP displayed as horizontal lines..... 55
- Figure 4-6 Intensities of the RDF maxima and number of the NCP-NCP contacts determined for rNCP (circles, black line) and aNCP (squares, red line) models in dependence on the Mg²⁺ content in the simulation cell:** (a) Core-Core RDF; (b) External tail-Core; (c) number of contacts per core. Data were averaged for the converged region of the trajectories (last 1.0×10^7 steps of the MD trajectory)..... 56
- Figure 4-7 Contacts per core in the “titration” systems of rNCPs (top) and aNCPs (bottom) with variation of Mg²⁺ content.** Representative snapshots show different degrees of aggregation in the systems with rNCPs (snapshots at the top of the figure) and aNCPs (bottom)..... 58
- Figure 5-1 Core-Core RDF (solid lines) and cumulative number (dashed lines) for two rNCP models in the system with Mg²⁺, computed from last 6×10^7 time steps:** Sphere NCP model (black curves) and the Sphere-Bead NCP model (red curves). Representative snapshots are shown as inserts: (top) The Sphere model and (bottom) The Sphere-Bead model. The conditions for the simulation set up (including the number of mobile ions) are identical to rNCP-MgCl₂ simulations..... 63
- Figure 5-2 Comparison of Core-Core RDFs (solid lines) and cumulative numbers (dashed curves) for the Sphere-Bead model of rNCP (black curves) and aNCP (red curves) calculated in the presence of Mg²⁺, computed from last 6×10^7 time steps and last 2×10^7 time steps respectively. Other conditions (box size, number of particles) are the same as for the systems rNCP-MgCl₂ and aNCP-MgCl₂... 64**
- Figure 6-1 Simulation results of Array with same concentration of KCl salt in three boxes of different size.** (A) Snapshots of the array at the end of the simulation run. (B) Variation of the radius of gyration (R_g , top graph) and sedimentation coefficient ($S_{20,w}$, bottom graph) in the course of the simulations. (C) RDFs calculated from the last 2×10^7 time steps of the simulation or longer. From left to right: Core-Core RDF including closest core in the array; Core-Core RDF excluding closest core in the array (exclusion of the closest core from Core-Core distribution highlights the folding of the array); and External tail-Core RDF..... 66
- Figure 6-2 Autocorrelation function of the system energy (Array-CoHex-1.31mM system) 66**
- Figure 6-3 Above)** Cation distribution (particles in green) in simulation box in the beginning (left) and the end (right) of simulation in Array-CoHex-1.31mM system. **Below)** Cation distribution (particles in yellow) in simulation boxes in the beginning (left) and the end (right) of simulation in Array-Spm-1.06mM system. 67
- Figure 6-4 Monitoring Mean Square Displacement (MSD) of the cations in the simulation of the “12-177” nucleosome array in the presence of various cations.** (A) MSD of K⁺ cations in titration systems containing mixture of K⁺ and various concentration of Mg²⁺, CoHex³⁺, Spd³⁺ and Spm⁴⁺ (identity and concentration of the cations are indicated in the graph; see **Table 6-5** for details) (B-E) MSDs of (B) Mg²⁺, (C) CoHex³⁺, (D) Spd³⁺, and (E) Spm⁴⁺ ions. For each cation, we picked two systems out of all “titration” simulations, one system has $S_{20,w} \sim 35S$ (unfolded) and another is associated with $S_{20,w} > 50S$ (folded). Exception is given to Mg²⁺ for having the 2nd system of $S_{20,w} \sim 40S$ because of its less capability to fold the array than tri-valent cations..... 68
- Figure 6-5 Comparison of the results of simulations of the “12-177” array in the presence of K⁺ (green line in the graphs B-E), Mg²⁺ (blue line) and CoHex³⁺ (red line).** (A) Representative snapshots illustrating typical degree of folding of the array in the presence of the respective cations observed in the end of the simulations. (B) Sedimentation coefficient ($S_{20,w}$) and (C) radius of gyration (R_g) along the course of the simulations. (D) Core-Core and (E) External tail-Core radial distribution functions

obtained by averaging the last 2×10^7 time steps or longer of each of the simulation runs. In (B) Core-Core RDF neighboring (connected by fixed-length linker DNA) NCPs were excluded from statistics. The three peaks clearly visible in Core-Core RDF of the array in the presence of CoHex^{3+} (red line) are related to three possible types of the NCP-NCP contact illustrated by sketches near the corresponding peaks. 70

Figure 6-6 Properties of the “12-177” nucleosome array in the presence of various amount of Mg^{2+} . Concentrations of Mg^{2+} in the simulated systems are indicated on the right-hand side of the graph. Variation of (A) sedimentation coefficient, $S_{20,w}$, and (B) radius of gyration, R_g , in the course of the simulations. (C-E) RDFs calculated from the last 4×10^7 steps of the simulation. From left to right: (C) Core-Core RDF including closest core in the array; (D) Core-Core RDF excluding closest core in the array; and (E) External tail-Core RDF. Number of ions and major results of the simulations are given in Table 6-5(B) below. 72

Figure 6-7 Summary of simulation results for the “12-177” array “titration” by Mg^{2+} (A-D); CoHex^{3+} (E-H); Spd^{3+} and Spm^{4+} (I-K). (A, E) Representative snapshots illustrating folding of the array at different stages of titration by (A) Mg^{2+} and (E) CoHex^{3+} ; numbers under the snapshots indicate bulk concentration (1st line) and average (2nd line) Mg^{2+} or CoHex^{3+} concentration. (B, F, I) Radius of gyration, R_g ; (C, G, J) sedimentation coefficient, $S_{20,w}$; (D, H, K) intensity of the maximum in External tail-Core RDF. In graphs B-D and F-H, data for the array with charged tails are drawn as solid circles and lines; similar results for the tailless array are shown as hollow points with dashed lines. X-axis is the average concentration of Mg^{2+} (B-D, C_{Mg}), CoHex^{3+} (F-H, C_{Co}); x-axis is bulk concentration for Spd^{3+} or Spm^{4+} (I-K, $C_{\text{Spd}}/C_{\text{Spm}}$). 73

Figure 6-8 Cation-DNA RDF cumulative number for wt-nucleosome array (A) K^+ ($C_K = 3.17$ mM), (B) Mg^{2+} ($C_{\text{Mg}} = 1.58$ mM) and (C) CoHex^{3+} ($C_{\text{Co}} = 1.06$ mM); (the same three systems shown in Figure 6-5) and tailless nucleosome array (D) Mg^{2+} ($C_{\text{Mg}} = 4.18$ mM) (E) CoHex^{3+} ($C_{\text{Co}} = 1.65$ mM) Data is shown up to radial distance of 30nm. Linker DNA and core DNA RDFs are calculated individually. 76

Figure 6-9 Simulations within Debye-Hückel approximation as a function of ionic strength for different models of the array. (A) Sedimentation coefficient, $S_{20,w}$; (B) Radius of gyration, R_g . Horizontal dashed lines indicate the values of $S_{20,w}$ and R_g calculated for the array model with uncharged particles. Hollow ovals in (A) and (B) show respectively $S_{20,w}$ and R_g values obtained in the simulations with explicit ions; cations are indicated near the points. Solid ovals in (A) show the maximal $S_{20,w}$ values observed in the AUC measurements for different cations (cation is indicated in the graph). Indices “ Mg^{2+} ” and “high Mg^{2+} ” mark the simulation data calculated respectively for $C_{\text{Mg}} = 4.024$ and 14.752 mM. 84

Figure 7-1 Core-Core RDFs and snapshots from simulation of wild-type NCP and g-NCP with the Sphere-Bead model in Mg^{2+} and CoHex^{3+} 92

Figure 7-2 (a) Core-Core contact number and magnitudes of maximum in (b) External tail-Core RDFs for titrations of systems of ten g-NCPs (upper), or wt-NCPs (lower), with Mg^{2+} (red, square) and CoHex^{3+} (black, hollow circle). 94

List of Tables

Table 1-1 Amino Acids Composition in NCP (according to PDB ID: 1KX5). Total Number of Amino Acids: 980. Net Charge at pH 6.0: +196e. Molecular Weight: 108.2687 kDa. Vol.: 131947 Å ³ .	3
Table 2-1 Histone tail amino acid sequence and charges depicted in (Cheung, Allis et al. 2000). Arginine and Lysine are in bold . The charges are calculated based on protonated amino acids at pH 6.0.	21
Table 2-2 Coordinates (in Å) of the first bead of the histone tails in the NCP model.	22
Table 2-3 Non-titration systems. *: Repeated simulations from (Korolev, Lyubartsev et al. 2006) and the names in the bracket are their previous names. ^: 6 simulations to study variants of NCP rNCP/aNCP/paNCP subjected to two salt conditions. #: 4 variants, “Lysine-site”, “even”, “head”, and “root”	30
Table 3-1 Coordinates (in Å) for the central particle, start/end DNA beads and the first beads of tails closest to the core in the Sphere-Bead NCP model.	36
Table 3-2 Comparison of array properties between systems of neutral core and negative core.	40
Table 6-1 Fiber dimensions for compact “12-177” arrays calculated from the analysis of the simulations	77
Table 6-2 Maximum concentration of cations giving reproducible sedimentation velocity curves and maximum $S_{20,w}$ values obtained in the AUC measurements.	80
Table 6-3 Selected results from simulations of arrays using Debye-Hückel approximation. Simulations of the uncharged model of the “12-177” array give $R_g = 18.4 \pm 1.0$ nm and $S_{20,w} = 43.5$.	83
Table 6-4 Modeling of electrostatic contributions to chromatin folding.	85
Table 6-5 Summary of simulations for the “12-177” nucleosome array.	88
Table 7-1 Simulation result summary with the Sphere-Bead model for wt- and tailless NCP. (Systems shown in Figure 7-1 are denoted by *)	93
Table 7-2 Summary of precipitation assay results. Cation concentrations at 50% NCP precipitation ($C_{50\%}$) from at least 3 independent measurements are shown, with maximal degree of NCP precipitation (Max) also indicated, adapted from Table 2 of (Liu, Lu et al. 2011).	96

1 Introduction

DNA is a highly charged polyelectrolyte but also highly compacted in the nucleus. If fully extended and linearly linked, human DNA is of 2m long but it can fit inside the nucleus of 6 μm in diameter, which translates to a compaction of 100,000 times¹ (Razin, Iarovaia et al. 2007). It has been known that histones and other proteins help achieving such compaction of DNA by wiring string-like DNA onto the bead-like histone octamers, all of which together form the chromatin; yet many mechanisms remain to be elucidated. How this polyelectrolyte can overcome the entropy loss and resist repulsion between the like-charged chain segments during compaction to fit in such a small volume? And at the same time, various DNA regions remain accessible for gene expression, interacting with regulatory factors and transcription/replication factories?

1.1 Nucleosome

In chromatin, nucleosomes are the fundamental repeating units, representing about 85% of eukaryotic chromatin (Noll 1974). The central part of the nucleosome is the nucleosome core particle (NCP). It consists of 145-147 bp of DNA wrapped around the histone octamer composed of two copies each of 4 histone proteins H2A, H2B, H3 and H4. Double-stranded linker DNA of variable length connects NCPs to each other forming “beads-on-string” nucleosome arrays, which further condenses into the 30-nm chromatin fibers (van Holde and Zlatanova 1996; Woodcock and Dimitrov 2001; Hansen 2002; Luger and Hansen 2005). A few intermediates were proposed, from naked-eye visible chromosome during cell division, to naked double helix DNA structure, shown in **Figure 1-1**.

Gene activation/deactivation changes the 3D structure of chromatin (reviewed in (Wegel and Shaw 2005)). There are experimental results showing large-scale chromatin

¹ ~seven-fold compaction from naked B-DNA to 10nm fiber.

40-50-fold compaction from 10nm fiber to 30nm fiber.

~500-fold compaction from 30nm fiber to 100- to 300-nm fiber

decondensation which results in drastic falling of compaction ratio from 30,000-fold to 1,000-fold after transcription activation in mammalian cells (Tumbar, Sudlow et al. 1999). As observed in different organisms, nucleosomes unfold completely (with loss of histones) at transcriptionally active promoters (Boeger, Griesenbeck et al. 2003).

Considering the length of the genome and majority of them existing in nucleosomes, in order to achieve targeted and controlled mechanism for complex gene regulation, we could assume that a ubiquitous mechanism exists in the nucleosome itself. Nucleosome is not merely a structure for DNA organization; it is a platform leading to various fundamental life processes.

Figure 1-1 Proposed Hierarchical Structures of Chromatin (adopted from Figure 1 in (Horn and Peterson 2002))

1.2 NCP

Since the discovery of the nucleosome in mid-1960s (reviewed in (Wu and Grunstein 2000; Olins and Olins 2003; Khorasanizadeh 2004)), structural biologists have embarked on a long journey to discover the high-resolution structure of NCP (reviewed in the introduction of (Harp, Hanson et al. 2000)). Their efforts have resulted in today's near-atomic-scale elucidation of the NCP structure in crystal (Luger, Mader et al. 1997; Harp, Hanson et al. 2000; Davey, Sargent et al. 2002), followed by tetranucleosome structure (Schalch, Duda et al. 2005). To which illustrates how DNA wraps around the octamer, **Figure 1-2** shows the structure of NCP from X-ray crystallography (PDB ID: 1KX5). The NCP is a well-defined protein-DNA complex containing 145-147 bp of DNA wrapped for 1.70-1.75 turns around a cylindrical disk-like histone octamer (Finch and Klug 1976;

Richmond, Finch et al. 1984; Arents and Moudrianakis 1995), There are two copies of each histone protein in nucleosome (Luger, Mader et al. 1997; Harp, Hanson et al. 2000; Davey, Sargent et al. 2002). It is noted that for nucleosome as a kind of highly dynamic and mobile molecule, such structure is just a static snapshot captured. **Table 1-1** below shows the statistics of amino acids in NCP, according to PDB ID: 1KX5. Positively charged Lysine and Arginine are in the top three in terms of the number of amino acids.

Table 1-1 Amino Acids Composition in NCP (according to PDB ID: 1KX5). Total Number of Amino Acids: 980. Net Charge at pH 6.0: +196e. Molecular Weight: 108.2687 kDa. Vol.: 131947 Å³.

Amino Acid			No. of Residues		Mol. Weight	Volume
Lysine	Lys	K	114	11.63%	14612	19220
Alanine	Ala	A	108	11.02%	7676	9568
Arginine	Arg	R	106	10.82%	16557	18380
Glycine	Gly	G	84	8.57%	4793	5048
Leucine	Leu	L	84	8.57%	9506	14002
Threonine	Thr	T	66	6.73%	6673	7662
Valine	Val	V	66	6.73%	6543	9240
Serine	Ser	S	56	5.71%	4876	4984
Glutamic Acid	Glu	E	50	5.10%	6456	6920
Isoleucine	Ile	I	48	4.90%	5432	8001
Proline	Pro	P	36	3.67%	3496	4417
Glutamine	Gln	Q	36	3.67%	4613	5180
Tyrosine	Tyr	Y	30	3.06%	4895	5808
Aspartic Acid	Asp	D	24	2.45%	2762	2666
Asparagine	Asn	N	24	2.45%	2738	2824
Phenylalanine	Phe	F	18	1.84%	2649	3418
Histidine	His	H	18	1.84%	2468	2757
Methionine	Met	M	10	1.02%	1312	1629
Cysteine	Cys	C	2	0.20%	206	217

Figure 1-2 Structure of the NCP (plot with PyMol (Schrodinger 2010) based on PDB ID: 1KX5 (Davey, Sargent et al. 2002) (Left) Side (Right) Front. DNA is shown with “surface” style and colored by surface potential (both are estimated based on van der Waals radius and simple charge calculation provided by PyMol). Histone proteins are shown with “cartoon” style and colored by different chains. “Tails” outside the DNA are not determined from the crystal but added arbitrarily.

1.3 Histone

All histone proteins share a highly similar structural motif, the histone fold (illustrated in **Figure 1-3**), formed by three connected helices (indicated in the gray box in **Figure 1-4**). The structured histone-fold domains form the “core” with four spirally arrayed dimers in “hand-shaking” manner, i.e. two H2A/H2B dimers and a (H3/H4)₂ tetramer. Such fold, by its occurrence in the four dimers, also gives rise to repetitive structural elements on the octamer side surface with dominance of positive charges. Regularly spaced sites on the surface of the histone octamer allow favorable interactions with the DNA backbone. The conformation of nucleosomal DNA reflects the dyad symmetry of octamer (Richmond, Finch et al. 1984) allowing the greatest number of favorable histone-DNA interactions.

Figure 1-3 H3-H4 Dimer, based on PDB ID: 1KX5 (Davey, Sargent et al. 2002). Two histone protein’s helical regions are associated in *hand-shaking* manner. The dash line indicates the pseudo-dyad axis of the dimer.

Figure 1-4 Tail and core domains in four histone (Adopted from (Zheng and Hayes 2003b)) Red **T** indicates the peptide bond closest to the histone fold domain susceptible to trypsin proteolysis in the nucleosome core. The vertical blue arrows indicate the approximate point where the tails exit either through (H2B/H3) or over/under (H2A/H4) the DNA superhelical gyres.

1.4 Histone Tails

The N-terminal domains for all four kinds of histones and C-terminal domains for H2A, counting for about 20% of the total protein, are unstructured flexible “histone tails”. **Figure 1-4** shows the tail amino acid sequences for human histone. The flexible nature of the tails was established in NMR experiments during 1980s (reviewed in (Zheng and Hayes 2003a)). In the high-resolution X-crystallography structure of NCP reported in (Luger, Mader et al. 1997), the large portion of tail domains which is exposed to the exterior of the nucleosome cannot be resolved in crystal, implying no fixed structure. “Tails” are found to protrude outside the NCP between the gyres of the DNA superhelix through the channels formed by the minor groove of DNA (H3 and H2B), or follow the minor grooves over or under the top/bottom of the DNA superhelix (H2A and H4).

According to the protein protonation status at pH 6.0 used in crystal preparation in (Davey, Sargent et al. 2002), core histones carry +56e, which can only partially neutralize the negative charge of the DNA (-292e from 146bp). Eight tails have a net charge in the range +9-14e each (originating from Lysines and arginines), and summed to +88e. Therefore, the NCP carries a significant net negative charge of about -148e, which can be visualized as a central particle (-236e=-292e+56e) to which eight flexible and positively charged histone tails (+88e) are attached.

In review (Luger 2003), histones have been described to play bipartite role in both structure and genetic function. The structured histone-fold domains, “core”, are responsible for creating a stable yet also accommodating binding with DNA. Histone Tails are necessary for formation and maintenance of both secondary and tertiary structure of chromatin and participate in both intra- and inter-array nucleosome-nucleosome interactions (van Holde and Zlatanova 1996; Hansen 2002). The histone tails are therefore essential for regulation of transcription and replication (Wolffe 1998; Horn and Peterson 2002).

1.4.1 Tail Interactions

There are eight possible tail interactions within chromatin (adopted from Figure 2 of (Zheng and Hayes 2003a)), divided into two groups.

Internal Interaction

- 1) Interact with each other, perhaps within the same nucleosome
- 2) Interact with nucleosomal DNA (within NCP)
- 3) Interact with linker DNA (may help to bend it for higher-order structure formation)
- 4) Potentially cause a concerted rearrangement of tail contacts

External Interaction

- 5) Interact with non-nucleosomal proteins
- 6) Posttranslational modifications could alter tail structure and interactions
- 7) Make inter-nucleosomal protein-protein contacts (tail-bridging)
- 8) Make protein-DNA interactions (tail-bridging)

Regarding items 1)-4), we can consider Lys-rich and Arg-rich “tails” bearing positive charges and DNA (an electrolyte) phosphates bear negative charges, this localized interaction could be due to electrostatic interactions, stabilizing the DNA binding inside nucleosome.

Items 7) and 8) arise from electrostatic origin, i.e. tails bearing positive charge and core/DNA bearing negative charges, as well as the structural featuring flexible tails.

Modification on the tail directly affects the binding ability of the tail via the change of charges and possible structural change.

Items 5) and 6) are of profound significance because they provide vast possibility from numerous modifications to alter all other interactions. From structural point of view, tails could be easily accessible for protein factors because they are exposed outside.

1.4.2 Tail Modification

Amino acids in the histone tails exposed outside DNA can be regarded as a platform for post-translational covalent modifications. The seemingly featureless “tails” of NCP can actually produce huge number of variants upon biochemical modification, leading to different pathways of gene expression, reviewed in (Khorasanizadeh 2004). There are at least eight distinct types of modifications found in histones. Numerous specialized proteins precisely and dynamically regulate the degree and positions of covalent modifications (Hansen, Tse et al. 1998; Mizzen, Kuo et al. 1998; Strahl and Allis 2000; Wu and Grunstein 2000; Waterborg 2002) so as to modulate the biological role played by the underlying DNA.

Notably, the most common reactions, *acetylation* of Lys⁺ (removal of a positive charge), *phosphorylation* of serine/threonine (adds negative $-2e$ charge), change the distribution and net amount of charged groups in the tails (Luger and Richmond 1998; Wolffe and Hayes 1999; Ren and Gorovsky 2003). Experimental observations suggest that these changes are crucial for functional regulation (Dou and Gorovsky 2000; Ren and Gorovsky 2001; Dou and Gorovsky 2002).

Histone acetyltransferases (HAT), which belong to N-alpha-acetyltransferases (NATs), a sub-family of the GNAT superfamily of acetyltransferases, are enzymes that acetylate conserved Lysine amino acids on histone proteins by transferring an acetyl group from acetyl CoA to form ϵ -N-acetyl Lysine. To reverse the modification, there is a factor called Histone deacetylase (HDAC). In general, histone acetylation is linked to transcriptional activation and associated with euchromatin. Phosphorylation is the addition of a phosphate (PO_4^{3-}) group to a protein or other organic molecule, which activates or deactivates many other protein enzymes.

Below we give a limited selection of possible citations with emphasis on early original work and on recent reviews. It is now generally acknowledged that: i) acetylation of the histone

tails correlate with transcriptionally active regions in chromatin, see for example (Turner 1991; Kurdistani, Tavazole et al. 2004; Calestagne-Morelli and Ausio 2006; Szerlong, Prenni et al. 2010); ii) acetylated regions of chromatin are more sensitive to nuclease digestion which indicates higher accessibility of DNA (Lewis, Guillemette et al. 1988); iii) acetylated chromatin shows higher solubility in salt (NaCl, MgCl₂) compared to the control cores, such solubility is more sensitive to smaller amount of Mg²⁺ than Na⁺ (Perry and Chalkley 1981; Perry and Chalkley 1982; Tse, Sera et al. 1998). This shows that cation valence is also significant to inter-nucleosome interactions.

1.5 Compact Structure of Nucleosomes

A mandatory feature of condensed nucleosome and chromatin structures is the formation of a close NCP-NCP contact between the rather flat surfaces of the globular histone octamer core on both sides of the cylindrical wedge-shaped NCP. Close stacking is observed in electron microscopy images (Finch, Brown et al. 1981; Pelta, Durand et al. 1996; Leforestier and Livolant 1997; Leforestier, Dubochet et al. 2001; Mangelot, Leforestier et al. 2003a; Bertin, Mangelot et al. 2007), in NCP crystals (Finch, Lutter et al. 1977; Finch, Brown et al. 1981; Luger, Mader et al. 1997; Davey, Sargent et al. 2002; Mangelot, Leforestier et al. 2003b; Mangelot, Leforestier et al. 2003a), in the crystal of the tetranucleosome, (Schalch, Duda et al. 2005) and in nucleosome arrays (Dorigo, Schalch et al. 2004; Robinson, Fairall et al. 2006).

It is well known that DNA spontaneously undergoes a collapse of single molecules and/or intermolecular aggregation upon addition of polycations of charge $>+2e$ (Bloomfield 1997; Iwaki, Saito et al. 2007). A distinctive feature of such DNA condensation is a prominent oligocation-charge dependency for a certain concentration of oligocation needed to induce the transition (Korolev, Berezhnoy et al. 2009). As histones compact DNA from basic unit nucleosome to higher-order structure, how similar polyelectrolyte effects influence chromatin folding and compaction beyond the effect of monovalent, and divalent cations has rarely been addressed (Widom 1986; Pollard, Samuels et al. 1999) and has yet to be established.

Recent experiments indicate that aggregation of NCPs (Bertin, Renouard et al. 2007) as well as intra-array compaction (folding) and inter-array aggregation of the chromatin arrays (Dorigo, Schalch et al. 2003; Zheng and Hayes 2003b; Zheng and Hayes 2003a; Gordon, Luger et al. 2005; Zheng, Lu et al. 2005; Shogren-Knaak, Ishii et al. 2006; Kan

and Hayes 2007; Robinson, An et al. 2008; Wang and Hayes 2008; Kan, Caterino et al. 2009) is facilitated by bridging of tails between adjacent NCPs, which is diminished upon removal of the tails or quenching of the basic tail charges.

These two factors, salt and tails, would be discussed further below in NCP aggregation and nucleosome array folding respectively.

1.5.1 NCP Aggregation

Livolant and colleagues analyzed the phase behavior of NCP solutions in the presence of cations and gave information on the mechanisms involved in the NCP interactions (Leforestier and Livolant 1997; Raspaud, Chaperon et al. 1999; Leforestier, Dubochet et al. 2001; Raspaud, Pitard et al. 2001; Mangelot, Leforestier et al. 2002; Mangelot, Raspaud et al. 2002; Mangelot, Leforestier et al. 2003b; Mangelot, Leforestier et al. 2003a; Bertin, Leforestier et al. 2004; Bertin, Durand et al. 2007; Bertin, Mangelot et al. 2007; Bertin, Renouard et al. 2007). They produced isolated NCP from selective digestion of H1-depleted calf thymus chromatin with micrococcal nuclease to digest DNA and further obtained tailless NCP with trypsin.

Cation and Aggregation

It was shown that cation-induced NCP aggregation displayed a behavior similar to other polyelectrolyte systems such as DNA (Raspaud, Chaperon et al. 1999). Aggregation of NCPs was found to be dependent on the valence and nature of cations present in solution (Raspaud, Chaperon et al. 1999; de Frutos, Raspaud et al. 2001; Mangelot, Leforestier et al. 2002; Bertin, Mangelot et al. 2007). A very small concentration of divalent ions of merely a few mM could induce aggregation of NCP.

In (Mangelot, Leforestier et al. 2002), the effects of salt were explored in further detail. It was suggested that the tails are “extended” upon high concentration of salt, in contrast to the “compact” form at low salt concentration. It was concluded that the histone tails mediate NCP-NCP contacts above a critical monovalent salt concentration for its extension outside the core and therefore its contacts with other neighboring nucleosomes (Mangelot, Raspaud et al. 2002; Bertin, Leforestier et al. 2004; Bertin, Durand et al. 2007; Bertin, Renouard et al. 2007), yet this view would be questioned and justified in this work below in “4.3.1 Mechanism of Tail Bridging and NCP-NCP attraction”, and published work, (Yang, Lyubartsev et al. 2009).

The shortcoming from Livolant group and some other groups' approach (Ausio 2000; de Frutos, Raspaud et al. 2001; Wang, He et al. 2001; Hansen 2002; Mangelot, Leforestier et al. 2003a; Mangelot, Leforestier et al. 2003b; Bertin, Durand et al. 2007) (and references cited therein) is the use of cell-extracted particles obtained by nuclease treatment of the *in vivo* chromatin, except later work (Bertin, Durand et al. 2007; Bertin, Renouard et al. 2007). These NCPs contain DNA heterogeneous in size and carry an unknown number of post-translational modifications on the histone tails, in particular acetylations (Jiang, Smith et al. 2007; Pesavento, Bullock et al. 2008). Hence, the detailed mechanism involving tail and cation in nucleosome-nucleosome interactions in ordered precipitated NCP systems remains unclear; and, how this interaction is influenced by histone tail acetylations remains unexplored.

Tail Modification and Aggregation

Using well-defined nucleosome core particles prepared by recombinant methods can overcome the abovementioned shortcomings. One of such work is (Liu, Lu et al. 2011) where five types of reconstituted NCP (wild-type unmodified, globular tailless, H4 histone tail having Lysine (K) acetylations, Lysine-to-glutamine mutations at positions K5, K8, K12 and K16, and single H4K16 acetylation) have been studied for aggregation upon addition of cations (Mg^{2+} , Ca^{2+} , Cobalt(III)Hexamine³⁺, Spermidine³⁺ and Spermine⁴⁺). NCP aggregation is found to be sensitive to the position and the nature of the H4 tail modification. The tetra-acetylation in the H4 tail significantly weakens the nucleosome-nucleosome interaction, while the H4 K→Q tetra-mutation displays a more modest effect. The single H4K16 acetylation also weakens the self-association of NCPs, which reflects the specific role of H4K16 in the nucleosome-nucleosome stacking. Tailless NCPs can aggregate in the presence of oligocations, which indicates that attraction also occurs by tail-independent nucleosome-nucleosome interaction and DNA-DNA attraction in the presence of cations. We discuss its relevance to our work in “4.3.2 Relevance of the NCP Simulation to Experiment” and “7.2.1 Experimental Results” below.

In summary, isolated NCPs display experimentally observed salt-induced aggregation, depending on the concentration and valence of salt present in solution, which is a similar polyelectrolyte behavior as DNA condensation; tails modifications alter such behavior with less attraction arising from reduction of tail charges. Chromatin fibers quantitatively share similar properties because nucleosome Array-Array aggregation is related to NCP-NCP attraction.

1.5.2 Nucleosome Array Folding

Raising monovalent cation concentration in a solution can induce condensation or compaction of an extended array of nucleosomes into a form resembling a contacting zigzag structure (Thoma, Koller et al. 1979; van Holde 1988; Widom 1989). Furthermore, even in the absence of linker histones, addition of small amounts of divalent cations (i.e. Mg^{2+}) can induce further compaction of oligonucleosomes to a structure with the same hydrodynamic shape as a fully condensed $\sim 30nm$ chromatin fiber (Hansen, Ausio et al. 1989; Hansen 2002). At higher salt concentrations, (2–5mM $MgCl_2$) arrays oligomerize into tertiary chromatin structures. The effects of salt of higher valence are less studied.

Experiments with Recombinant Nucleosome

There is a growing number of works performed on precisely defined chromatin and nucleosomes prepared by recombinant methods: Histones and DNA components are individually produced by over-expression in bacteria (Luger, Rechsteiner et al. 1999; Dyer, Edayathumangalam et al. 2004; Huynh, Robinson et al. 2005) and can be assembled *in vitro* into pure NCPs and well-defined segments of chromatin arrays² in large quantity. Using these objects, determination of the contributions of individual tails and specific alterations within them (acetylation, phosphorylation, mutations etc.) to the self-association of chromatin and NCPs is possible.

Analytical ultracentrifugation (AUC) sedimentation velocity measurement is the standard experiment for detecting the degree of folding. In (Dorigo, Schalch et al. 2003), the sedimentation coefficient for a wild-type 12-mer array (here denoted 12-177-601 array, having the “601” high affinity nucleosome positioning sequence (Lowary and Widom 1998)) changes continuously from about 35S in low monovalent salt buffer solution to

² A template containing 12 tandem repeats of the 177 bp DNA sequence is available. It consists of the 147 bp “601” (Lowary and Widom 1998) positioning sequence flanked by two 15 bp DNA linkers (Schalch 2004) reconstituted with recombinant histone octamers to a 12-mer chromatin array, enables the production of large quantities (10–100 mg) of such a model chromatin array system.

about 53S at 1 mM MgCl_2 , implying higher divalent salt concentration induces aggregation. The general behavior of Mg^{2+} -induced chromatin condensation suggests a dominating contribution of electrostatic interactions to chromatin compaction. The authors also showed that among all four kinds of histone tails, the H4 tails are essential to the formation of fully compact chromatin. Noteworthy, one of the H4 K16-R23 regions (of the two H4 tails) is clearly involved in NCP-NCP stacking since it is located between the two NCPs in practically all reported NCP crystal structures. H3 and H4 tails (Dorigo, Schalch et al. 2003; Bertin, Renouard et al. 2007; Kan, Lu et al. 2007; Kan, Caterino et al. 2009; McBryant, Klonoski et al. 2009) and specifically H4K16 (Shogren-Knaak, Ishii et al. 2006; Robinson, An et al. 2008; Allahverdi, Yang et al. 2011) have been found critical for intramolecular folding of individual nucleosome arrays.

Another two reports of work performed on 5S nucleosome arrays (Gordon, Luger et al. 2005; Zheng, Lu et al. 2005) are in accordance with the present theoretical picture of tail mediated nucleosomal interactions. Using radiolabeled H3 tails, Zheng et al. (Zheng, Lu et al. 2005) found that, in low to moderate monovalent salt, the tails are associated with the DNA of its own nucleosome core in an extended array, whereas upon addition of Mg^{2+} , the tails interact with the DNA of the other nucleosomes in the condensed nucleosome secondary and tertiary arrays. Gordon et al. (Gordon, Luger et al. 2005) enumerate all possible combinations of “tailless” nucleosome arrays and demonstrated the independent (with some exception for the H4 tail) and additive (related to the charge on the tails) character of the Mg^{2+} -induced chromatin inter-array oligomerization. This indicated that all NTDs of core histone proteins participate in the oligomerization process.

In (Korolev, Allahverdi et al. 2010), sedimentation velocity measurements were performed to monitor the folding provoked by addition of cations Na^+ , K^+ , Mg^{2+} , Ca^{2+} , CoHex^{3+} , Spermidine³⁺, and Spermine⁴⁺. Typical polyelectrolyte behavior was observed, with the critical concentration of cation needed to bring about maximal folding covering a range of almost five orders of magnitude, from 2 μM for Spermine⁴⁺ to 100 mM for Na^+ .

Acetylation and Array Folding

There are other published results indicating that Lysine acetylation or absence of tails destabilize NCPs or arrays in the solution state over self-association and promote the unfolding of compact nucleosome arrays (Pollard, Samuels et al. 1999; Shogren-Knaak, Ishii et al. 2006; Bertin, Renouard et al. 2007; Kan and Hayes 2007; Kan, Lu et al. 2007;

Wang and Hayes 2007; Allahverdi, Yang et al. 2011). However, there is only a limited amount of data available for systems with modifications of the individual histones, such as mutants lacking tails, containing only the globular part of the histone molecule.

1.5.3 Compaction in Biopolyelectrolyte and Colloidal System

Colloidal systems composed of negatively charged macroion with attached (or free) positively charged chain-like oligocations (similar to polyamines) also show experimental attraction behavior (Claesson, Poptoshev et al. 2005). Such systems have been studied with theoretical modeling and computer simulations, see (Miklavic, Woodward et al. 1990; Rescic and Linse 2000) and (Another model) Podgornik's (Podgornik 2003; Podgornik 2004). From these models, it is clear that attraction can occur not only due to mobile counterion fluctuation but also due to chain bridging which is driven by both electrostatic and entropic origins. The bridging is favored by entropy gaining, as tails not only stay on their own charged surface, but extend to the charged surfaces of other particles as well, and also turns electrostatic repulsion to attraction by neutralizing it and other particles. Furthermore, these flexible cations also exhibit attractive dynamic correlation behavior. The resemblance of colloidal systems to the nucleosome has enlightened the hypothesis of how nucleosome aggregation is mediated by histone tails.

1.6 Theoretical Modeling

Modeling is a process of crafting the levels of abstraction and approximation in structure and interaction, based on experiment results, to separately investigate different physical mechanisms and assess them, both qualitatively and quantitatively.

DNA is a highly charged polyelectrolyte and it is highly packed inside nucleosomes. The histone protein also carries a large amount of positive charges but overall, nucleosomes is electrostatic negative, surrounded by cation cloud. In a simplistic but not oversimplified summary of experimental findings, nucleosomes are held together by mechanisms involving tails and cations, once triggered electrostatically by change in salt and tail modifications, leading to change of its physical properties and hence conformation. The modification of tails causes changes to the charges on tails and cation distribution around array varies with different valence and concentrations. There is certainly electrostatics (van Holde 1988; Clark and Kimura 1990; van Holde and Zlatanova 1996) in the description of the physics of the nucleosome. It is also equivalently to re-state this as following, if a

model correctly describes electrostatic interactions, comparison of experiment and modeling data could be very revealing for the understanding of how different forces contribute to chromatin structure and dynamics.

1.6.1 Previous Models

Modeling of nucleosome fiber with computer simulation has been cultivated for many years. Langowski and co-workers have contributed many models in the fiber conformational dynamics. There are models developed by Wedemann and Langowski (Wedemann and Langowski 2002) and later by Wedemann, Rippe and co-workers (Kepper, Foethke et al. 2008; Stehr, Kepper et al. 2008), (Kepper, Foethke et al. 2008); without explicit description of flexible tails and cations, their models use Debye-Hückel theory (define below in “3.2.4 Chromatin Array Modeling in Debye-Hückel Approximation”) for Linker DNA-Linker DNA interaction and effective “nucleosome-nucleosome interaction”. The later potentials, despite its name, captures mainly electrostatic interactions, and also the contributions from solvent and hydrophobic effects, hydrogen bonding and tail-mediated interactions. The separated treatment of DNA-DNA interaction and internucleosomal interaction is advantageous. The authors highlighted the importance of nucleosome shape and the magnitude of nucleosome-nucleosome potential to the folded chromatin structure. Considering the scale of the simulation systems, approximation in these mechanical models is justified at given salt conditions.

Following is a list of recent models, which have included description of the histone tails in coarse-grained models of the NCP and nucleosome arrays:

- 1) Muhlbacher: (Muhlbacher 2004; Muhlbacher, Holm et al. 2006; Muhlbacher, Schiessel et al. 2006)
- 2) Schlick and co-workers group:
 - a) Preliminary model: NCP modeled as flat cylinder (Beard and Schlick 2001)
 - b) Updated model with discrete charges on irregular surface (Sun, Zhang et al. 2005), method is described in (Zhang, Beard et al. 2003).
 - c) Updated model with flexible tails, linker DNA and histone H1.
 - d) Model development and monovalent salt (Arya and Schlick 2006; Arya, Zhang et al. 2006)
 - e) Model with divalent salt (Arya and Schlick 2009; Grigoryev, Arya et al. 2009)
- 3) Coarse-grain NCP model with flexible tails and explicit ions:

- a) Preceding work (Korolev, Lyubartsev et al. 2006) and
- b) The Sphere model to model tail modification (Yang, Lyubartsev et al. 2009)
- c) The Sphere-Bead model for NCP and nucleosome array (with linker DNA) (Korolev, Allahverdi et al. 2010; Liu, Lu et al. 2011).

1.6.1.1 Muhlbacher

Muhlbacher et al. (Muhlbacher, Schiessel et al. 2006) built a model with flexible positively charged chains grafted on the negatively charged surface of a sphere. Their di-nucleosome MD simulation result shows that tail-bridging effects contribute to the attractive NCP-NCP interaction measured in the experiment, more likely than “charge patch” mechanism. Another important finding is that attraction between two eight-tail colloids can be tuned by changing the fraction of the charged monomers on the tails. DH has been used to treat electrostatic interactions.

1.6.1.2 Schlick and Co-workers

Schlick and colleagues started to build the model having a well-justified idea to simplify extremely complex experimental system of huge number of atoms (chromatin array of sufficient length in diluted salt water solution) with a “meso-scale” model which tries to take into account the major forces acting in the system.

The nucleosome charged irregular surface is modeled with discrete normalized charges set by the DiSCO algorithm (Zhang, Beard et al. 2003) placed on the surface to reproduce the electrostatic field. The electrostatic interactions are treated with a mean field Debye-Hückel approximation. Tails are fixed in (Sun, Zhang et al. 2005) and became flexible in later models (Arya and Schlick 2006; Arya, Zhang et al. 2006) and (Arya and Schlick 2009; Grigoryev, Arya et al. 2009).

With the series of work cited above, they have presented detailed models, which were applied to theoretical studies of the salt-induced chromatin folding by using a DH approach to describe the effects of monovalent salt. Use of DH screening potentials may be justified in modeling systems with monovalent salt only and do well at reproducing the experimental folding behavior.

Considering the modeling process, the charge normalization removes the major physical contribution from the ligand-DNA interaction namely a decisive gain in entropy due to

the release of the counterions upon ligand binding. Therefore, the crucial component of the chromatin dynamics, dissociation/binding of the histone tails to DNA and to the other histones, e.g. to the acidic patches of the H2A globular domain, is described inadequately if fixed and reduced “normalized” charges are used in the coarse-grained model. Although for the case of monovalent salt their simulation results produce good agreement with experiment this might be a result of mutual cancellation of entropic loss/gain in dissociation/binding events. When cations of higher charge (Mg^{2+} etc.) come into play the model fails completely.

In (Sun, Zhang et al. 2005), the “fixed” protruding tails could artificially enhance the internucleosome interactions in the simulation when compared to a “flexible” variant of the model. And somehow, such configuration has enabled the model to achieve equilibrium with MC simulation with short computation time. Their conclusion on H3 tail’s superior significance over other tails is also arguable because H3 is by chance the longest tail in the X-ray crystal structure. Neutralizing it will apparently weaken the internucleosome interactions.

Recent works (Arya and Schlick 2009; Grigoryev, Arya et al. 2009) includes divalent ions and linker histone H1. To take into account the inadequacy of description of divalent ions in the DH theory, they used a “phenomenological approach towards modeling the effect of 1 mM Mg^{2+} ”. In detail, this includes, first, setting the Debye screening to 2.5 nm^{-1} for interactions between linker DNA, the inverse of the diameter of DNA, and second, reducing the persistence length of the linker DNA from 50nm to 30nm. In our view, this is a wrong way to fix the model to produce “correct” result, ignoring the inappropriate use of DH theory and ignoring the valence of ions. In addition, bridging mechanism of the tails is understood incorrectly, merely following (Mangenot, Leforestier et al. 2002), questioned and justified in this work below in “4.3.1 Mechanism of Tail Bridging and NCP-NCP attraction”.

1.6.1.3 The “Spider” Model

Korolev *et al* (Korolev, Lyubartsev et al. 2006) designed a model with flexible tails and explicit cations. Its NCP model is composed of 89 beads, mimicking the real NCP structure, shown in **Figure 1-5(a)**. One bead is in the center carrying -236e charge and other 88 beads (each with +1e) are in the form of eight tails. Beads on the tails are connected by harmonic bond. Ions and NCP models are placed into a cube with periodic

boundary conditions, shown in **Figure 1-6**. Dielectric continuum model (with $\epsilon=78.0$) containing explicit mobile counterions and various amounts of added ions describe the salt conditions performed in the experiments. NVT MD simulation was carried out to generate a canonical ensemble, producing the same kind of results as Monte-Carlo simulations. Ewald summation, which counts infinite, repeating periodic interactions, is employed. It incorporates ion-ion correlation attraction mechanisms into the model, similar to full-atomic simulation. The price for such inclusion is the high demand for computational costly algorithm for high-accuracy representation of electrostatic forces. Four kinds of salt conditions (no salt, low Mg, High K, and Low K) were simulated, and the results are in good agreement with experiment data in salt-dependent aggregation as well as in tails' role of bridging.

Figure 1-5 NCP model in (Korolev, Lyubartsev et al. 2006) Left) Atomic structure of NCP with tails extended, adopted from (Wolffe and Hayes 1999). Right) Top view of NCP model (1+88 beads).

Figure 1-6 Simulation system built from NCP models (Korolev, Lyubartsev et al. 2006)

1.6.2 Our Approach

A common drawback of most models, with the exception of (Korolev, Lyubartsev et al. 2006), is that the electrostatic interactions and the presence of salt is treated highly approximately, either with effective potentials or with a Debye-Hückel (DH) potential, sometimes including a mean field Poisson-Boltzmann (PB) model. Debye-Hückel treatment cannot describe the distinct effects of divalent ions, where all effects of ion valence and salt concentration are contained in the single descriptor, ionic strength.

It is well established that for polyelectrolytes in general and for DNA in particular, fluctuation induced dynamic correlation of cations shared by different poly-anions gives rise to attractive electrostatic force contributions that explains experimental aggregation (Oosawa 1968; Ray and Manning 1997; Gelbart, Bruinsma et al. 2000). The effect is more pronounced in the presence of multivalent ions. To model the compaction of nucleosomes, particularly in the presence of multivalent ions, which are exceptionally better in inducing compaction, should incorporate mobile ions.

Also, experimental results on nucleosome arrays indicate that tails relocate from mainly intra-nucleosome interactions with its own core particle in the extended array at low salt, to primarily bridging-like inter-nucleosomal interactions upon formation of condensed secondary and tertiary chromatin aggregates (Dorigo, Schalch et al. 2003; Zheng, Lu et al. 2005). This so-called “tail-bridging” mechanism is also enlightened from colloidal system for its entropic origin. A model with flexible charged tails is necessary.

Overall, mobile cation and flexible tails ensure an uncompromised treatment to electrostatics as well as address contributions from ion competition, dynamic ion-ion correlation (Oosawa 1968) due to both fluctuations in tail configuration and in mobile cation distribution, and the influence of entropic tail bridging (Podgornik and Licer 2006). They shall be simultaneously taken into consideration. This decision led us to a model of very simple structural description, but very expensive but worthy electrostatic computation.

In this work, we further develop the “Spider” model to include description of tail modifications, and to have more details of the shape and charges in the nucleosome. We progressively developed two models for the NCP. The first is the Sphere model “2 The Sphere Model” (pg. 20), followed by the Sphere-Bead model “3 The Sphere-Bead Model”

(pg. 35). The Sphere model incorporate tail modification on top of “spider” model where beads of neutral charge or negative charge and of different sizes represent acetylated and phosphorylated histone tail amino acid residues. The Sphere-Bead model features explicit modeling of DNA wrapped around octamer and linking NCP. Both models have been used to simulate NCP in salt solution and the Sphere-Bead model has been used to simulate a nucleosome array in salt solution.

Although we can establish typical polyelectrolyte properties of nucleosome systems, which imply the importance of the amount and nature of salt present, previous theoretical description of interactions in chromatin lags behind the knowledge gathered in experimental studies. In particular, following two points are unclear:

- 1) Regarding salts, in particular, when it comes to the description of the hallmark Mg^{2+} induced NCP and chromatin array compaction and aggregation experiment, the majority of previous chromatin and NCP models fail to describe the electrostatics at a level necessary to properly account for physics of the systems.
- 2) Regarding tails, the underlying mechanism that leads to attraction between NCPs and subsequent aggregation is not known in detail in how flexible tail mediates with interplaying of different concentration and valance of salts, and how tail modifications alter such mechanisms.

For our quest on elucidating mechanisms of nucleosome aggregation or array folding in relation to salt and tails, we performed simulations with wild-type and modified (acetylated/phosphorylated and tailless) NCP system “4 NCP Simulation with The Sphere Model” (pg. 49), “6 Array Simulation with The Sphere-Bead Model” (pg. 65), and “7 Tailless NCP Simulation with The Sphere-Bead Model” (pg. 92) and nucleosome array system (“6 Array Simulation with The Sphere-Bead Model”, pg. 65) subjected to different concentration (in “titration” fashion) and to salts of different valences.

The original work described in this thesis has been published in (Yang, Lyubartsev et al. 2009; Korolev, Allahverdi et al. 2010; Liu, Lu et al. 2011).

2 The Sphere Model

The Sphere model first appeared in (Korolev, Lyubartsev et al. 2006). We extended it with parameters describing modified tail beads and also ported the model to another simulation software package, called ESPResSo (Limbach, Arnold et al. 2006). The model is described below.

The atomic model from X-ray crystal structure at 1.9Å resolution (Davey, Sargent et al. 2002) (PDB ID: 1KX5) is taken as a reference for the model, on the size, geometric structure and charge of the beads. There are a few proposed positions where the flexible part shall begin, albeit they are of tiny differences in a few amino acids. In prior work (Korolev, Lyubartsev et al. 2006), the amino acid definition is referenced from yeast nucleosome described in (Hansen, Tse et al. 1998) but we opted for the tail definition and locations of acetylation and phosphorylation sites from experimental work on the signaling function through histone tail modification (Cheung, Allis et al. 2000) (reproduced in **Figure 2-1**). In this later publication (Cheung, Allis et al. 2000) which describes the human nucleosome, the H2B tail has a longer sequence/higher +charge (+13e $\rightarrow$ +14e) and H3 tail has a shorter sequence/lower +charge (+12e $\rightarrow$ +11e) than in (Hansen, Tse et al. 1998), shown in **Table 2-1**.

Figure 2-1 Core histones terminal tails as signaling platforms. Red lollipops represent known phosphorylation sites; green flags represent known acetylation sites (Thorne et al., 1990); red Serine and Histidine residues represent potential phosphorylation sites. Blue Ac-NHs represent N-terminal acetylation of H2A and H4 tails. (Reproduced from Figure 1 in (Cheung, Allis et al. 2000))

Table 2-1 Histone tail amino acid sequence and charges depicted in (Cheung, Allis et al. 2000). Arginine and Lysine are in **bold**. The charges are calculated based on protonated amino acids at pH 6.0.

	Histone Amino Acid Sequence	Net charge
H2A	H3N+-SGRG K QG GK ARAKAKSRSSRAGL-gH2A	+9
H2B	H3N+-PEPSKSAPAP K KG S KKAITKAQ K KG K RRSRK-gH2B	+14
H3	H3N+-ART K QTARKSTGGKAPRKQLATKAARKSAP-gH3	+11
H4	H3N+-SGRG K GG K GLG K GGAKRRHRKVLR-gH4A	+10
Total		+88

2.1 Model Setup

Figure 2-2 Scheme of the Sphere Model

The Sphere model approximates the recombinant NCP (rNCP) with a spherical particle of effective radius 4.85 nm with a net charge $-236e$ combined with 8 strings of linearly connected $+1e$ charges of effective radius 0.25 nm and bond length 0.7 nm, which represent the histone tails, shown in **Figure 2-2**. The charge of the central particle reflects the electrostatic balance between the 147bps of the double-stranded DNA (charge $-294e$) and the net positive charge ($+58e$) of the amino acids comprising the globular domains of the histones. The number of charged groups in each tail is set according to the charges of the real histone tails (see **Table 2-1**). The total charge of histone tails is $+88e$, which, together with the charge of the central unit, yields $-148e$ for the total charge of the NCP. The distance between charged groups of the histone tails (0.7 nm) is obtained from

extended conformation of the polypeptide chain and the approximation of even distribution of the Lys⁺/Arg⁺ amino acids.

2.1.1 Geometry

The exit points of the tails are determined for corresponding C α atoms of the histone tails in the atomic structure of the NCP determined by X-ray crystallography (Luger, Mader et al. 1997; Wolffe and Hayes 1999) (see Figure 2 in ref. (Wolffe and Hayes 1999)). The Sphere model approximates the crystal structure with “idealized symmetry”: the two histone of the same kind are symmetric about x-, y-, and z-axis and the origin of center sphere. **Figure 2-3** shows the geometry of the connections between the first (closest to the negative core) charged group of the histone tails to the center of the core particle and to each other. **Table 2-2** tabulates the coordinates.

Table 2-2 Coordinates (in Å) of the first bead of the histone tails in the NCP model.

Tail	X	Y	Z
H3	55.0	0.0	0.0
H3	-55.0	0.0	0.0
H2B	0.0	55.0	0.0
H2B	0.0	-55.0	0.0
H4	34.64	-34.64	25.0
H4	-34.64	34.64	-25.0
H2A	-42.43	-24.5	25.0
H2A	-42.43	-42.43	-25.0

In **Figure 2-3**, three schemes on the left-hand side are views of NCP with all charges in the histone tails. On the right-hand side, three lines schemes show the connection of the histone tails to the negatively charged core and to each other (first 2 beads of each tail closest to the core are shown as blue spheres). The central core particles are in the same projections for the two schemes on two sides. The middle schemes show view from the top along z-axis; top and bottom schemes display projections resulting from respective counter- and clockwise 90° rotations of the scheme in the middle.

The center of the core particle is drawn as red ball. The coordinate system has its origin in the center of the core particle with exit points of the H2B and H3 tails forming x, y-plane

($z=0$). The distance from the first charged group of the tails to the center of the core is set to 55 Å for all tails. The points of exits for the 4 “histone tails” (H2B/1, H2B/2, H3/1, H3/2), which protrude through the gyres of the DNA superhelix in the real NCP, are set in the x, y -plane ($z=0$) with 90° between them. The exit points of the H4/1 and H4/2 histone tails are placed 25 Å below and above the x, y -plane, opposite to each other and perpendicular to the dyad axis. Finally, the H2A/1 and H2A/2 tails exit the core particle 25 Å below and above the x, y -plane and are positioned between H2B/1 and H2B/2 tails.

Figure 2-3 Schematic drawing for tail start bead location (distance in Å).

2.1.2 Variants of Tail

The location for the modifications corresponded to the most conserved and frequently observed in human eurochromatin (Cheung, Allis et al. 2000; Turner 2002) are shown in **Figure 2-1** and summarized below:

- H2A: The N-terminal serine (Ser1) is acetylated and phosphorylated and modeled as two particles; one more acetylation site corresponds to the modified Lys5.

- H2B: It contains two uncharged particles closer to its end that corresponds to the acetylation of Lys12 and Lys15; phosphorylation site is located closer to the central particle and matches modification of Ser32.
- H3: It contains a cluster of modifications in the middle part of the tail, which correspond acetylation of Lys9 and Lys14 and phosphorylation of Ser10.
- H4: It carries a phosphorylation site at Ser1. Two acetylation sites are in the middle of the tails matching modifications of Lys5 and Lys12.

Based on these modification sites locations, we introduce three variants of the Sphere model, “recombinant” (rNCP), “acetylated” (aNCP) and “phosphorylated+acetylated” (paNCP). aNCP and paNCP variants mimic the nucleosome with significant covalent modifications in the tails (typically associated with transcriptionally active chromatin, euchromatin).

rNCP variant represents the NCP with unmodified fully charged histone tails; aNCP variant mimics the situation when two positively charged amino acid in each of the histone tails have been “acetylated” which means that their charges are switched off to zero and the radius increased from 2 to 3Å; paNCP variant addresses the situation of “extreme” modification of the NCP when in addition to the acetylation of 2 Lysines in the each tail, there is one phosphorylated serine on each of the histone tails, which means one additional particle with charge -2e and radius 4Å is inserted (“phosphorylation”). As a result, the net positive charge of the tails is reduced from +88e to +72e (aNCP), and to +56e (paNCP) and the tails acquire more complex charge distribution. The total charge of histone tails in rNCP variant is +88e, which yields the total charge of rNCP -148e, together with the charge of the central unit. aNCP variant has net charge -164e, and paNCP variant -180e.

In **Figure 2-4**, we show those model variants in detail. The positions of modification are shown in the graphic form and listed beside. Blue beads of +1e charge, green beads of no charge, and red beads of -2e charge represent unmodified, acetylated and phosphorylated amino acid bead positions respectively.

For aNCP variant, the basic variant is to have neutralized sites matching the sites of Lysine acetylation frequently observed *in vivo* (top section in **Figure 2-4**). There are three more variants. First variant is to have acetylation sites evenly distributed (named as “even” and presented in the middle section of **Figure 2-4**). Two more variants are designed with two

uncharged beads placed at the one of the two ends of the tails (named as “head” and ”root”; not shown in the figure).

rNCP	H2A(9)		
	H2B(14)		
	H3(11)		
	H4(10)		
aNCP	H2A(9)		N-ter, K5 (1,3)
	H2B(14)		K12,K15 (3,4)
	H3(11)		K9, K14 (5,6)
	H4(10)		K5, K12 (3,5)
	H2A(9)		Evenly Distributed
	H2B(14)		
	H3(11)		
	H4(10)		
paNCP	H2A(10)		Ser1 (1)
	H2B(15)		Ser32 (13)
	H3(12)		Ser10 (6)
	H4(11)		Ser1 (2)

● 1e
● 0
● -2e

Figure 2-4 Charges on the tails in variants of the Sphere model

2.2 Dynamics and Simulation

2.2.1 Interaction

The interaction potential of the coarse-grained model consisted of three parts: electrostatic, short-range and bond potentials. The electrostatic interaction is defined as sum of Coulombic potentials from all charged particle pairs in the system, in dielectric media with Bjerrum length $\lambda_B = 7.13 \text{ \AA}$, corresponding to water's dielectric constant at $\epsilon = 78$ at $25C^\circ$ with $\lambda_B = \frac{e^2}{4\pi\epsilon_r\epsilon_0k_B T}$. This long-range electrostatic potential is calculate by P3M method with optimal parameter determination (Deserno and Holm 1998a; Deserno and Holm 1998b). ESPResSo source code (Limbach, Arnold et al. 2006) has been modified to implement this type of potential. The short-range potential, acting between any pair of elements is:

$$U_{short}(r) = \begin{cases} kT \left(\frac{\sigma}{r - r_{ij}} \right)^9 & r > r_{ij} \\ \infty & r \leq r_{ij} \end{cases} \quad \text{Eq. (1)}$$

where r is the distance between elements i and j , $r_{ij} = r_i + r_j$ is the sum of their hard core radii, and $\sigma = 3 \text{ \AA}$, which is a parameter offering effective thickness of the soft repulsion potential. Similar interaction potential has been used in Monte-Carlo simulation of ions around DNA in (Montoro and Abascal 1995; Stehr, Kepper et al. 2008). For hard-core radii, we used $r_{core} = 47 \text{ \AA}$ for NCP central unit and $r_{tail} = 1 \text{ \AA}$ for +1e particles in the histone tail, K^+ and Cl^- ions. To account for larger sizes of hydrated Mg^{2+} ion, acetylated Lysine and phosphate group in the histone tail corresponding hard-core sizes are assigned $r = 2, 3 \text{ and } 4 \text{ \AA}$ respectively. The potential in Eq. (1) has effective interaction radius $r_i + r_j + \sigma$, which is the sum of effective radii of each component $(r_i + \sigma/2) + (r_j + \sigma/2)$. So the effective radius of NCP core is $48.5 \text{ \AA}$ and single-charged ion/+1e tail beads $2.5 \text{ \AA}$ and effective radius of Mg^{2+} ion, acetylated Lysine and phosphate group are $3.5 \text{ \AA}$, $4.5 \text{ \AA}$ and $5.5 \text{ \AA}$ respectively.

From the ion-ion RDF plotted from full-atomic MD simulations with ions in water, we found those effective radii are: Cl^-/K^+ $2.5 \text{ \AA}$, Mg^{2+} $4\text{-}5 \text{ \AA}$. Taking these values as reference, the values used in the coarse-grained simulation are decided as a result of adequately differentiating various types of particles with approximation to their physical sizes.

The bond potential acting between connected particles is of such form:

$$U_{bond}(r) = kT \left(\frac{r - r_{eq}}{D} \right) \quad \text{Eq. (2)}$$

where characteristic fluctuation $D = 1 \text{ \AA}$, force constant $k_{bond} = 25 \text{ kT per \AA}^2$. Such force constant is set to allow the distance for slight changes only. Equilibrium distances are $r_{eq} = 7 \text{ \AA}$ for neighboring Lysines (acetylated and positively charged) and $r_{eq} = 5 \text{ \AA}$ between phosphate group and closest bead in the tail, respectively.

Same kind of potential acts between the first beads of the histone tails and the core as distance constraints. They are introduced to preserve the initial geometry of the histone tails' exits from the core particle and making the core and 8 first beads of tails a rigid body. The number of the constraints is $3 \times (8 + 1) DOF - 6 = 21$.

2.2.2 Dynamics

Since the solvent is not included explicitly, the molecular dynamics should be considered as a tool to generate the canonical ensemble, producing the same kind of results as Monte-Carlo simulations. For strongly charged polyelectrolyte systems the molecular dynamics method is a more efficient way to sample the configuration space compared to the standard Metropolis Monte-Carlo algorithm, which displays convergence problems for simulations of highly charged system with moving polyions. In the last two decades, constant-temperature molecular dynamics with thermostats of different kinds are often used in polyelectrolyte and other coarse-grained simulations (Stevens and Kreker 1995; Limbach and Holm 2001; Stevens 2004; Lyubartsev 2005).

We found the Langevin MD thermostat within the ESPResSo package (Limbach, Arnold et al. 2006) to be very effective. In this thermostat, the total force acting on each particle, is a sum of conservative force from the expression of the potential energy, friction force $-gv_i$ (where g is friction coefficient and v_i is particle velocity), and random force, the intensity of which is determined by the fluctuation-dissipation theorem (Grest and Kremer 1986). We set Langevin friction coefficient "gamma" $g=0.01$.

Mass is assigned in reduced units and is set to 1u for histone tail beads, 1u for ions, and 100u for the NCP core. The ratios of the masses of the NCP core to the masses of tail beads and ions was made substantially smaller than the real one with the purpose of providing faster sampling of the configuration space.

With such a low value and low mass, this parameter does not account for real friction of moving particles through the solvent, but provides only a small correction to the Newtonian equations of motion keeping the temperature constant. The dynamics in such a simulation is artificially accelerated, due to the low friction and the low mass of NCPs and may be qualitatively reconstructed by scaling the time with some factor, but provides fast and efficient sampling. Since we are interested only in equilibrium configurations, we do not evaluate this factor and report only the number of MD steps made. However, it

should be pointed out, that the configurations generated in this way are not artificial, and do represent the equilibrium ensemble for sufficient convergence of the simulation runs.

Comparing to our model's predecessor described in (Korolev, Lyubartsev et al. 2006), our new implementation uses same NVT ensemble, fundamentally similar electrostatic algorithm (Ewald v.s. P3M), (Previous model v.s. our model, same for below), different dynamics (MD/Nóse-Hoover v.s. Langevin Dynamics), and units (actual unit v.s. reduced unit).

2.2.2.1 Reduced Units

Under the assumption that the chosen reduce mass unit corresponds to 100 a.u., the time step used in simulations is equivalent to 16fs. The relationship between reduced unit and standard unit is as following (verified in online reduced unit calculator: <http://www.pumma.nl/index.php/Manual/ReducedUnits>):

Given Boltzmann constant $k = 1$, Unit mass $m = 100 \text{ a.u.}$, Temperature $T = 1$, unit length $l = 1 \text{ \AA}$ and $E = 2.4 \text{ kJ/mol}$ at 300K, we have the time unit in our simulation:

$$E = \frac{1}{2}mv^2 = \frac{1}{2}m\left(\frac{l}{t}\right)^2$$

$$[t] = \sqrt{\frac{[m][l]^2}{[E]}}$$

$$[t] = \sqrt{\frac{(100)(1.66054 \times 10^{-27})(10^{-10})^2}{\frac{2.4 \times 10^3}{N_A}}}$$

$$[t] = 0.645 \times 10^{-12} = 0.645 \text{ ps} = 1 \text{ t.u.}$$

As we set the time step in simulation as 0.25, it is equivalent to $0.25 \text{ t.u.} = 16.125 \text{ fs}$, which is in the order of 10fs. Trajectory is saved in the interval of 100 time steps, which is about 1ps. Simulation lasts for about $200\text{-}600 \times 10^5$ time steps, equivalent to 200-600ns in simulation with non-reduced unit.

2.2.3 Systems

2.2.3.1 NCP Concentration

Figure 2-5 Simulation box with the Sphere model, so-called “ten spiders in a can”.

The cubic simulation cell of side length 40 nm contains 10 NCPs, subject to periodic boundary condition. **Figure 2-5** shows 10 NCP in well-ordered positions inside simulation box. NCP concentration is about 260 μM , corresponding to the higher limit of nucleosome concentration inside eukaryotic cell nucleus in the interphase estimated by combined fluorescent correlation spectroscopy and confocal imaging (Weidemann, Wachsmuth et al. 2003).

During simulation, each NCP occupies a spherical space of radius 6-7nm, which is translated to a volume of $\sim 14.3 \times 10^3 \text{ nm}^3$, or 22.4% of the box volume for ten NCPs. Artifacts could be formed when there are insufficient inter-NCP interactions before aggregation. We prevent this by having well-separated (every inter-core distance $> 15 \text{ nm}$) randomized starting locations for NCPs, which is an improvement from previous work (Korolev, Lyubartsev et al. 2006).

2.2.3.2 Non-titration Groups

We re-performed simulations for the group of four systems with rNCP studied in (Korolev, Lyubartsev et al. 2006). They differ from each other by salt type and concentration and are listed as following.

- 1) Salt free: 1480 K^+ ions to neutralize the NCP charge;
- 2) Low salt: 320 KCl (8.3 mM) + 1480 K^+ ;
- 3) Normal salt: 2520 KCl (65.4 mM) + 1480 K^+ ;

4) Magnesium salt: 160 MgCl_2 (4.15 mM) + 740 Mg^{2+} .

Salt-free and low-salt systems contain concentrations of K^+ (or Na^+) ions are used in many *in vitro* studies of the NCP. ‘Normal salt’ system represents conditions of physiological ionic strength with monovalent cations; ‘Magnesium salt’ system matches the concentration typical for *in vitro* studies and observed also *in vivo* for NCP- Mg^{2+} studies.

With the inclusion of tail-modified variants, we have 12 more systems only with salt conditions 3) normal salt and 4) Mg salt because the other two salt conditions cannot induce much aggregation with rNCP. The 12 systems are: 2 with rNCP, 8 with aNCP (Lysines sites set to “native”, “even”, “root”, “head”) and 2 with paNCP. Additional cations were added to aNCP and paNCP systems to compensate for the increase of negative charge of the NCP due to neutralized or negated tail beads. The concentration of KCl salt in all systems with different NCP variants is 65.4 mM KCl, for 2520 KCl molecules; for systems with Mg^{2+} , the concentration of MgCl_2 is 4.15 mM, for 160 MgCl_2 molecules. These 12 systems are studied in two groups (with overlapping): the first group is rNCP/aNCP (“native”)/paNCP, six systems, where “native” means Lysine sites at *in vivo* locations; and the second group comprising of four variants of aNCP. These two groups are called “non-titration” groups, summarized in following table.

Table 2-3 Non-titration systems. *: Repeated simulations from (Korolev, Lyubartsev et al. 2006) and the names in the bracket are their previous names. ^: 6 simulations to study variants of NCP rNCP/aNCP/paNCP subjected to two salt conditions. #: 4 variants, “Lysine-site”, “even”, “head”, and “root”.

Name	NCP	K	Mg	Cl
*rNCP (Salt Free)	10	1480	0	0
*rNCP (Low Salt)		1480+320	0	320
*^rNCP-KCl (Normal Salt)	10	1480+2520	0	2520
*^rNCP-MgCl ₂ (Mg Salt)		0	740+160	320
^#aNCP-KCl		1640+2520	0	2520
^#aNCP-MgCl ₂		0	740+80+160	320
^paNCP-KCl		1800+2520	0	2520
^paNCP-MgCl ₂		0	740+160+160	320

2.2.3.3 Titration Groups

In studies of NCP and chromatin array systems, the standard experiment to monitor the condensation behavior is the titration by multivalent salt, in particular MgCl_2 (Hansen 2002) (and references therein) as well as recent works (Shogren-Knaak, Ishii et al. 2006; Bertin, Mangelot et al. 2007; Kan, Lu et al. 2007; Wang and Hayes 2007). To model such kind of experiment correctly in simulations, a very large system (large simulation cell) must be employed or preferably, simulations within a grand canonical ensemble in equilibrium with a defined bulk salt solution should be carried out (Lyubartsev and Nordenskiöld 1995). In order to qualitatively mimic the titration, we performed a group of simulations with systems comprising K^+ and Mg^{2+} in different ratio, yet keeping the total amount of charges the same.

We have another two group of simulations, called “titration” groups, one for rNCP and another for aNCP. paNCP is excluded because paNCP- MgCl_2 system shows very weak aggregation. These systems are named using the indication of the NCP variant and percentage of Mg^{2+} relative to the total positive charge of $\text{K}^+ + \text{Mg}^{2+}$ (+1480+50e for rNCP systems; +1640+50e for aNCP systems). The system also contains 50 Cl⁻. For example, the notation “aNCP-50Mg” is used for a system containing 422 Mg^{2+} and 846 K^+ (plus 10 aNCPs and 50 Cl⁻). In this way, the variation of Mg^{2+} content from 0 to 100% may be consider to qualitatively mimic a titration experiment whereby an NCP solution starting with a low KCl salt is titrated with gradually replacing K^+ with Mg^{2+} in the ratio of 2:1(K: Mg).

2.2.3.4 Single NCP Group

The last group of simulations is used to investigate the relationship between participation of the histone tails in NCP-NCP interaction and the salt-dependent dissociation of the tails from the core of the NCP. They are performed with a single NCP in a cubic simulation cell of side length 23nm with the presence of varying amount of KCl or MgCl_2 . The concentration of salt in the simulations with one NCP in the cell is varied from 5 mM to 200 mM of KCl or MgCl_2 .

2.2.3.5 Simulation Time

The deca-nucleosome systems were simulated for $(160-320 \times 10^5)$ time steps (non-titration system) and for (480×10^5) time steps or longer (titration systems). For single-NCP

systems, convergence is achieved in shorter trajectory, for 80×10^5 to 160×10^5 time steps. Configurations for analysis were collected after each 100 steps.

2.3 Analysis

Results are averaged from the final 160×10^5 time steps or longer, after achieving convergence in the NCP-NCP radial distribution functions (RDF) and External tail-Core RDF.

2.3.1 RDF

The RDF (radial distribution function) between particles of type A and B is defined in the following way:

$$g_{AB}(r) = \frac{\langle \rho_B(r) \rangle}{\langle \rho_B \rangle_{local}} = \frac{1}{\langle \rho_B \rangle_{local}} \frac{1}{N_A} \sum_{i \in A} \sum_{j \in B} \frac{\delta(r_{ij} - r)}{4\pi r^2}$$

with $\langle \rho_B(r) \rangle$ the particle density of type B at the distance r around particle of type A, and $\langle \rho_B \rangle_{local}$ the particle density of type B averaged over all spheres around particles A with radius r_{max} which is usually taken as half of the box length, and also averaging among all timeframes. In practice, the analysis divides the system into spherical slices (from r to $r + dr$) and a histogram of the number of particles in the slice along the radial distance, divided by the local density, becomes the RDF instead of δ -function. The minimum non-zero RDF value indicates the minimum distance between two types of particles; the peak or plateau in RDF indicates an interaction distance or distance range for two types of particles.

The analysis program included in GROMACS (van der Spoel, Lindahl et al. 2005a; van der Spoel, Lindahl et al. 2005b) is used for RDF calculation. Its source code has been examined and improvements were feedback to the development team³.

³ GROMACS mailing list thread:
<http://oldwww.gromacs.org/pipermail/gmx-developers/2007-May/002112.html>

In this work, we used various RDFs between different groups of atoms. Following is a list of them.

- 1) **Internal tail-Core RDF** averages distribution of tail beads relative to the center of their “host” NCP they are attached to, calculated from all NCPs in the system. The intensity of the peak in the RDF describes how tails dislocate from their NCP.
- 2) **External tail-Core RDF** averages distribution of the beads of the tails relative to cores, which they are not attached. The intensity of the peak on this RDF characterizes the “tail bridges” between NCPs.
- 3) **Core-Core RDF** shows how cores are distributed relative to each other. The formation of aggregates of NCPs or the folding of array can be characterized from the location and magnitude of the peak on Core-Core RDF.
- 4) **Ion-Core RDF, Ion-X RDF, Ion-Ion RDF etc.** They show how ions distribute around the core and other selected group of particles.

Not all RDFs necessarily converge to 1 because it is subjected to the distribution of the selected particles and occasionally due to scaling. For example, when cores and tails are aggregated, converging value could be 0 at the long-distance end.

2.3.2 Cumulative Number

To quantitatively describe RDF all containing one peak at similar radial distance, peak value is sufficient. With RDFs showing multiple peaks, it is better to compare cumulative number, which is an integration of RDF over a certain radial distance r . Cumulative number is able to smooth out the various peaks of the RDF, giving the averaged number of particles of type B found inside the Sphere of radius r centered at one particle of type A. The distance r to be integrated is determined usually on the right-hand side to the peaks.

2.3.3 Contacts per Core

Two NCPs are considered as being in contact if Core-Core distance between them is less than 12 nm, a distance based on the position of the right border of the first maximum of Core-Core RDF, shown in **Figure 4-3(a)** below. The number of contacts per core is

From: Yang Ye
Subject: [gmx-developers] proposed change to g_rdf with -nopbc
Date: 2007 May 3

defined as: $n = N_{pair}/N_{core}/N_{frame}$, where $N_{core} = 10$ and N_{pair} is the sum of contacts counted from one trajectory segment containing $N_{frame} = 4000$ frames.

2.3.4 Trajectory Output

XTC is a compressed trajectory file format, containing coordinates time and box vector information from simulations. We ported the implementation of XTC output in GROMACS (Lindahl, Hess et al. 2001) to ESPResSo (Limbach, Arnold et al. 2006). This also enables the use of GROMACS analysis tools to our trajectories.

3 The Sphere-Bead Model

We improved the Sphere model with the inclusion of DNA beads for both nucleosome DNA and linker DNA, adding more geometric details named as “Sphere-Bead model”. The Sphere-Bead model is used to simulate both discrete NCP and nucleosome array (“12-177” array).

In the Sphere-Bead model, the smaller spherical core approximating the shape of a cylinder with hemispherical caps on top and bottom, rather than flat surfaces. Connected negative charged DNA beads of 10Å diameters are on the jacket of the center sphere. In comparison, the Sphere model uses a sphere of diameter 97Å to represent both histone and DNA.

3.1 Model Setup

3.1.1 Modeling Single NCP: Charges, Sizes and Geometry

A central neutral spherical particle of effective radius 35 Å represents the globular domain of the histone octamer. 25 beads each of effective radius 10 Å models the DNA wrapped around the core. Each DNA bead model represents one 6 bp DNA fragment. The original -12e charge of each DNA bead is reduced to -9.44e assuming that all positive charges of the histone octamer core are to partially neutralize DNA (discussed below in “3.1.3 Adoption of Neutrally Charged Core”). 25 DNA beads carry -236e charge, which is equal to the sum of the charges on 147 bp DNA (-294e) and the globular part of the histone octamer (+58e). The DNA beads are separated by 20 Å, an approximate to 6bp×3.4(rise/bp) Å. They wrap around the core for about 1.75 turns, very close to the number of turns in high-resolution NCP model (PDB: 1KX5).

The first and the last DNA beads are placed correspondingly close to the “root” (the binding location to the core) of H3/1 histone tail and to the root of H3/2 tail respectively. In a local coordinate system, fixed at the NCP, the z-coordinates of the DNA beads from 1 to 9 all are -12 Å; whereas for the DNA beads from 15 to 25, $z = +12$ Å; DNA beads 10-14 bridge two segments by being equally positioned between -12 Å and +12 Å in z-direction. The x- and y-coordinates of the DNA beads follow the circle of radius 44.5 Å

with angle 26.2° between neighboring beads. The described combination of the bead sizes and positioning closely represents the crystallographic structure of the NCP (Luger, Mader et al. 1997; Davey, Sargent et al. 2002), shown in **Figure 3-1**.

Figure 3-1 The Sphere-Bead model of the NCP. Top view and front view of the Sphere-Bead NCP model illustrated for variant model aNCP, with two acetylation sites shown in larger sized green beads on each of the histone tails. The non-acetylated (recombinant) model shares the same geometry. The histone tails are modeled in the same way as the Sphere model with updated coordinates listed in **Table 3-1**.

Table 3-1 Coordinates (in Å) for the central particle, start/end DNA beads and the first beads of tails closest to the core in the Sphere-Bead NCP model.

Particle	X	Y	Z
Core	0.0	0.0	0.0
DNA (Start)	44.5	0	-12.0
DNA (End)	0	44.5	12.0
H3/1	48.0	0.0	0.0
H3/2	0.0	48.0	0.0
H2B/1	-40.0	0.0	0.0
H2B/2	0.0	-40.0	0.0
H4/1	31.0	-1.0	16.0
H4/2	-1.0	31.0	-16.0

H2A/1	-34.0	-1.0	21.0
H2A/2	-1.0	-34.0	-21.0

3.1.2 Modeling Chromatin Array

To model the nucleosome array containing 12 nucleosomes with nucleosome repeat length 177 bp, twelve NCPs are connected by a string of five beads, representing 30bp linker DNA fragment. The linker DNA beads are of the same size but of full charge -12e. The distance between the beads is set to 20 Å (the same as for DNA wrapped around NCPs). The constructed nucleosome array contains 12 NCPs connected by $11 \times 5 = 55$ beads of linker DNA. The total charge of all DNA beads (DNA attached to the histone core plus linker DNA) is -3492e ($-236 \times 12 + -60 \times 11$). When combined with +1056e ($+88 \times 12$) positive charge of the histone tails, it gave net total charge of the whole array equal to -2436e. The “unit” nucleosome and array is shown in **Figure 3-2**.

Figure 3-2 The Sphere-Bead model of the NCP Array. (A), (B) Partially neutralized DNA beads are in red; fully charged beads modeling the linker DNA are in orange. (C) Model of the “12-177” nucleosome array, in initial configuration in the simulation cell.

3.1.3 Adoption of Neutrally Charged Core

Most of the positive charges are located on the periphery of the jacket of the cylinder-like histone octamer core, facing and neutralizing the DNA wrapped around the octamer.

After subtracting the positive charge of the basic amino acids involved in neutralization of the DNA charge, the net charge of the histone core appears to be neutral or negative (up to $-12e$ and depending on the choice of cutoff). We also used APBS (Baker, Sept et al. 2001), the Adaptive Poisson-Boltzmann Solver, to compute the electrostatic potential molecule surface for NCP in vacuum (amino acids pronation status set to $\text{pH}=6.0$, the condition for crystal formation), using PDB ID 1KX5 for histone protein and whole NCP respectively, shown in **Figure 3-3**. Histone protein surface is strongly positively charged but the existence of negative charged DNA causes the removal of almost all positively charged surfaces. For the “core” histone proteins, it brings our attention that the periphery areas of the core are more positively charged than some central areas, or in other words, the periphery is less negative than some central areas.

Figure 3-3 Electrostatic potential molecule surface computed by APBS (Baker, Sept et al. 2001) and rendered in PyMol for protein model 1KX5 at $\text{pH}=7.0$. (Left) Histone protein alone (Right) Whole NCP.

Our analysis is in line with earlier observation (Arents and Moudrianakis 1993) and recent analysis (Cherstvy 2009). Therefore, assigning a positive charge to the central particle representing the core domain of the histone octamer does not reflect the charge distribution of the real NCP; instead, we can use a neutral or a negative core and reduce the charge of the DNA wrapped on the core to account for the neutralization.

We made a few comparison test simulations with variations of the NCP models where a negative or positive charge of the central particles is applied, with results shown in **Figure 3-4**. For positively charged core, there is strong attraction between core and DNA from

neighbor NCPs. Because DNA is wrapped around the core on the periphery, NCP stacks with DNA from neighbor NCP with bias to DNA-core interaction (shown as stacking form 1, 2 in **Figure 3-4**) to form local aggregation albeit more artificially; comparatively, neutralized core NCP avoids such stacking (stacking form 3, 4).

Figure 3-4 Core-Core RDF between model with neutral core and model with positive core. The scheme of two NCPs shown near the peak in RDF represents how NCPs stacks locally to be shown as a peak in RDF.

There is an “acidic islet” (Luger, Mader et al. 1997; Luger and Richmond 1998) on each of the surfaces of the NCP “cylinder” which is formed by a cluster of 5 Glu and Asp amino acids of the H2A histones, which appear as the more negatively charged surface in the core in **Figure 3-3**. These acidic patches might play an important role in formation of ordered chromatin structures (Chodaparambil, Barbera et al. 2007; Zhou, Fan et al. 2007). The $pK_a=4$ for both Asp and Glu so at physiological pH conditions, they are deprotonated and carry negative charge. For the present description of the histone octamer as a spherical particle, assigning a negative charge of $-10e$ to it can approximate the effects of the acidic patch.

We then made some estimates of the simulations to the choice of the charge of the central particle testing two models with neutral (DNA beads is of charge $-9.44e$) or negatively charged ($-10e$) central particle (DNA beads is of charge $-9.04e$). The array properties obtained for two systems are tabulated in **Table 3-2**, which are very close.

Table 3-2 Comparison of array properties between systems of neutral core and negative core.

Salt Concen. System	$C_{Mg^{2+}}=1.58$ mM		$C_{Mg^{2+}}=3.89$ mM		$C_{CoHex^{3+}}=1.055$ mM	
	Array is unfolded.		Array is folded.		Array is folded.	
	$S_{20,w}$ (S)	R_g (nm)	$S_{20,w}$ (S)	R_g (nm)	$S_{20,w}$ (S)	R_g (nm)
Neutral Core	38.6	22.9	47.4 ± 1.3	16.5 ± 0.7	58.9 ± 0.9	12.0 ± 0.3
Negative Core	38.3	23.1	47.5 ± 2.2	16.0 ± 1.4	56.3 ± 0.8	12.6 ± 0.4

Using a positive charged core is not justified by the electrostatic surface and may create simulation artifacts. With the level of details in our model, we chose to use the neutralized core, assuming that all positive charges of the histone octamer core are fully neutralized (by the DNA) and therefore partially neutralizes DNA and not including the influence brought by “acidic islet”.

3.2 Dynamics and Simulation

3.2.1 Non-bonded Interaction

In simulations with explicit ions, all charged entities described above (NCP core, DNA, tails and ions) interact by a Coulombic potential in a dielectric medium with permittivity of 78. Ewald summation method implemented in P3M algorithm (Deserno and Holm 1998a) is used to treat the long-range electrostatic interactions in a periodic box.

To represent steric effects, we employed a short-range variant of a Lennard-Jones potential as inter-particle potential function for every pair of particle types.

$$U_{short}(r) = \begin{cases} \infty & r < r_{ij} \\ 4\varepsilon \left(\left(\frac{\sigma}{r - r_{ij}} \right)^{12} - \left(\frac{\sigma}{r - r_{ij}} \right)^6 + \Delta \right) & r_{ij} < r < r_{ij} + \sqrt[6]{2}\sigma \\ 0 & r > r_{ij} + \sqrt[6]{2}\sigma \end{cases}$$

where $\varepsilon = kT = 1$, $\sigma = 0.4nm$ and shift parameter $\Delta = 0.25$. In this function, the lowest point of the original Lennard-Jones curve $[\sqrt[6]{2}\sigma, -\varepsilon]$ is translated to $[r_{ij} + \sqrt[6]{2}\sigma, 0]$. This

potential energy function represents a soft purely repulsive potential, which smoothly goes down to zero at a cutoff distance $r_{ij} + \sqrt[6]{2}\sigma$ ⁴.

The parameter r_{ij} for each pair of particles represents the distance of minimal approach, which is equal to the sum of “hard” radii r_i of each particle: $r_{ij} = r_i + r_j$. The typical contact distance between a pair of particles (defined by the condition that the short range potential of Eq. (1) is equal to kT) is $r_{ij} + \sigma$, thus the effective radius of each particle can be determined as $r_i + \sigma/2$.

The particles are designated to have an effective radius according to the crystallographic structures NCP (Luger, Mader et al. 1997; Davey, Sargent et al. 2002) and DNA: histone core, 35.0 Å; DNA, 10.0 Å. The sizes of mobile ions and particles in the histone tails are set using the data of earlier Monte-Carlo (Lyubartsev and Nordenskiöld 1997; Korolev, Lyubartsev et al. 1999; Korolev, Lyubartsev et al. 2001a; Korolev, Lyubartsev et al. 2001b) simulations: tail bead, 3 Å; Acetylated tail bead, 4 Å; Mg^{2+} , 2.5 Å; K^+ and Cl^- , 2 Å; CoHex³⁺, 3.5 Å; NH^+ -group in Spermidine³⁺ and Spermine⁴⁺, 2.5 Å; r_i is set respectively for each kind of particles to satisfy above effective radius by subtracting $\sigma/2 = 2$ Å from effective radius. Some of these values differs slightly from those used in Sphere model. The guideline is the same, “adequately differentiating various types of particles with approximation to their physical sizes”.

⁴ It is derived as following:

$$(U_{ij})' = 24\epsilon\sigma^{12} \left(\frac{1}{r - r_{ij}} \right)^7 \left(1 - 2 \left(\frac{\sigma}{r - r_{ij}} \right)^6 \right)$$

So,

$$(r - r_{ij})_{\min} = \sqrt[6]{2}\sigma = 1.212246\sigma$$

At minima of U_{ij}

$$U_{ij} = 4\epsilon(0.25 - 0.5 + \Delta) = 4\epsilon(\Delta - 0.25)$$

To make it zero, shift shall be

$$\Delta = 0.25$$

3.2.2 Bond Interaction

The integrity of the histone core and 25 DNA beads (nucleosome DNA beads) wrapped around the central particle is maintained by assigning harmonic bonds with corresponding equilibrium distances between the DNA beads of the first and second turn, and between each DNA bead and the central core particle. A set of constraints with force constant $5 \text{ kT per } \text{\AA}^2$ is applied to these bead pairs to make the “Core Sphere + DNA + First bead of all tails” a rigid body.

Bonds between neighboring beads of histone tails uses a harmonic potential $(k_{\text{bond}}/2)(r - r_0)^2$ with equilibrium distance r_0 equal to $7 \text{ \AA}$, and force constant $k_{\text{bond}} = 25 \text{ kT per } \text{\AA}^2$. Two values of the force constant are set in order to keep the inter-bead distance unchanged. The force constant for wiring-up the core is smaller because there could be more than one constraint on each involved beads and to anticipate software error only.

The particles representing the amino groups of the Spd^{3+} and Spm^{4+} molecules are attached by a harmonic bond potential $V_{\text{bond}} = (k_{\text{bond}}/2)(r - r_0)^2$, with an equilibrium distance $4.7 \text{ \AA}$ and a force constant $k_{\text{bond}} = 10 \text{ kT per } \text{\AA}^2$. A harmonic angle potential between three particles is described by an equilibrium angle of 120 degrees (determined from full-atomic MD simulations) with a cosine potential energy of the form $(k/2)(1 - \cos(\Theta - 120^\circ))$ with $k = 120 \text{ kT per rad}^2$. In the limit of small deviation of the angle from 120° , it is equivalent to a harmonic bending potential with an angular force constant 30 kT per rad^2 .

To take into account the rigidity of the DNA, a bending potential for each three consecutive DNA beads is applied. This bending potential is of energy form $(k/2)(1 - \cos \Theta)$, where $k = 250 \text{ kT per rad}^2$ and Θ is the angle between the vectors going from each "i" to "i+1", and from "i" to "i+2" beads of DNA. In the limit of small bending angle, this potential corresponds to a harmonic bending potential with angular force constant 31 kT per rad^2 .

Linker DNA torsion is not included because of following two reasons.

- a) Realistic values of torsion force penalty are negligible because even for 11bp linker DNAs between the NCP, total twist cannot be more than $2/10\pi$. Even with a very

rigid linker DNA, not turning such small angle cannot prevent core-core attraction and the aggregation of the NCPs.

- b) For the real nucleosome array, rotating or twisting the DNA on the surface of the octamer may easily nullify twist of the linker DNA. The fact that the nucleosomal DNA is more flexible than one can suggest can be supported by an observation that the NCP crystal are obtained for the different DNA length from 145 to 147bp which means that nucleosomal DNA can be easily adjusted to more than 100° twist and 10 Å stretching, while stacking contact between the DNAs of different NCP in the crystal lattice remain intact.

From two points above, linker DNA twist may affect the structure only when nucleosomal DNA cannot accommodate it, which shall be the case during the later stage of nucleosome aggregation. Because our objective is about studying early-stage intra-array aggregation, inclusion of twist on linker DNA is not necessary.

3.2.3 Tailless NCP

We obtain tailless NCP by setting tail beads charge neutral to make long-range electrostatic forces zero; and hard radius $r_i = 0$, and $\sigma = 0$ to turn off their short-range interactions to other particles.

3.2.4 Chromatin Array Modeling in Debye-Hückel Approximation

We also performed simulations with Debye-Hückel (DH) approximation with the Sphere-Bead model. A harmonic angle potential for each two consecutive DNA bonds $(k_{bond}/2)(\phi - \phi_0)^2$ is applied with an angular force constant $25 \text{ } kT \text{ per } rad^2$ and $\phi_0 = \pi$. This setup corresponds to a non-electrostatic persistence length, L_p of 500 Å , based on the relation $L_p = kR_{bond}$, where R_{bond} is the bond length between DNA beads. This is the same bending rigidity of the linker DNA used by Schlick and co-workers (Zhang, Beard et al. 2003; Sun, Zhang et al. 2005; Arya and Schlick 2006; Arya, Zhang et al. 2006; Arya and Schlick 2009; Grigoryev, Arya et al. 2009). It differs slightly in potential energy form and angular force constant with our explicit ion simulations, i.e. 31 (explicit ion) v.s. 25 (DH) $kT \text{ per } rad^2$.

With DH approximation, the effect of ions of all types is expressed by a single parameter, the ionic strength that determines the screening parameter in all electrostatic interactions.

The electrostatic interactions between all charges of NCP, i.e. NCP core particles, DNA and histone tail beads, are described by a screened Coulombic potential:

$$U_{DH}(r) = \frac{q_i q_j}{4\pi\epsilon_0\epsilon} \frac{\exp(-r/r_D)}{r}$$

where ϵ_0 is vacuum dielectric permittivity, relative dielectric permittivity of water $\epsilon=78$, and Debye radius r_D defined by the ionic strength I :

$$r_D = \sqrt{\frac{\epsilon_0 \epsilon k T}{2 N_A e^2 I}} \text{ where } I = \frac{1}{2} \sum_{\alpha} z_{\alpha}^2 C_{\alpha}$$

where z_{α} and C_{α} being the valence and concentration of ion species respectively.

Simulations have been performed with different charged core:

- 1) Full charge, neutral core
- 2) Scaled charge (0.5), neutral core
- 3) Tailless, neutral core
- 4) Uncharged

3.2.5 Systems

3.2.5.1 NCP Systems

The first group is composed of two Mg^{2+} systems with rNCP and aNCP, with the same system configuration like the Sphere model. This group is studied to show how two models perform comparatively.

Once confirmed the feasibility of the Sphere-Bead model, we simulated four more groups of systems. They are all in “titration” manner for the 2x2 combinations between two kinds of NCP: “wt” (wild type) and tailless NCP, and two kinds of cations: Mg^{2+} and $CoHex^{3+}$.

3.2.5.2 Array Systems

For the nucleosome array simulations with the Sphere-Bead model, most of them (with a few systems in larger boxes for K^+ to justify the box sizes) were carried out in a periodic cubic simulation cell with a size of 120 nm containing single “12-177” nucleosome array. The starting structure of the simulation has the two ends of DNA going out from each

NCP forms an angle 90° , and consequently the array generated in this way adopts a “semi folded” conformation as displayed in **Figure 3-2(C)**.

The concentration of the array in (120 nm) cubic box is $0.961 \mu\text{M}$ (about $1 \mu\text{M}$), which corresponds to 4.024 mM in terms of DNA phosphate group concentrations. In most of the simulation systems, in addition to the cations neutralizing the negative charge of the array ($-2436e$), added salt ($M^{n+} \cdot nCl^-$) is present with 800 Cl^- ions corresponding to a Cl^- concentration of 0.77 mM . The influence of Mg^{2+} , $CoHex^{3+}$, Spd^{3+} and Spm^{4+} on the properties of the “12-177” nucleosome array is investigated in a number of simulations modeling substitution (“titration”) of K^+ for the multivalent cations in the array solution. For pure K^+ -, Mg^{2+} -, and $CoHex^{3+}$ -containing systems, additional simulations were carried out with increased concentration of multivalent salt. These systems with “excess” cations are included in the “titration” groups and analyzed together. Also, to highlight the importance of the histone tails for the chromatin folding, a number of simulation runs were performed for solutions of tailless array: all NCPs in the array are lacking histone tails; increase of the negative charge of the array is neutralized by addition of monovalent cations in the simulation cell.

To summarize, following groups of systems are simulated:

- 1) Array with K^+
- 2) Array with Mg^{2+}/K^+ : titration
- 3) Tailless Array Mg^{2+}/K^+ titration
- 4) Array with $CoHex^{3+}/K^+$: titration
- 5) Tailless Array with $CoHex^{3+}/K^+$ titration
- 6) Array with Spermidine $^{3+}$: titration
- 7) Array with Spermine $^{4+}$: titration

The simulated systems with number of ions in the simulation cell are summarized in the tables attached after “6 Array Simulation with The Sphere-Bead Model”.

3.2.6 Dynamics

For all systems, Langevin dynamics simulations at 300 K were carried out using the ESPResSo package (Limbach, Arnold et al. 2006). The dynamics setup is similar to the Sphere model described in “2.2.2 Dynamics” with Langevin Friction coefficient increased from $\gamma=0.01$ (used in the Sphere model simulations) to 0.05 to account for a larger

and more complex system, and meant to provide small correction to the Newtonian equations of motion to keep the temperature constant. Masses are given in reduced units and are set to 2u for the histone tail beads, 1u for the ions, 10u for the DNA beads, and 100u for the NCP core.

The systems are simulated for 3×10^7 - 8×10^7 simulation time steps. Configurations for analysis are collected after each 100 steps. Average values of all parameters are calculated at least for the final 2×10^7 time steps or longer, after achieving convergence.

3.3 Analysis

3.3.1 RDF, Cumulative Number and Core-Core Contact Number

Like the Sphere model, we employ RDF, cumulative number in our analysis. For array, we show Core-Core RDFs, but two adjacent cores along the linker DNA contributes a peak at distance around 18nm due to constraints from linker DNA, shadowing other cores' contribution. **Figure 6-6**(C) and (D) below show the difference in two Core-Core RDFs. To remove this peak, we adopt Core-Core RDF (excluding neighboring cores) and use its cumulative number at 30nm (determined from RDF as the boundary for core-core close contact) to define array's Core-Core contact number.

3.3.2 Sedimentation Coefficient

The use of sedimentation velocity analysis has been applied to nucleosome arrays (Hansen, Ausio et al. 1989; Garcia-Ramirez, Dong et al. 1992). It is calculated according to (Sun, Zhang et al. 2005) using the method developed by Bloomfield et al (Bloomfield, Dalton et al. 1967) following the Kirkwood approach (Kirkwood 1954).

$$S_{20,w} = S_1 \left(1 + \frac{R_1}{N} \sum_{i>j} \frac{1}{R_{ij}} \right)$$

where summation is taken over all $N = 10$ NCP pairs, R_{ij} is the distance between two nucleosomes, and parameters $R_1 = 5.5$ nm and $S_1 = 11.1$ Svedberg (S ; $1S = 10^{-13}$ sec) are adopted from previous works (Garcia-Ramirez, Dong et al. 1992).

3.3.3 Radius of Gyration

To measure the compaction of array, we used radius of gyration counting for all cores. It is defined using the standard definition:

$$R_g^2 = \frac{1}{N} \sum_{i=1}^N (R_i - R_{COM})^2$$

where R_{COM} is coordinate of the geometric center of array core particles.

3.3.4 Fiber Length

Fiber dimensions (diameter, length and density) are obtained by the averaged sizes of the enclosing rectangular box of the array when the array is aligned along its principal axis. The height of the box becomes the length of the fiber, where the geometric average of the two side lengths of the box is taken as the fiber diameter.

3.3.5 Bulk Concentration

AUC experiments are performed for a mean solution concentration, C_{Mg} , which is equal to the bulk concentration, C_{Mg}^{bulk} , far away from the array, which also corresponds to the concentration that one would have Donnan membrane equilibrium between the array solution and a solution with this bulk concentration. The “bulk concentration” in simulations is the average concentration of the cation in the region of the simulation cell with low electric field that is far away from the nucleosome array.

Array could be in a densely compacted state and also be in extended “beads-on-string” structure. In either case, it only occupies partial space of the simulation box. To calculate the bulk concentration, each frame of the simulation trajectory in convergence is treated as following: shifting the array to the center of the box so that, even when the array conformation is morphing from frame to frame, the bulk phase is always located in the inner side of the box boundary. We then count the number of ions in each slice along one direction perpendicular to the array extension direction and calculate ion density. A bell-shape curve of cation distribution is usually obtained with flat lines on two sides. The average value of the ion concentration obtained from the flat curves is the “bulk concentration”, according to **Figure 3-5**.

Figure 3-5 Extracting bulk cation density from density across simulation box. The arrows indicate the flat parts.

4 NCP Simulation with The Sphere Model

With the Sphere model of NCP, we investigated two factors affecting NCP aggregation: salt-dependency and tail modifications.

4.1 Salt-dependent Aggregation

Simulations in previous work (Korolev, Lyubartsev et al. 2006) have been reproduced and results have been redrawn in **Figure 4-1**. The definitions of the RDFs are described above in “2.3 Analysis”.

Figure 4-1 RDFs from four rNCP simulation system in different salt conditions. (a) Core-Core RDF. (b) RDFs of the counterions (K⁺ or Mg²⁺) relative to the center of NCPs. In the inset graph showing the far-end of the curve, there are drops in cation concentration in the Normal Salt and Mg Salt systems, implying a boundary between phase with condensed NCP particles and the bulk solution phase. (c) External tail-Core RDF. Appearance of maximum below 6nm distance indicates NCP-NCP aggregation. (d)

Internal tail-Core RDF. Normal Salt and Mg Salt systems are similar while Low Salt and Salt Free are similar. The crossover between two groups indicates the detachment of the histone tails from the negative core caused by interaction with other NCP particles observed in the Normal Salt and in the Mg Salt systems.

Mg Salt and Normal Salt systems form aggregates while Salt Free and Low Salt systems rarely form even small aggregates (2-3 particles). External tail-Core RDFs in **Figure 4-1(c)** show that the NCP aggregate is maintained by crosslinking; on the other hand no crosslinking occurs in the Salt Free and Low Salt systems. Values of the peak in Core-Core, External tail-Core and Internal tail-Core RDFs versus the time steps are showing coordinated movements, as demonstrated in **Figure 4-2**. These results from our newly performed simulations are identical to (Korolev, Lyubartsev et al. 2006).

Figure 4-2 Peak values of (from top to bottom) External tail-Core, Internal tail-Core and Core-Core RDFs (unscaled RDF value) versus the time steps.

Previous work (Korolev, Lyubartsev et al. 2006) mentioned that the results for rNCP model under four salt conditions are in agreement with experimental data which reported a sharp change of the second virial coefficient in solutions of NCPs at salt concentration above 50 mM NaCl (Mangenot, Raspaud et al. 2002; Bertin, Leforestier et al. 2004), which is caused by the emergence of NCP-NCP interactions via histone tail cross-linking. Similarly, in the presence of millimolar concentration of Mg^{2+} and high concentration of monovalent salt, aggregation of isolated (linker-free) NCPs has been observed experimentally (de Frutos, Raspaud et al. 2001).

4.2 Tail Modification

4.2.1 Modification of Histone Tails Weakens NCP-NCP Interaction

Figure 4-3 Results of the simulations for rNCP, aNCP, and paNCP models for the systems with KCl and MgCl₂ salts. (a) Core-Core RDF and (b) External tail-Core RDF collected for the converged region of the trajectories (last 1.6×10^7 simulation steps). The insert in (b) shows the change in the maximum intensity of External tail-Core RDF during the course of the simulation run. The numerical value of the RDF corresponds to the concentration of particles relative to the value expected for a uniform distribution. (c) Histograms of the averaged peak values of Core-Core RDF and (d) of External tail-Core RDFs under two salt conditions, 4.15 mM MgCl₂ (grid) and 65.4 mM KCl (shade). (e) Close-up view of a snapshot of three NCPs in aggregation, each nucleosome with its core and tails are in one color, showing the presence of tails participating in inter-core interactions.

Figure 4-3(a)(b) display radial distribution functions (RDFs) which are highly indicative of the NCP-NCP interaction. **Figure 4-3(c)(d)** are the histogram from the values of the RDF maxima of the curves displayed in **Figure 4-3(a)(b)**. Pronounced maxima are observed in Core-Core and External tail-Core RDFs for rNCP-KCl and rNCP-MgCl₂ systems at a distance of close Core-Core and histone tail-core contact. Aggregation of rNCPs proceeds to a different extent in the presence of K⁺ and Mg²⁺ ions.

Core-Core RDF

In rNCP-MgCl₂ system, a single conglomerate of 10 NCPs is formed with occasional dissociation of 1-2 NCPs whereas in rNCP-KCl system, several aggregates of 2-5 NCPs formed with continuous dissociation and exchange of NCPs during simulation. The histone tails play a decisive role in formation of the NCP-NCP contacts. For aNCP, only a small degree of NCP-NCP aggregation is observed in the presence of KCl. In aNCP-MgCl₂ system, the intensities of both Core-Core and External tail-Core RDFs are lower than in rNCP-MgCl₂ system. For paNCP model under the same salt conditions, no stable NCP-NCP contacts are observed.

It may be also noted that the intensity of the first maximum of Core-Core RDF in the Mg²⁺ systems is always higher than in the corresponding K⁺ systems, shown in **Figure 4-3(a)(c)**, despite the fact the ionic strength of the Mg²⁺ systems (50-59 mM, including neutralizing counterions) is always below the ionic strength in the K⁺ systems (84-89 mM). This observation demonstrates that the Debye-Hückel approximation, having ionic strength as the only parameter describing the effect of the ions, cannot provide an adequate description of interactions in strongly charged polyelectrolyte systems.

External tail-Core RDF

NCP aggregates are maintained by both cross-linking supported by the External tail-Core RDF shown in **Figure 4-3(b)** and by the screening of the negative charge of the core surface by the positive particles of the tails and cations, K⁺ or Mg²⁺. **Figure 4-3(e)** gives a close-up view of crosslinking in a cluster of three NCPs obtained from a snapshot similar to that in **Figure 4-7** below (upper right hand corner, labeled as rNCP-Mg100). It is noted that the extended tails of the red NCP is not free but also in contact with another NCP from the mirror periodic cell not shown in the snapshot. Still, most of the positively charged beads of the tails remains localized in the vicinity of their “host” core with only a fraction of them taking part in cross-linking even in the most condensed rNCP-MgCl₂

system, shown in **Figure 4-3(e)** and more clearly in **Figure 4-4(a)** below and will be discussed there.

Distribution of Ions

For all three NCP model variants (rNCP, aNCP, paNCP), mobile cations (K^+ and Mg^{2+}) accumulate near the negative surface of the core up to concentrations $\sim 1.4M$ K^+ and $\sim 1.2M$ Mg^{2+} with no significant difference between the models; a 10-20% difference is observed between the variants (rNCP, aNCP, and paNCP; data not shown). When stable NCP aggregates are formed, most of the cations are gathered inside the condensed phase and there is a drop in concentration of the cations between bulk and NCP aggregates. Interaction of oligocationic tails with chlorine anions is reduced due to the confinement of the tails in the area of the negative electrostatic field from the central particle. Although some accumulation of Cl^- ions is seen in the vicinity of the tails, the intensities of the maxima in the corresponding RDFs are lower than those of the K^+-Cl^- and $Mg^{2+}-Cl^-$ RDFs (data not shown).

4.2.2 Tail-Core Interaction in Single NCP

Figure 4-4 (a) Internal tail-Core RDFs of rNCP, aNCP, and paNCP in the presence of K^+ and Mg^{2+} salts for the systems with 10 NCPs in the simulation cell. (b) The intensity of the first maximum of Internal tail-Core RDF of a single-NCP system versus the salt concentration (KCl or $MgCl_2$). The maxima in (a) are labeled in (b) as filled symbols of the same shape as the corresponding systems in (b). The single unfilled red square with cross and unfilled blue triangle with cross, represent data from the systems corresponding to **Figure 4-6** below at zero percentage of Mg^{2+} in “titration” group of simulations.

The tail detachment is modest (albeit definitely present) for a system of a single rNCP or aNCP, even at very high Mg^{2+} salt, shown in **Figure 4-4(b)**. The detachment of the tails from the core is therefore not directly induced by increased salt concentration *per se*. The extension of tails is entropically favorable, but the electrostatic energy loss would prevent this unless the presence of other NCPs and conditions diminishing NCP-NCP repulsion are at hand; therefore, it is the presence of the other NCPs, the screening of NCP-NCP repulsion and the induced fluctuation in the ion distribution, which favors bridging and thus induces tail extension outside their cores. The existence of explicit water molecules in physical environment would make the extension of tails more entropically favorable although the partially hydrophobic tails, rich in Lys, Arg could be preferred to be buried. Yet, the electrostatic attraction from the excessive amount of negative charges on DNA is difficult to overcome.

The two additional points at 1.8 mM KCl, unfilled red square with cross (rNCP) and unfilled blue triangle with cross (aNCP), represent data from systems with 10 NCPs where the salt conditions favor NCP-NCP repulsion, i.e. low concentration of K^+ and an absence of Mg^{2+} (data are taken from the simulations described in **Figure 4-6** below and corresponding to zero percentage of Mg^{2+}). Notably, there is still a reduction in the maximum of the RDF for these 10-NCP systems as compared to the single NCP even when NCPs do not aggregate in the simulation cell. This is likely because in this rather dense system, frequent random contacts between different NCPs as well as a higher ionic strength from the neutralizing counterions cause some extension of the tail.

The data presented in **Figure 4-4** furthermore illustrates a seemingly paradoxical point about the mechanism of NCP-NCP interactions: “Stable NCP aggregation is only possible when tail interaction with the negative core is strong.” Acetylation and phosphorylation of the tails (leading to the reduction of positive charge) as well as addition of salt allow greater extension of the tails from their cores and may imaginarily facilitate contacts of the tails with the other NCPs. However, the strength of the tail-core contact becomes weak and the tails lose ability to hold NCPs close to each other with reduced positive charges. It can be observed at the initial stage of the simulations with 10 NCPs in rNCP- MgCl_2 and aNCP- MgCl_2 systems (when NCPs placed away from each other) that the internal RDFs of the tails are similar (data not shown) to those of the single rNCP or aNCP in the simulation cell at low salt, shown in **Figure 4-4(b)**. In later stage, they relocate from being associated with its “host” core and becoming associated to the neighboring nucleosome only after NCPs approach each other. Therefore, we conclude that the effect of the

screening on the attractive interaction between tails and their host core is not so important for the interaction between NCPs, as previously suggested (Mangenot, Leforestier et al. 2002; Mangenot, Raspaud et al. 2002; Bertin, Leforestier et al. 2004).

4.2.3 Tail Modification Positions

Figure 4-5 Simulation results for aNCP model with two acetylation sites in each tail but with different positions of acetylation. Values of the first maximum in (a) Core-Core RDF (b) External tail-Core RDF and (c) Internal tail-Core RDF are shown for the four aNCP models, under two salt conditions, 4.15 mM MgCl₂ (grid) 65.4 mM KCl (shade). Data for the “root”, “head”, and “even” aNCP models shown as columns; results of the native aNCP displayed as horizontal lines.

To understand how the position of the acetylation sites in the tails influences the NCP-NCP interaction, we look at six additional simulation runs carried out with different locations of the uncharged beads in the histone tail (“head”, “root” and “even”) for aNCP model in KCl and MgCl₂, in comparison to the “native” aNCP systems.

Values of the maximum of Core-Core, External tail-Core and Internal tail-Core RDFs are shown in **Figure 4-5**. The columns represent the three new aNCP models; horizontal lines represent the original, “native” aNCP model. As observed from **Figure 4-5(a)(b)**, four models display small differences in KCl solution. In the presence of MgCl₂, “even” aNCP significantly weakens NCP aggregation and tail bridging and “root” aNCP makes the weakest influence on NCP aggregation. Remarkably as concluded from **Figure 4-5(c)**, in spite of the rather complicated behavior in NCP-NCP interactions and formation of bridges displayed by the different models for aNCP, the “native” distribution of the acetylation sites within the histone tails demonstrates the greatest decrease in interaction of the tails with their host core.

Although the coarse-grained NCP model approximates to the atomic model, it is clear from the data that the relative position of the charge reduction has a pronounced influence on the behavior of the NCP systems. This demonstrates that the inherent electrostatics of the NCP system gives considerable room for external regulation by acetylation and phosphorylation, to alter electrostatically induced condensation properties of the system.

4.2.4 Titration Simulations of NCP

Figure 4-6 Intensities of the RDF maxima and number of the NCP-NCP contacts determined for rNCP (circles, black line) and aNCP (squares, red line) models in dependence on the Mg^{2+} content in the simulation cell: (a) Core-Core RDF; (b) External tail-Core; (c) number of contacts per core. Data were averaged for the converged region of the trajectories (last 1.0×10^7 steps of the MD trajectory).

Figure 4-6 (a) and (b) show peak values of Core-Core and External tail-Core RDFs for the “titration” groups and (c) presents contacts per core. The curves in each sub-figure are of almost identical. rNCP curves show more drastic changes in its values with increasing percentage of Mg^{2+} content; aNCP curves are smoother but they are all well beneath the rNCP curve, implying that the aNCP system is less sensitive to the Mg^{2+} in terms of aggregation.

Figure 4-7 shows the “dynamics” of the NCP-NCP contacts in rNCP and aNCP systems with different percentages of Mg^{2+} content. rNCP systems show a rather static behavior with little fluctuation in the number of NCP-NCP contacts, shown in **Figure 4-7** (top). A contrasting behavior is observed in aNCP systems where large fluctuations in the number of NCP-NCP contacts are observed during simulations, shown in **Figure 4-7** (bottom).

The degree of NCP aggregation depends on the “acetylation” state of the histone tails and on the relative amount of Mg^{2+} . Below 50% Mg^{2+} , no aggregation is observed for both rNCP and aNCP. It is noted that there is a sudden increase of value for rNCP curves between 50% to 60% marks in **Figure 4-6**. Aggregation of rNCPs occurs between 60% and 80% of Mg^{2+} and above 80% Mg^{2+} , a single aggregate containing all 10 rNCPs is formed. In aNCP systems, the existence of stable NCP-NCP aggregates consisting of 2-5 aNCPs is seen only above 80% Mg^{2+} , observed from the snapshots in **Figure 4-7**. Interestingly, there is a decrease of the NCP-NCP interaction for rNCP from 90% to 100% of Mg^{2+} . This effect might reflect a real event of “loosening” of the NCP aggregates in excess of Mg^{2+} which leads to re-dissolution of the NCP and chromatin at high concentrations of $MgCl_2$ (de Frutos, Raspaud et al. 2001; Schalch 2004).

Figure 4-7 Contacts per core in the “titration” systems of rNCPs (top) and aNCPs (bottom) with variation of Mg^{2+} content. Representative snapshots show different degrees of aggregation in the systems with rNCPs (snapshots at the top of the figure) and aNCPs (bottom).

4.3 Discussion

4.3.1 Mechanism of Tail Bridging and NCP-NCP attraction

With the Sphere model, we are able to reproduce previous results of four NCP systems (Korolev, Lyubartsev et al. 2006); furthermore, we performed simulations to study the effect of tail modifications and aggregation mechanism in “titration” systems. It is possible to draw conclusions on the mechanism of tail bridging and NCP-NCP attraction from our comprehensive study on tail modification, salt concentration and presence of divalent cations like Mg^{2+} in both deca-nucleosome system and single-NCP system.

From recent experimental work performed on solutions of linker free recombinant NCPs (Mangenot, Raspaud et al. 2002; Bertin, Renouard et al. 2007), it was argued that NCP-NCP interaction becomes possible when (or after) the histone tails extend from the NCP core. The tail extension is suggested to occur due to the salt screening, which leads to the dissociation of tails from NCPs (Mangenot, Leforestier et al. 2002; Mangenot, Raspaud et al. 2002; Bertin, Leforestier et al. 2004). NMR data of an earlier work (Walker 1984) reported an increase of dissociation of the histone tails from the chromatin array only when the salt concentration is raised well above 200 mM NaCl. Remarkably, aggregation and precipitation of chromatin is observed under conditions where no dissociation of the tails is registered (between 100 and 200 mM NaCl). This is in agreement with experimental data obtained in the NCP solutions (Mangenot, Raspaud et al. 2002; Bertin, Renouard et al. 2007) and with the results of our simulations.

However, we do not regard dissociation alone is the key to tail bridging. We propose that tails can promote NCP-NCP interaction through tail bridging only if the tail binding to the NCP (no matter to the host or to the foreign NCP) is strong. The function of salt (addition of KCl, presence of Mg^{2+} or other multivalent cation) is to screen the mutual repulsion of NCPs and promote fluctuation-induced attraction. When repulsion between NCPs is screened by salt, the tail-core interaction is still much attractive. The gain in entropy, due to the possibility of tails sampling the volume outside the core, allows close

contacts to be favorable. When the tail extension from the core can occur more freely due to the reduction in the positive charge of the tails in the cases of aNCP and paNCP, the attraction between core and tail is also reduced so the tails shall prefer to stay away from all cores, in the favor of entropy gain. The formation of crosslinking becomes both weak and entropically unfavorable so the aggregation is reduced.

4.3.2 Relevance of the NCP Simulation to Experiment

Modeling of NCP aggregation is highly relevant to the question of chromatin fiber formation. It is particularly in regards to one of its stages: inter-array aggregation. The process of aggregation during an Mg^{2+} titration experiment is known to be preceded by the compaction of individual arrays to a folded state (at about 1 mM Mg^{2+}) (Dorigo, Schalch et al. 2003), while inter-array aggregation then follows at higher concentrations. In the compact state, the linker DNA is folded with limited flexibility. The 12-mer arrays with linker length of 30bps DNA can be considered as a pre-aggregated cluster of 12 NCPs. The major factor that determines the inter-array aggregation of such clusters is expected to be the interactions between the nucleosomes of different arrays as they approach each other. This is a process, which is very similar to the NCP aggregation in the simulations.

In order to assess the validity of simulation results, it is desirable to make direct comparison with biochemical data. Ideally, NCPs containing specifically neutralized Lysine tails should be investigated with respect to the effects on Mg^{2+} -induced inter-particle aggregation (precipitation). One approach to this is the preparation of histones with selective Lysine to glutamine (K→Q) mutations, bringing neutralization like acetylation. The highly relevant result is from recent experiment work (Liu, Lu et al. 2011). It used well-defined tail charge neutralizations (either by acetylation or K→Q mutation) to study their salt-dependent aggregation behavior. It shows that tetra-acetylation (K5, K8, K12, and K16) in the H4 tail and single H4K16 acetylation weakens the NCP self-association. Cation concentration at 50% NCP precipitation ($C_{50\%}$) for Mg^{2+} is 1.4 ± 0.2 mM for wt-NCP, 2.3 ± 0.1 mM for AcK16-NCP, 3.1 ± 0.2 mM for quadO-NCP (four K→Q mutations on H4x2) and 5.9 ± 0.6 mM for quadAc-NCP (4 acetylated K on two H4). There is a trend of increasing Mg^{2+} concentration to induce aggregation with more neutralization applied to the tails.

On the other hand, Wang and Hayes (Wang and Hayes 2007; Wang and Hayes 2008) published a comprehensive study of the inter-array aggregation for all combinations of the

four histones where selective K→Q mutations had been introduced in each histone tail, characterized by EC_{50} . EC_{50} is the concentration of Mg^{2+} needed to precipitate 50% of the arrays. An increase in EC_{50} , i.e. a higher concentration of Mg^{2+} needed to induce aggregation would be expected for an electrostatic mechanism and is implied for NCPs based on the above simulation results that mimic the Mg^{2+} “titration” experiment.

The number of K→Q mutations is 2, 6, 6 and 4 for H2A, H2B, H3 and H4 histones respectively. 16 different histone mutant/wild-type octamer combinations were assembled to nucleosome arrays with 4 to 36 quenched Lysine charges in each NCP. The effect of these mutations on the EC_{50} values in Mg^{2+} -induced aggregation as detected by a precipitation assay is established. Although no simple relation between number of charges neutralized and EC_{50} is found, the results of this study showed that in general, the more positive charges that are neutralized, the higher is the amount of Mg^{2+} needed to induce aggregation. For arrays with combination of all histones mutated (36 charges neutralized), the EC_{50} value increased to 16.42 mM as compared to 2.38 mM for the wild type. For systems with only H3 and H4 histones mutated (20 charges neutralized), there is moderate increase of EC_{50} to 4.63 mM (data refer to a 12-mer nucleosome array with repeat length of 177 bp).

Even though direct quantitative comparison with our simulation data cannot be made, the general trend observed in these experiments is in agreement with the simulation results, which indicate a higher concentration of Mg^{2+} needed to induce aggregation of NCPs having 16 tail charges neutralized (**Figure 4-6** and **Figure 4-7**) with Lysine sites at *in vivo* positions. However, the effects observed in these array experiments (including a small unexpected decrease in EC_{50} for the H2A single mutant histone case) likely depend on additional factors. The relative tail exit position on the nucleosome core as well as a possible cooperation in the combined contributions of the individual tails (Wang and Hayes 2007; Wang and Hayes 2008) may affect the aggregation process.

Our present study illustrates the crucial importance of an explicit presentation of the charged particles in the histone tails and the mobile ions in solution. This principal advantage of the Sphere model should be preserved in further development.

4.3.3 Approximations in the Sphere Model

The Sphere model contains several approximations in order to cope with the complexity of the system with limited computational power. The continuum description of solvent water has the effect of neglecting solvent mediated effects, but is consistent within the level of a coarse-grained model. Solvent hydration of charged ions is approximately accounted for. We chose not to include coarse-grain water model because they are not needed in our electrostatic interactions-driven system. Furthermore, even a coarse-grained water model would increase computation time considerably. Core particle is described by a single-point charged sphere and approximates excluded volume and packing effect of the core and DNA wrapped. It's a highly coarse description. In reality, the shape of this core is more like a flat cylinder with the majority of the negative charges on the periphery of the cylinder jacket. The rationale for our approach is the tremendous net charges of this particle, which correspond to about -236e of the DNA + globular part of the histones with additional +88e in the flexible tails. The electrostatic forces involved in the interaction between NCPs are of long range. It is therefore expected that the general qualitative properties of aggregation and its dependence on histone tail charge and on the amount of small ions present, should not be highly affected by the details of the shape and charge distribution of the core particle. However, NCP-NCP interaction is of short-range and highly relevant with refined structure and charge features in model. Therefore, a more refined NCP structure and distribution of the charges is of importance. Recent experimental studies have implied importance of the wedge-like NCP shape inside the so-called 30-nm chromatin fiber (Robinson, Fairall et al. 2006) as well specific functions of the so-called "acidic islets" in the globular domain of the H2A histones, localized on the top and bottom of the NCP cylinder (Barbera, Chodaparambil et al. 2006; Chodaparambil, Barbera et al. 2007; Zhou, Fan et al. 2007).

Recent coarse-grained models of NCP and chromatin, e.g. (Sun, Zhang et al. 2005; Arya and Schlick 2006; Arya, Zhang et al. 2006) have a detailed description by assigning effective values to the nucleosome surface and tail charges by "charge renormalization" methods with omission of explicit mobile cations. This kind of collective treatment of the polyelectrolyte and its counterions is less rigorous and loses the specificity from multivalent cations.

4.3.4 The Sphere Model and Beyond

Present simulations have clearly demonstrated electrostatically salt-induced condensation of NCPs mediated by tail bridging for isolated (linker free) NCP. Our simulation system is a coarse grained NCP modeled as a charged sphere with flexible charged tails attached, surrounded by mobile counterions, in electrostatic continuum. This description is based on the rationale of using a model for which the electrostatic problem of ions is to be explicitly solved with precision and with a method (Langevin dynamics in NVT ensemble) that generates the correct equilibrium properties. The Sphere model captures two crucial physical mechanisms in NCP-NCP interactions:

- i) Attractive ion-ion correlation effects due to fluctuations in the ion cloud; and
- ii) Attractive entropic and energetic tail bridging.

In fact, the use of the Sphere model, which can be considered as the simplest possible representation of the NCP yet incorporating the major aspects of the electrostatics of the system, highlights the strength in our current approach.

5 NCP Simulation with The Sphere-Bead Model

We performed some control simulations to verify whether the general features and conclusions from the Sphere model are well preserved in the Sphere-Bead model. The systems simulated are rNCP and aNCP with K^+ and Mg^{2+} respectively. Results are shown in **Figure 5-1** and **Figure 5-2** below.

Figure 5-1 Core-Core RDF (solid lines) and cumulative number (dashed lines) for two rNCP models in the system with Mg^{2+} , computed from last 6×10^7 time steps: Sphere NCP model (black curves) and the Sphere-Bead NCP model (red curves). Representative snapshots are shown as inserts: (top) The Sphere model and (bottom) The Sphere-Bead model. The conditions for the simulation set up (including the number of mobile ions) are identical to rNCP- $MgCl_2$ simulations.

RDFs for the Sphere-Bead model are considerably broader than the corresponding RDFs for the Sphere model. It also displays two broad peaks. These features are caused by added details to the globular octamer to increase ways of how two NCP do close contact in the Sphere-Bead model. For example, two NCP can join either with the DNA chains adjacent to each other, or with the octamer cores in close proximity, giving rise to the two peaks, illustrated earlier in **Figure 3-4**.

Integration of the RDFs, i.e. cumulative number of NCP in the cluster, is used to compare the number of NCPs in clusters formed in the two models. From the cumulative number,

we can see that there are about four NCP cores within close distance from the central particle, for the NCP-MgCl₂ system with the Sphere model; and about three NCPs are coordinated to the central NCP with the Sphere-Bead model. Despite different intensities of Core-Core RDFs, systems with the Sphere and the Sphere-Bead models, the cumulative numbers are similar with small deviation for $r > 15$ nm which is the distance typical for the NCP-NCP separation in the second layer of densely aggregated NCPs.

Figure 5-2 Comparison of Core-Core RDFs (solid lines) and cumulative numbers (dashed curves) for the Sphere-Bead model of rNCP (black curves) and aNCP (red curves) calculated in the presence of Mg²⁺, computed from last 6×10⁷ time steps and last 2×10⁷ time steps respectively. Other conditions (box size, number of particles) are the same as for the systems rNCP-MgCl₂ and aNCP-MgCl₂

“Acetylation” of the tails also significantly reduces NCP-NCP aggregation in the Sphere-Bead model like the Sphere model. The amount of reduction is different from that but sufficiently large to show differences between rNCP and aNCP. **Figure 5-2** shows RDFs, cumulative numbers as well as representative snapshots for rNCP and aNCP in Sphere-Bead model.

In summary, independent of the details of the shape and charge distribution of the NCP core, we have observed that for both models, first, multivalent cations like Mg²⁺ can induce tail-mediated aggregation of NCPs (**Table 7-1(A)(B)** presents more data for Mg²⁺- and CoHex³⁺-induced NCP aggregation); second, reduction of tail charge (“acetylation”) reduces aggregation propensity under similar conditions of Mg²⁺ concentration. All these observations are in agreement with available experimental observations for NCPs.

6 Array Simulation with The Sphere-Bead Model

With the Sphere-Bead model for nucleosome array, we investigated its self-folding behavior with two variants: wt- and tailless array with different kinds of salt at various concentrations.

6.1 Preliminary Analysis

Some preliminary analysis have been performed to have a general understanding of model's behavior and to justify our setup (described in “3.2.5 Systems”) because our simulation systems have become more complicated and heterogeneous with the introduction of a more refined and structured model.

6.1.1 Size of the Simulation Cell

We first concern the size of the simulation cell because the length of the linker DNA, if fully extended, is $(12 \times 5) \times 20 \text{ \AA} = 1200 \text{ \AA}$, the same as the cubic length of the simulation cell. The head-to-tail distance of extended array in K^+ salt is about 100nm, which is close to the capacity of the cubic box of side length 120nm. So we run three simulations with monovalent (K^+) cations of varying simulation cell size (120, 160 and 180nm) but of similar ionic strengths. Detailed results can be found in **Table 6-5(A)**. Nucleosome array snapshots resemble one another and all calculated properties (Radius of gyration, and sedimentation) converged to similar values of three systems (see **Figure 6-1** below).

Figure 6-1 Simulation results of Array with same concentration of KCl salt in three boxes of different size. (A) Snapshots of the array at the end of the simulation run. (B) Variation of the radius of gyration (R_g , top graph) and sedimentation coefficient ($S_{20,w}$, bottom graph) in the course of the simulations. (C) RDFs calculated from the last 2×10^7 time steps of the simulation or longer. From left to right: Core-Core RDF including closest core in the array; Core-Core RDF excluding closest core in the array (exclusion of the closest core from Core-Core distribution highlights the folding of the array); and External tail-Core RDF.

It can be concluded that a 120nm cell is sufficient to accommodate the most extended structure of the “12-177” nucleosome array without alteration of the behavior of the array.

6.1.2 Quality of Simulation

We calculated the autocorrelation function of the system energy for (Array-CoHex-1.31mM) system in **Figure 6-2**. The AC function is close to zero for the whole course of our simulation time.

Figure 6-2 Autocorrelation function of the system energy (Array-CoHex-1.31mM system)

6.1.3 Cations Distribution

For cations, especially multi-valent cation, they may stick to unrealistic positions or regions during simulation. In our system, cations would be either attracted to negatively charged DNA beads or anions. The geometry and sizes in our models are carefully assigned to prevent cation stuck into cavities inside Sphere-Bead NCP model or form salt crystal with anions.

The 2D distribution of cations can be found in **Figure 3-5**, which shows that they highly aggregated with nucleosome array and density in other space is drastically low and uniform. This formed the basis to calculation “bulk concentration”. We also plot the locations of cation in 3D in **Figure 6-3** below, for two systems: one is Array-CoHex and another is Array-Spm system. For both systems, in the beginning of the simulation, all cations are uniformly distributed in the simulation box; in the end of simulation, a substantial amount of cations are attracted to the core, leaving the rest uniformly distributed in other space. Visually, cations in “bulk phase” are more sparsely distributed in the right-hand side figures than the left-hand side figures.

Figure 6-3 Above) Cation distribution (particles in green) in simulation box in the beginning (left) and the end (right) of simulation in Array-CoHex-1.31mM system. **Below)** Cation distribution (particles in yellow) in simulation boxes in the beginning (left) and the end (right) of simulation in Array-Spm-1.06mM system.

6.1.4 Convergence of Simulations

Figure 6-4 Monitoring Mean Square Displacement (MSD) of the cations in the simulation of the “12-177” nucleosome array in the presence of various cations. (A) MSD of K^+ cations in titration systems containing mixture of K^+ and various concentration of Mg^{2+} , $CoHex^{3+}$, Spd^{3+} and Spm^{4+} (identity and concentration of the cations are indicated in the graph; see **Table 6-5** for details) (B-E) MSDs of (B) Mg^{2+} , (C) $CoHex^{3+}$, (D) Spd^{3+} , and (E) Spm^{4+} ions. For each cation, we picked two systems out of all “titration” simulations, one system has $S_{20,w} \sim 35S$ (unfolded) and another is associated with $S_{20,w} > 50S$ (folded). Exception is given to Mg^{2+} for having the 2nd system of $S_{20,w} \sim 40S$ because of its less capability to fold the array than tri-valent cations.

Mean Square Displacement (MSD) of the particles is a function of the number of simulation steps. It is a good indicator of convergence, or state of equilibrium. The results of MSD calculation are summarized in **Figure 6-4**. Note that for the highly compacted systems (end of titration) the averaged mobility of the cation is higher than that observed in the middle of titration. This reflects the fact that at the end of the “titration”, where the cation concentration is relatively high, a substantial fraction of the ions are “free” and not constrained to the vicinity of the array.

One can see that in all cases, including the quad-valent Spm^{4+} , the MSD reach a value corresponding to the square of the size of the compacted array ($50^2 = 2500 \text{ nm}^2$) in less than 5×10^6 simulation steps. This means that on the scale of simulations (3×10^7 - 8×10^7

simulation steps) all ions can move along the whole array many times, showing that sampling is adequate. The nucleosome array stays essentially at zero MSD, and it seems that a certain amount of three- and higher valence ions cannot leave its vicinity as the divergence of two lines after both of them reach ~ 2500 after similar number of time steps, seen from C), D) and E). This amount of cations is determined by how many cations negatively charged nucleosome array and surrounding anions can hold from for leaving for outer space, which is favored by entropic gain.

If the motion of ions is restricted by the condition that they are always close to the array, the MSD cannot exceed the maximum distance between two distant points of the array (which is around 50-60nm for a half-folded array). It is important that MSD reaches this distance $(50\text{nm})^2$ at a point of time, which is short relative to the total simulation time. This means that, during the whole course of simulation, ions can travel many times between distant points of the array, providing good sampling over the space. Mono- and di-valent ions, and probably three-valence Spd^{3+} , can leave the array and go through the periodic boundaries. That is why the MSD for them increases with the number of steps. For the mono- and divalent ions, a linear regime is observed corresponding to the diffusion limit, the ions are not condensed on the array. They can leave it and go through periodic boundaries, which is why the MSD linearly increases with time.

6.2 Salt-dependent Folding of the Array

For the starting structure shown in **Figure 3-2(C)**, R_g is ~ 19 nm. This number is between the extended state ($R_g \sim 34$ nm for the array in K^+ salt, from our simulation results) and fully folded state ($R_g \sim 12$ nm array in the presence of CoHex^{3+} , also from our simulation results). In experiment, electron microscopy visualization (Schalch 2004) shows that a value of $S_{20,w} \approx 53\text{-}55\text{S}$ corresponds to a dense compaction of the “12-177” array with gyration radius, $R_g \approx 12\text{-}15$ nm.

We analyzed three systems and show their results in **Figure 6-5**. Configuration and results summary are in **Table 6-5(A)(B)(D)**. Each system contains 800 Cl^- and one kind of cation to show their individual capability to induce aggregation of the nucleosome array. The average concentrations in those systems are $3.17\text{mM}(\text{K}^+)$, $1.58\text{mM}(\text{Mg}^{2+})$, and $1.06\text{mM}(\text{CoHex}^{3+})$. And ionic strengths are $1.97\text{mM}(\text{K}^+)$, $3.55\text{mM}(\text{Mg}^{2+})$, and $5.13\text{mM}(\text{CoHex}^{3+})$.

Figure 6-5(A) shows the snapshots of “12-177” array at the end of the simulation trajectory when equilibrium is reached. The “beads-on-a-string” structure is observed at low monovalent ion concentrations, whereas the presence of multivalent cations leads to folded structures. Higher density structure is formed in the presence of trivalent cations compared to divalent cations. Array folding characteristics such as the sedimentation coefficient $S_{20,w}$, shown in **Figure 6-5**(B), or the radius of gyration, R_g , shown in **Figure 6-5**(C), fluctuate around an equilibrium values after 2×10^7 simulation steps. Other radial distribution functions (RDFs) such as Core-Core RDF, shown in **Figure 6-5**(D), or External tail-Core shown in **Figure 6-5**(E) as well as Internal tail-Core (not shown); mobile ion-core and cation-DNA (shown later in **Figure 6-8**) give more information on the differences in the compaction and mechanism of folding in different ionic environments.

Figure 6-5 Comparison of the results of simulations of the “12-177” array in the presence of K^+ (green line in the graphs B-E), Mg^{2+} (blue line) and $CoHex^{3+}$ (red line).

(A) Representative snapshots illustrating typical degree of folding of the array in the presence of the respective cations observed in the end of the simulations. (B) Sedimentation coefficient ($S_{20,w}$) and (C) radius of gyration (R_g) along the course of the

simulations. (D) Core-Core and (E) External tail-Core radial distribution functions obtained by averaging the last 2×10^7 time steps or longer of each of the simulation runs. In (B) Core-Core RDF neighboring (connected by fixed-length linker DNA) NCPs were excluded from statistics. The three peaks clearly visible in Core-Core RDF of the array in the presence of CoHex^{3+} (red line) are related to three possible types of the NCP-NCP contact illustrated by sketches near the corresponding peaks.

The systems with trivalent CoHex^{3+} ions and di-valent Mg^{2+} ions display the largest intensity and more types of structures in the peaks of Core-Core RDF, shown in **Figure 6-5(D)**. This reflects that the highly folded state is obtained even at very low ionic strength with multivalent cations but is not possible with monovalent ions. We performed a simulation with average concentration of $5000\text{K}^+/7.20\text{mM}$ and ionic strength 6.00mM and obtained sedimentation coefficient at merely $34.1 \pm 0.4\text{S}$, see last row in **Table 6-5(A)**. The larger peaks in External tail-Core RDFs for multivalent cation systems, shown in **Figure 6-5(E)**, highlight the importance of the tail bridging events in the folding process and folded structure.

6.3 “Titration” Simulation

“Titration” simulations refer to a set of simulations with gradual replacement of monovalent K^+ with multivalent cations up to complete substitution and even to have excessive cations. From the changes in properties of nucleosome array, we can better reveal what factors are affecting the aggregation process. Each kind of multivalent ions, Mg^{2+} , CoHex^{3+} , Spd^{3+} , and Spm^{4+} has a set of “titration” simulations and there are two sets for Mg^{2+} , CoHex^{3+} with tailless nucleosome array respectively.

Figure 6-6 Properties of the “12-177” nucleosome array in the presence of various amount of Mg^{2+} . Concentrations of Mg^{2+} in the simulated systems are indicated on the right-hand side of the graph. Variation of (A) sedimentation coefficient, $S_{20,w}$, and (B) radius of gyration, R_g , in the course of the simulations. (C-E) RDFs calculated from the last 4×10^7 steps of the simulation. From left to right: (C) Core-Core RDF including closest core in the array; (D) Core-Core RDF excluding closest core in the array; and (E) External tail-Core RDF. Number of ions and major results of the simulations are given in **Table 6-5(B)** below.

We plotted the “raw” analysis results from Mg^{2+} “titration” group in **Figure 6-6**. For $S_{20,w}$ shown in (A) and R_g shown in (B), they are drawn as time series with the number of time steps as x-axis. (C), (D) are to illustrate the difference in Core-Core RDF including and excluding the adjacent cores. In **Figure 6-6(E)**, with the increase of Mg^{2+} concentration, RDFs transform from a broader plateau to a narrower peak and also moves closer to the core. This shows that tail-bridging mechanism becomes operative when cores are approaching sufficiently close.

To better visualize the results, we extract the key descriptor of each curve and draw one curve for one set of “titration” system. For sedimentation coefficient $S_{20,w}$ and radius of gyration R_g , their descriptors are averaged values calculated over the converged trajectory segment. For RDFs, we take Core-Core contact number, which is Core-Core RDF (excluding neighboring cores) cumulative number as (discussed above in “3.3.1 RDF, Cumulative Number and Core-Core Contact”), and peak values for External tail-Core and Internal tail-Core RDF to represent them. These results are plotted in **Figure 6-7** and tabulated in **Table 6-5(B)-(G)**.

Figure 6-7 Summary of simulation results for the "12-177" array "titration" by Mg^{2+} (A-D); $CoHex^{3+}$ (E-H); Spd^{3+} and Spm^{4+} (I-K). (A, E) Representative snapshots illustrating folding of the array at different stages of titration by (A) Mg^{2+} and (E) $CoHex^{3+}$; numbers under the snapshots indicate bulk concentration (1st line) and average (2nd line) Mg^{2+} or $CoHex^{3+}$ concentration. (B, F, I) Radius of gyration, R_g ; (C, G, J) sedimentation coefficient, $S_{20,w}$; (D, H, K) intensity of the maximum in External tail-Core RDF. In graphs B-D and F-H, data for the array with charged tails are drawn as solid circles and lines; similar results for the tailless array are shown as hollow points with dashed lines. X-axis is the average concentration of Mg^{2+} (B-D, C_{Mg}), $CoHex^{3+}$ (F-H, C_{Co}); x-axis is bulk concentration for Spd^{3+} or Spm^{4+} (I-K, C_{Spd}/C_{Spm}).

6.3.1 Folding of wt-Nucleosome Array by Mg^{2+}

Figure 6-7(A-D) shows the results from simulations of Mg^{2+} . Increasing C_{Mg} from zero to 3.89 mM leads to a gradual folding of the array as indicated by a decrease in R_g from 34.3

to 16.5 nm and increase in $S_{20,w}$ from $31.2 \pm 0.3S$ to $47.4 \pm 1.3S$. Unlike tri- and tetravalent cations, Mg^{2+} does not induce complete compaction under these ionic conditions, like $CoHex^{3+}$ at 1.31mM with $57.2 \pm 0.6S$.

In experiment, maximal folding is achieved at 1 mM Mg^{2+} (Korolev, Allahverdi et al. 2010). This is due to the inequality of the ionic conditions in experiments and simulations. First, the average concentration in simulation cannot be directly compared to the concentration found in experiment, detail is described in “3.3.5 Bulk Concentration”. Second, there is a difference in the charge ratio of C_{mg}/C_p . There is higher phosphate concentration C_p in simulation.

We attempted to see how simulation performs with charge ratio difference mitigated. The analytical ultracentrifugation (AUC) experiments measuring sedimentation coefficients, in the presence of magnesium have been performed at a low array DNA phosphate concentration of 0.12 mM and the sedimentation velocity coefficient, $S_{20,w}$ at maximum folding, is obtained at 1.0 mM added $MgCl_2$. The ratio of added cation charge to array DNA phosphate concentration is $R_{exp} = 2C_{Mg}/C_p \approx 17$. Hence, an excess of cation charge relative to array charge is necessary to induce the full folding. Simulations under such conditions of array concentration are difficult due to computer limitations (box size and number of particles). However, we performed one additional simulation for conditions closer to the experimental charge ratio of 17, at $R_{sim} = 2C_{Mg}/C_p = 7.9$ and obtained an $S_{20,w}$ value of 52S (Table 6-5(B), last row with 16474 Mg^{2+}) in close agreement with the experimental value of 52.5S at maximum folding.

6.3.2 Folding of wt-Nucleosome Array by $CoHex^{3+}$

Figure 6-7(E-H) presents the results of simulations of the “12-177” array obtained from simulations in the presence of $CoHex^{3+}$. In agreement with experiments (Korolev, Allahverdi et al. 2010) the trivalent cation shows much greater efficiency than Mg^{2+} in folding the array, comparing to **Figure 6-7(A-D)**. Even at an average concentration $C_{Co} = 0.74mM$, practically full compaction of the array is reached which is reflected by small R_g (12-13 nm, shown in **Figure 6-7(F)**) and high value of $S_{20,w}$ (close to 52S, shown in **Figure 6-7(G)**). The folding of the arrays occurs when the bulk concentration of $CoHex^{3+}$ is in the micro molar range. The middle snapshot in **Figure 6-7(E)** is displayed for the corresponding $C_{bulk} = 2.35\mu M$ (this corresponds to an average concentration $C_{Co} = 0.74mM$) and for this or larger concentrations, practically full folding has occurred.

Maximum folding is obtained at a bulk concentration of 0.28mM (1 mM average concentration) with a sedimentation coefficient of 59S, larger than the maximal value observed in experiments. This is probably because aggregation in experiments involves more arrays on simulations while only a single array is present and measured in the simulation,

The contributions from the histone tail bridging to array folding is observed in the intensity of the external core–tail RDF shown in **Figure 6-7(H)**, which exhibits a steep increase that is correlated with the folding of the array, also seen in the change of gyration radius and sedimentation velocity, shown in **Figure 6-7(F)(G)**. This is explained by the following argument: in the initial stages of the titration, folding of the array is caused by increased shielding by cations of the negative charge on NCPs and linker DNA. When the NCP-NCP distances become sufficiently close, the tail bridging is operative. Because CoHex³⁺ is able to fold array denser than Mg²⁺, such step-wise behavior of measured quantities is more pronounced and thus more visible with CoHex³⁺ than Mg²⁺.

6.3.3 Folding of wt-Nucleosome Array by Spd³⁺ and Spm⁴⁺

We performed a number of simulations in the presence of the natural polyamines, Spermidine³⁺ and Spermine⁴⁺. The mobile polyamines are modeled as a chain of three and four charged beads. The results are displayed in **Figure 6-7(I-K)** as a function of the ‘bulk’ concentrations. In agreement with experiment, Spd³⁺ shows lower efficiency to fold the array than Spm⁴⁺ and it is also less potent in inducing the folding than is CoHex³⁺. Significant compaction of the array is reached at bulk C_{Spd} of tens to one hundred μM . The same level of chromatin folding is observed for C_{Co} around 10 μM and for C_{Spm} in the range 1 - 10 μM bulk concentrations. The maximal folding with sedimentation coefficients of 58.8S for Spm⁴⁺ and 55S for Spd³⁺ is in good agreement with those observed in the experiments.

6.3.4 Folding of Tailless Array

The tailless array shows rather similar folding characteristics to the normal system in the lower concentrations in the ‘titration’ experiments, but it is clear that the tailless system has a much-reduced folding propensity according to **Figure 6-7(B-C)** and (F-G), in agreement with the published AUC results (Dorigo, Schalch et al. 2003). In terms of $S_{20,w}$, Mg²⁺ and CoHex³⁺-tailless NCP systems can reach ~80% of their wt-nucleosome array

counterparties. Titration curves of tailless systems show the transition to a folded state at higher concentration of cation. Mobile Ion-DNA interactions take the place of tail-DNA interaction to contribute to the folding of tailless array, as discussed below in “6.3.5 Cation-DNA RDF Cumulative Number”

6.3.5 Cation-DNA RDF Cumulative Number

Figure 6-8 Cation-DNA RDF cumulative number for wt-nucleosome array (A) K^+ ($C_K = 3.17$ mM), (B) Mg^{2+} ($C_{Mg} = 1.58$ mM) and (C) $Co(Hex)^{3+}$ ($C_{Co} = 1.06$ mM); (the same three systems shown in **Figure 6-5**) and tailless nucleosome array (D) Mg^{2+} ($C_{Mg} = 4.18$ mM) (E) $Co(Hex)^{3+}$ ($C_{Co} = 1.65$ mM) Data is shown up to radial distance of 30nm. Linker DNA and core DNA RDFs are calculated individually.

In **Figure 6-8** above, we show the cation-DNA RDF cumulative number (defined in “2.3.2 Cumulative Number”) of two most folded tailless systems, one for Mg^{2+} and one for $Co(Hex)^{3+}$, coupled with three wt-nucleosome array systems, which are the K^+ , Mg^{2+} and $Co(Hex)^{3+}$ systems found in **Figure 6-5**. The wt-nucleosome array K^+ system is unfolded with sedimentation coefficient $S_{20,w}$ at 31.2S; for Mg^{2+} systems: $S_{20,w}$ 38.6 $\pm$ 0.9S (wt-) and 36.8 $\pm$ 0.7S (tailless), both in loosely folded statue; for $Co(Hex)^{3+}$ systems, $S_{20,w}$ at 57.2 $\pm$ 0.6S (wt-) and 52.4 $\pm$ 2.0S (tailless).

We can observe that, from all figures, linker DNA attracts more cations than the DNA wrapped around histone core because linker DNA is fully exposed while nucleosomal

DNA is exposed less than half, and linker DNA is fully charged at $-12e$, compared to $-9.44e$ for nucleosome DNA. More importantly, we can observe that there are more cations gathered around the DNA beads in the tailless NCP systems than in the wt-array systems. For Mg^{2+} , the cumulative number at 30nm is almost double in the tailless system than in the wt-system. For $CoHex^{3+}$, the same holds and the gradient of the curve for tailless-system attenuate slower than that of the wt-system.

6.3.6 Estimation of Nucleosome Fiber Dimension

Table 6-1 Fiber dimensions for compact “12-177” arrays calculated from the analysis of the simulations

Cation Conc. (mM)	Fiber Length (nm)	Fiber Width (nm)	Density (Nucleosome/11 nm)
Analysis excluding the ends (first and last three nucleosomes):			
Mg^{2+} 3.89	48.6	25.02	4.14
Mg^{2+} 15.84	47.1	24.04	4.07
Co^{3+} 0.74	53.7	26.42	2.78
Co^{3+} 0.95	33.4	21.90	4.23
Co^{3+} 1.06	39.6	22.09	4.76
Co^{3+} 1.31	42.8	23.20	4.72
Spd^{3+} 0.79	48.1	22.56	4.30
Spd^{3+} 1.06	37.5	24.24	4.67
Spm^{4+} 0.59	39.5	22.69	4.07
Spm^{4+} 0.79	36.1	22.30	4.58
Analysis based on all 12 nucleosomes:			
Mg^{2+} 3.89	48.6	29.2	2.7
Mg^{2+} 15.84	47.1	26.5	2.8
Co^{3+} 0.74	53.7	28.4	2.5
Co^{3+} 0.95	33.4	27.0	4.0
Co^{3+} 1.06	39.6	24.2	3.3
Co^{3+} 1.31	42.8	25.9	3.1
Spd^{3+} 0.79	48.1	26.7	2.7
Spd^{3+} 1.06	37.5	25.5	3.5

Spm ⁴⁺ 0.59	39.5	27.8	3.3
Spm ⁴⁺ 0.79	36.1	25.5	3.7

In general, the fiber appearance bears some resemblance with electron microscopy images reported in literature, e.g. in (Routh, Sandin et al. 2008) where longer (80 NCPs) arrays with 167bp NRL were visualized in the absence and in the presence of linker histones with fiber diameter in the range 18-25 nm and array linear density 5-7 NCP/(11 nm). We made a statistical analysis of the fiber parameters for concentrations corresponding to compact arrays in the systems of all four cations, see **Table 6-1** above. For Mg²⁺ in the most compact state, the fiber diameter is about 24 nm and the nucleosome density is about 4.1 NCPs/(11 nm) for analysis excluding the first and last three nucleosomes. Including the ends significantly decrease the density, demonstrating the end effects. This fiber width is close to 25 nm expected for an idealized two-start model, while the density predicted from the tetra-nucleosome structure is about 5.8 NCPs/(11 nm) (Schalch, Duda et al. 2005). The discrepancy in density is likely due to end effects for the present short array and also strongly influenced by the lack of close stacking imposed by our spherical core model. In the presence of the cations of higher valence, more compact globular states are formed. Little is known about chromatin structure in the presence of multivalent ions, but due to strong binding, more compact fibers would be expected.

6.4 Discussion

6.4.1 Mechanism of Array Folding

The difference between NCP and array is the addition of negatively charged linker DNA. Although it needs more energy to overcome the electrostatic repulsion for aggregation, the advantages that it brings are 1) mechanically, the whole array is linked together and core beads' movement is constrained 2) the attraction of cations to gather around the array with resulting screening repulsion and inducing attractive ion-ion correlation 3) the attraction to tails (H3 and H4 tails are near the exits of DNA) to extend out and to bind to linker DNA beads near to the "foreign" core.

Our points in previous sections ("4.2.2 Tail-Core Interaction in Single NCP and Deca-NCP Systems" and "4.3.1 Mechanism of Tail Bridging and NCP-NCP attraction") are supported with the tailless array simulations. The dissociation of tails is a phenomenon but is not the key to the tail bridging. It requires cations to screen the mutual repulsion of

NCPs and promote fluctuation-induced attraction first; afterwards, the stronger attraction from the foreign core promotes bridging and stable NCP aggregation. In the case of array, both core and DNA are of negative charge.

The function of salt (addition of KCl, presence of Mg^{2+} or other multivalent cation) is to screen the mutual repulsion of NCPs and promote fluctuation-induced attraction. The advantage of multivalence cations is that they give larger attraction due to ion correlations.

Tails are constrained to have one bead fixed to the histone octamer core and the rest of the beads are flexible and extended. This gives mono-valence tail beads greater advantage than freely moving multivalent cations to neutralize the DNA and to contribute the “tail-bridging” mechanism. For tailless nucleosome array, mobile ions, taking the place of tails, gather around linker DNA to make mobile ion correlations contribute to the salt-dependent folding of array.

6.4.2 Comparison with Experimental Sedimentation Velocity Studies of the 12-177-601 Array in the Presence of Various Cations

In (Korolev, Allahverdi et al. 2010), a wide variety of cations: K^+ , Na^+ , Mg^{2+} , Ca^{2+} , $CoHex^{3+}$, Spd^{3+} , and Spm^{4+} have been tested to induce aggregation for “12-177-601” array, characterized by AUC. Results are summarized in **Table 6-2** and described below. This experiment is a systematic study of the compaction for a range of oligocations, of varying nature and charge ranging from +1 to +4e. Array folding is found to exhibit typical polyelectrolyte behavior, with the critical concentration of cations needed to bring about maximal folding and cation valences differentiate their effects, covering a range of almost five orders of magnitude (from 2 μM for Spermine⁴⁺ to 100 mM for Na^+).

With monovalent counterions, maximum compaction is achieved at 100 mM NaCl and KCl with $S_{20,w} = 54.5S$ and $46.7S$ respectively. Remarkably, K^+ and Na^+ , showed considerable difference in $S_{20,w}$ values. Moreover, among all ions studied, only K^+ displays an $S_{20,w}$ value not reaching 53-55S. For Mg^{2+} and Ca^{2+} , maximum compaction is observed at 1.0 and 0.8 mM ($S_{20,w} = 53.7S$ and $54.5S$ respectively). Ca^{2+} shows slightly higher efficiency in folding of the array but this difference with Mg^{2+} is rather small compared to the difference in the $S_{20,w}$ values of the Na^+ - K^+ pair. Maximum folding of the array by $CoHex^{3+}$ and Spd^{3+} is recorded at 20 μM and 60 μM respectively, with the compact state ($S_{20,w} \sim 50S$) being reached at about 10 μM $CoHex^{3+}$ and 40 μM Spd^{3+} . The observed

higher efficiency of CoHex^{3+} in inducing compaction compared to Spd^{3+} is in agreement with the established effects of these cations on DNA condensation (Korolev, Berezhnoy et al. 2009). As expected, the cation with the highest positive charge, Spermine⁴⁺, showed the highest efficiency in the folding of the array. Sub-micro molar concentration of Spm^{4+} increased the value of the sedimentation coefficient with maximal $S_{20,w}$ achieved at $C_{\text{Spm}} = 2 \mu\text{M}$.

A remarkably wide range in the observed concentrations needed to induce the maximum folding can be observed: from 100 mM for Na^+ to 2 μM for Spm^{4+} of $\sim 50,000$ in magnitude! This behavior strongly suggests a polyelectrolyte nature of the nucleosome aggregation process due to cation neutralization.

Table 6-2 Maximum concentration of cations giving reproducible sedimentation velocity curves and maximum $S_{20,w}$ values obtained in the AUC measurements.

Added cation	Concentration, (mM)	$S_{20,w}$ (S)
Reference (TEK buffer)**	--	35.4±0.3
K^+	100	46.9±0.3
Na^+	100	54.5±0.6
Mg^{2+}	1.0	52.5±0.7
Ca^{2+}	0.8	54.5±0.3
$\text{Co}(\text{NH}_3)_6^{3+}$	0.02	54.5±0.3
Spermidine ³⁺	0.06	54.8±0.5
Spermine ⁴⁺	0.002	53.25±0.5

*: For all cations except TEK buffer and Mg^{2+} , the AUC experiment has been repeated twice; the data for the TEK buffer are averaging of 7 measurements; there are 5 data sets of AUC data for Mg^{2+} at 1.0 mM.

** : TEK buffer: 10 mM KCl; 10 mM Tris, pH 7.6; 0.1 mM EDTA. (At pH 7.6 Tris is a mixture of protonated and basic forms with protonated form contributing 8.0 mM of the Tris cations).

The following considerations should be noted when comparing the experimental and theoretical results:

- i) The limitations set by computational capability means there are differences in ionic and array concentrations between the experimental and simulation systems. For this reason, we cannot model the monovalent salt-induced chromatin array folding.
- ii) Values of $S_{20,w}$ obtained in the simulations with monovalent salt are lower than those for the lowest experimental KCl salt (TEK buffer). This difference in the ($S_{20,w}$) values is well borne with different ionic strengths applied in the AUC experiment and in the simulations.
- iii) In the AUC experiment on the other hand, the salt concentration (for the cases of mono- and divalent cations) is significantly larger than the array DNA concentration ($C_p = 0.12$ mM). In simulations, there are comparable numbers of salt and array charges to phosphate concentration in simulation is $C_p = 4.024$ mM.
- iv) In the simulations with single array in the cell, intra-array folding cannot be hindered by inter-array interaction. Hence, it is possible to achieve additional folding, due to even stronger attractions induced by ion-ion correlations at higher concentrations of multivalent ions.

Quantitative reproduction of experimental values is not the main purpose of our simulation. Instead, we aim at predicting qualitative trends of the salt-induced compaction in order to understand the underlying mechanism at the level of the basic electrostatic interactions.

Qualitative comparison of the experimental results and simulation results are performed by comparing the threshold bulk concentrations of the cations where sedimentation coefficient values corresponding to a compact state ($S_{20,w} \sim 50S$) are reached. In the simulations, these approximate concentrations are ~ 5 mM (Mg^{2+}), ~ 30 μM (Spermidine³⁺), ~ 2 μM (Co^{3+}) and ~ 2 μM (Spermine⁴⁺). The corresponding experimental values are 1 mM, 40 μM , 10 μM and 1.5 μM . This is considered as an excellent agreement, considering all the approximations in the model and simulations.

From the discussion above, simulation and experiments suggest that multivalent cation can induce aggregation at lower concentrations.

6.4.3 Comparison with Other Models

A number of chromatin models with structural details have been developed and presented in the literature (Wedemann and Langowski 2002; Sun, Zhang et al. 2005; Arya and

Schlick 2006; Arya, Zhang et al. 2006; Korolev, Lyubartsev et al. 2006; Langowski and Heermann 2007; Kepper, Foethke et al. 2008; Arya and Schlick 2009; Grigoryev, Arya et al. 2009). These models often aim at predicting and understanding the structure of the elusive “30nm” chromatin fiber and consider a realistic shape of the NCP. However, the flexible tails are often omitted or fixed, and the nucleosome-nucleosome interactions are usually modeled with implicit effective attractive potentials, to mimic a given salt condition where the folded structure is formed. The issue of the salt-dependent process of folding in the sense of different kind of salt of different valences is usually not considered in these models and therefore a simplified treatment of the electrostatics may be warranted for these studies.

Schlick and co-workers studied the salt-dependent nucleosome array folding with computer simulations by using a Discrete Surface Charge Optimization (DiSCO) model of the nucleosome core particle and, recently by including the flexible histone tails (Arya and Schlick 2006; Arya, Zhang et al. 2006; Arya and Schlick 2009; Grigoryev, Arya et al. 2009). The treatment combines a detailed shape of the NCP with a more approximate modeling of the histone tails, linker DNA and linker histone H1. The effects of added salt are described by a Debye-Hückel treatment, not taking into account ion-ion correlations. The monovalent salt dependence of the array folding is well described by these simulations. A charge renormalization is used to change the nominal charges of the particles in order to match the PB potential at a given salt condition. This results in concentration-dependent charge adjustment of the NCP, tails, linker DNA, and linker histone. At low ionic strength the net (absolute) charge of the arrays seems to be reduced below its nominal value, while at high ionic strength it is larger (Arya, Zhang et al. 2006; Arya and Schlick 2009). The modeling of the presence of divalent ions is using a phenomenological approach, which is to introduce a modification of the Debye screening length for linker DNA interactions, and another modification of the DNA persistence length. In our view, the physical picture of the origin of chromatin array folding caused by divalent ions, which seemingly emerges from this approach, remains unclear.

6.4.4 Chromatin Array Modeling in Debye-Hückel Approximation

Debye-Hückel (DH) is a popular way to represent the effect of added salt in modeling. However, this approach suffers from several drawbacks (see below) in strongly charged systems like chromatin, which is built from DNA polyelectrolyte. We present our DH simulations results in **Table 6-3** and **Figure 6-9**.

Table 6-3 Selected results from simulations of arrays using Debye-Hückel approximation. Simulations of the uncharged model of the “12-177” array give $R_g = 18.4 \pm 1.0$ nm and $S_{20,w} = 43.5$.

Ionic strength, mM	Full charge, neutral core		Scaled charge (0.5)		Tailless, neutral core	
	R_g , nm	$S_{20,w}$	R_g , nm	$S_{20,w}$	R_g , nm	$S_{20,w}$
1.0	45.9 ± 1.0	27.6	42.8 ± 1.0	28.5	46.7 ± 0.9	26.9
1.98	38.5 ± 0.7	28.9	36.3 ± 0.7	30.6	--	
3.59	35.9 ± 0.9	30.0	34.4 ± 0.8	31.7	35.4 ± 0.7	29.9
4.75	34.6 ± 0.5	30.8	30.7 ± 1.2	32.4	--	
6.79	31.1 ± 0.8	31.5	29.5 ± 1.1	33.1	--	
10.0	29.9 ± 0.9	32.7	--		30.7 ± 0.7	31.8
16.7	26.1 ± 0.8	35.3	25.6 ± 0.7	35.7	--	
30	22.6 ± 0.8	38.1	--		26.4 ± 0.7	34.6
50	19.4 ± 1.5	41.9	20.2 ± 1.4	41.3	24.5 ± 0.6	36.7
70	16.2 ± 1.3	46.7	19.7 ± 0.8	41.5	21.4 ± 0.6	39.2
100	15.2 ± 0.8	49.8	18.7 ± 0.6	42.9	21.1 ± 0.7	39.8
200	15.8 ± 1.1	48.0	18.1 ± 1.3	43.2	18.0 ± 0.9	43.5

Our DH model shows the dependence on $S_{20,w}$ and R_g for wt and tailless arrays for a range of ionic strengths from 1 to 200 mM. In addition, the results for a 50% reduced charge (scaling) model, sometimes used in DH modeling to represent ion condensation and charge renormalization (Arya, Zhang et al. 2006; Arya and Schlick 2009), are presented. The result for an uncharged model is shown for reference as a dashed line. For the normal array, the folding becomes similar to the uncharged model at 50-80 mM ionic strength. In the range 100-200 mM the array adopts a more compact form with R_g values less than that in the uncharged model and $S_{20,w}$ approaches a value of 50S, which is characteristic for a compact (folded) structure. This compaction is caused by the histone tails, which are present explicitly and are able to promote attraction through a correlation in their positions and by bridging. The tailless model does not show full compaction effect and in fact never reaches compaction beyond that displayed by the uncharged model. These results are also in qualitative agreement with previous results obtained by a DH model (Sun, Zhang et al. 2005; Arya, Zhang et al. 2006).

DH approximation fails when multivalent ions are present in the system. The experiments demonstrate array compaction at divalent ion concentrations of about 1 mM, which corresponds to an ionic strength below 20 mM (including KCl and buffer). In all DH calculations, the array is essentially stretched for such values of the ionic strength but in experiments and in our the Sphere-Bead model simulations, three- and four-valent counterions cause array folding at submillimolar concentrations.

Figure 6-9 Simulations within Debye-Hückel approximation as a function of ionic strength for different models of the array. (A) Sedimentation coefficient, $S_{20,w}$; (B) Radius of gyration, R_g . Horizontal dashed lines indicate the values of $S_{20,w}$ and R_g calculated for the array model with uncharged particles. Hollow ovals in (A) and (B) show respectively $S_{20,w}$ and R_g values obtained in the simulations with explicit ions; cations are indicated near the points. Solid ovals in (A) show the maximal $S_{20,w}$ values observed in the AUC measurements for different cations (cation is indicated in the graph). Indices “ Mg^{2+} ” and “ high Mg^{2+} ” mark the simulation data calculated respectively for $C_{\text{Mg}} = 4.024$ and 14.752 mM.

6.4.5 Shortcomings of Debye-Hückel Approximation

As specified in **Table 6-4** below, different contributions of electrostatic origin to chromatin statics and dynamics are categorized as following:

Table 6-4 Modeling of electrostatic contributions to chromatin folding

Model	Electrostatic contribution					
	Cation-anion interactions	Screening /charge neutralization	Cation multi-valence/structure	Mobile ion correlations	Tail-tail correlations	Tail bridging
	1	2	3	4	5	6
Array with charged tails; explicit ions	+	+	+	+	+	+
Tailless array; explicit ions	+	+	+	+	-	-
Array with charged tails; DH approximation	+/-*	+/-*	-	-	+/-*	+/-*
Tailless array; DH approximation	-	+/-*	-	-	-	-
Array without charges	-	-	-	-	-	-

*: +/- means that this type of interaction is included, but incompletely.

- 1) **Cation-anion attraction** is the static Coulombic interaction between positively (histone tails and mobile cations) and negatively (mainly DNA) charged units in a given configuration. It is partly described in DH models since there are no explicit mobile cations in this treatment. The DH treatment with charged histone tails includes tail-DNA interactions.
- 2) **Screening/charge neutralization** refers to the reduction in electrostatic interactions due to the ion atmosphere (screening) and counterion condensation and binding on the DNA polyelectrolyte (charge neutralization), leading to an effective neutralization of its negative charge. This is present in the full explicit ion as well as partly in the DH models. The importance of counterion release, which is absent in the DH descriptions, is mostly ignored in the context of chromatin, but is a well-known entropic effect for DNA condensation (Bloomfield 1997; Korolev, Berezhnoy et al. 2009). It is the entropy gain due to the small ion release caused by an increased oligocation association to DNA in the collapsed state (as compared to the random coil).

- 3) **Cation multivalence/structure** refers to effects of the chain-like structure of the polyamine oligocations and accounts for differences between oligocations of the same charge, but of different nature (e.g. CoHex³⁺ and Spd³⁺).
- 4) **Mobile ion correlation** is the dynamic correlation of the fluctuating delocalized binding of the mobile counterions to the neighboring polyions (like NCP or linker DNA). It is probably the most important contribution to attraction between nucleosomes neglected in the DH/PB models.
- 5) **Tail-tail correlation** refers to the contribution caused by attractions induced by the correlated fluctuations in the positions of tails belonging to different nucleosomes and is included in both treatments if a charged flexible tail model is used. It is the same as mobile ion correlations but applied to the histone tails of the neighboring NCPs. However, a DH treatment is approximate in this respect since this correlation is coupled to correlations with small mobile ions, while in the DH treatment it always occurs as the background of a screened interaction.
- 6) **Tail bridging** is the binding of the histone tails to the neighboring NCP and to the linker DNA. It may not be clearly distinguished from the tail-tail correlation. It refers to the entropy gain upon tail binding to negatively charged sites on adjacent nucleosomes rather than its own core. This mechanism is clearly observed by the increase in the intensity of External tail-Core RDFs upon folding, shown in **Figure 6-7**.
- 7) This interaction is also incompletely described in a DH model for the same reasons as just mentioned in the case of tail-tail correlations, i.e. no explicit mobile ions.

All contributions listed above are either absent or lack adequate representation within chromatin models based on Debye-Hückel approximation.

6.5 Conclusion

The Sphere-Bead model contains more detailed description for NCP and includes linker DNA. We employed it to run simulations for nucleosome array. Observed from simulations, higher valent cations are able to fold nucleosome array more effectively than lower valent cations with lower concentration and denser structure. Even in the absence of the tails, partial folding of chromatin is observed, both in experiment (Dorigo, Schalch et al. 2003) and in simulations. The mechanism of chromatin array folding requires a sufficient neutralization of the negative charge of the nucleosome by screening and oligocation binding as well as attractive contributions due to tail-related interactions and

multivalent ion-ion correlations, comparing with DNA condensation in (Korolev, Berezhnoy et al. 2009). It is clear that a physically grounded theory of the cation-induced chromatin compaction should include the entropy, correlations and competition effects of mobile ions, only possible with explicit ions.

The current approach that we used treats electrostatics in a consistent manner and manifests the differences in various cation nature and valence. It consequently contributes to an understanding of the physical origin of structural and phase transitions of these highly important systems. As indicated in **Table 6-4**, important contributions to chromatin folding are either absent or incompletely represented in the DH models and hence the inability of DH model to show the folding of array induced by the multivalent cations and the folding of tailless array.

There are several approximations in the present modeling, one of the most important ones being the approximate description of the shape and charge distribution of the core particle of the nucleosome array, which limit the “stacking” between the surfaces of the real wedge-shaped cylinder-like NCPs. Another serious approximation is the lack of hydration effects, i.e. the continuum description of the aqueous solvent with a uniform dielectric constant. Other simplifications of the real system include the absence of hydrogen bonding and hydrophobic effects.

Table 6-5 Summary of simulations for the “12-177” nucleosome array.

Table 6-5(A) Monovalent (K⁺) cations (^: Three systems shown in **Figure 6-5**)

Cell size, nm	Number of ions		Average K ⁺ , mM	Bulk K ⁺ (*), mM	Ionic strength, mM	Sedimentation coefficient, S	Radius of gyration, nm	(+) Core- Core contact number	RDF maximum, rel. units	
	K ⁺	Cl ⁻							(++) ExtTail- core	(+++) IntTail- core
180	5196	2700	1.48	0.804	1.12	30.8±0.3	34.6±1.1	3.07	1.15	74.0
160	4392	1896	1.78	0.812	1.27	30.7±0.3	35.6±0.8	2.92	1.15	74.1
120	^3296**	800	3.17	0.874	1.97	31.2±0.3	34.3±1.2	2.88	1.19	74.0
120	7496	5000	7.20	4.879	6.00	34.1±0.4	28.9±0.9	3.91	1.53	73.0

* Total charge of the DNA (number of the phosphate groups) in the “12-177” array is $12 \times (-294-60) = -4188e$. For the 120nm box used in the most of the simulations , volume is 1.728×10^{-18} liter. The concentration of the array is $0.96096 \mu\text{M}$ (one particle per cell, $=1/\text{Vol}/N_A$), concentration of phosphate group $C_P = 4.024 \text{ mM}$ ($=4188/\text{Vol}/N_A$). The histones neutralize charge of the DNA wrapped in the NCP from $-294e$ to $-236e$ ($-294+58$) and this neutralization is accounted by assigning charge -9.44 to each of 25 DNA beads; additionally, NCP carries the tails of positive charge $+88e$. As a result, net charge of the array is $-236 \times 12 + -60 \times 11 + 88 \times 11 = -3492 + 1056 = -2436e$; that gives concentration of the negative charge in the 120 nm box 2.34 mM and the neutralization of the DNA charge (combined neutralization from the globular histones and the histone tails) is 0.4141 , so 58.59% of the DNA charge remains un-neutralized by the histones. For the tailless array in the 120 nm cell, C_P remains the same $C_P = 4.024 \text{ mM}$ but DNA is neutralized only by the globular histones so “12-177” array carries the charge $-3492e$, concentration of the negative charge is 3.356 mM ; neutralization of the DNA charge is equal to 0.1602 , so 83.98% of the DNA charge remains un-neutralized by the histones.

** Concentration of 2496 K⁺ ions neutralizing negative charge of the array is 2.34 mM

+ Core-Core contact number is the cumulative number from Core-Core RDF (excluding neighboring cores) at 30nm radial distance. (Integration of RDF for distance [0, 30nm]) (Discussed above in “3.3.1 RDF, Cumulative Number and Core-Core Contact”)

++ Peak value of External tail-Core RDF.

+++ Peak value of Internal tail-Core RDF.

Table 6-5(B) Array with histone tails; Mg^{2+} titration

Number of ions		Average Mg^{2+} , mM	Bulk Mg^{2+} , mM	Ionic strength, mM	Sedimentation coefficient, S	Radius of gyration, nm	Core-Core contact number	RDF maximum, rel. units	
K ⁺	Mg^{2+}							ExtTail-core	IntTail-core
3296	0	0.0	0.0	1.97	31.2±0.3	34.3±1.2	2.88	1.19	74.0
2472	412	0.40	0.0001	2.36	32.1±0.3	31.7±1.05	3.21	1.17	64.4
1648	824	0.79	0.0066	2.76	33.4±0.4	30.3±1.1	3.92	1.22	57.0
824	1236	1.19	0.1130	3.16	36.1±0.6	25.45±0.9	4.61	1.81	52.5
330	1483	1.43	0.2798	3.39	37.2±0.7	24.8±0.85	5.09	2.05	51.6
^0	1648	1.58	0.4262	3.55	38.6±0.9	22.9±1.0	5.71	2.11	51.3
0	1748	1.68	0.5141	3.84	39.5±1.0	21.9±1.1	6.12	2.26	51.1
0	1848	1.78	0.5954	4.13	40.3±1.3	21.2±1.1	6.56	2.33	50.9
0	2048	1.97	0.8088	4.70	40.5±1.0	21.4±1.0	6.62	2.41	50.6
0	4048	3.89	2.6955	10.47	47.4±1.3	16.5±0.7	8.25	3.68	48.0
0	16474	15.83	14.752	46.29	52.5±0.8	14.3±0.4	8.13	4.63	45.1

Table 6-5(C) Tailless array; Mg^{2+} titration*

Number of ions		Average Mg^{2+} , mM	Bulk Mg^{2+} , mM	Ionic strength, mM	Sedimentation coefficient, S	Radius of gyration, nm	Core-Core contact number
K ⁺	Mg^{2+}						
4352	0	0.00	0.0	2.48	30.7±0.3	36.2±0.8	2.15
3264	544	0.52	0.000054	3.00	31.3±0.2	34.4±0.9	2.29
2176	1088	1.05	0.00244	3.52	31.6±0.3	34.0±0.9	2.53
1088	1632	1.57	0.058	4.04	34.5±0.4	28.0±0.7	3.26
436	1958	1.88	0.220	4.36	36.4±0.5	25.6±0.7	3.79
0	2176	2.09	0.432	4.57	36.8±0.7	24.9±0.9	3.93
0	3264	3.14	1.440	7.70	38.8±0.7	22.6±0.8	4.50
0	4352	4.18	2.500	10.84	39.7±1.1	21.6±1.5	4.94

*: Concentration of the negative charge of the tailless array is 3.356 mM (in charge units)

Table 6-5(D) Array with histone tails; CoHex³⁺ titration

Number of ions		Average CoHex ³⁺ , mM	Bulk CoHex ³⁺ , μM	Ionic strength, mM	Sedimentation coefficient, S	Radius of gyration, nm	Core-Core contact number	RDF maximum, rel. units	
K ⁺	CoHex ³⁺							ExtTail-core	IntTail-core
3296	0	0.0	0.0	1.97	31.2±0.3	34.3±1.2	2.88	1.19	74.0
2966	110	0.11	0.28	2.29	31.4±0.3	34.5±0.9	2.68	1.20	68.9
2309	329	0.32	0.53	2.92	31.8±0.4	33.6±1.1	3.12	1.16	61.8
1649	549	0.53	1.105	3.55	33.3±0.4	31.1±0.7	3.68	1.27	55.3
1319	659	0.63	1.27	3.87	34.9±0.4	28.7±0.9	4.11	1.46	52.0
989	769	0.74	2.35	4.18	51.9±0.9	14.3±0.5	10.73	4.47	45.8
329	989	0.95	154.2	4.82	57.0±1.2	12.5±0.5	11.63	6.31	43.4
^2	1098	1.06	284.0	5.13	59.0±0.9	12.0±0.3	11.95	7.33	42.8
0	1365	1.31	500.0	6.67	57.2±0.6	12.6±0.2	11.70	6.90	43.0

Table 6-5(E) Tailless array; CoHex³⁺ titration

Number of ions		Average CoHex ³⁺ , mM	Bulk CoHex ³⁺ , μM	Ionic strength, mM	Sedimentation coefficient, S	Radius of gyration, nm	Core-Core contact number
K ⁺	CoHex ³⁺						
4352	0	0.00	0.0	2.48	30.7±0.3	36.2±0.8	2.15
3264	363	0.35	0.202	3.52	31.1±0.3	36.1±0.8	2.80
2177	725	0.70	0.576	4.57	32.2±0.4	33.2±0.9	2.80
1740	871	0.84	0.721	4.99	33.4±0.3	29.8±1.0	3.94
1088	1088	1.05	1.153	5.61	38.1±0.9	23.3±1.1	5.61
435	1306	1.26	126.3	6.24	49.9±1.4	15.2±0.6	10.36
0	1451	1.39	275.5	6.66	49.4±1.5	14.7±0.9	10.56
0	1717	1.65	524.1	8.19	52.4±2.0	14.0±0.9	11.12

Table 6-5(F) Spermidine³⁺ titration

Number of cations		Average Spd ³⁺ , mM	Bulk Spd ³⁺ , μM	Ionic strength, mM	Sedimentation coefficient, S	Radius of gyration, nm	Core-Core contact number	RDF maximum, rel. units	
K ⁺	Spd ³⁺							ExtTail-core	IntTail-core
3296*	0	0.0	0.0	1.97	31.2±0.3	34.3±1.2	2.88	1.19	74.0
2474	274	0.26	0.0769	2.76	33.1±0.2	30.4±0.2	3.91	1.18	68.2
1649	549	0.53	0.2857	3.55	33.9±0.2	30.3±0.4	4.03	1.21	64.0
824	824	0.79	30.83	4.34	50.5±1.0	15.5±0.6	9.89	4.26	55.3
0	1099	1.06	256.8	5.14	55.0±0.9	13.3±0.4	11.54	6.00	54.1
No tails, 0	1451	1.39	261.1	6.66	45.7±1.4	18.2±0.9	8.39	--	--

Table 6-5(G) Spermine⁴⁺ titration

Number of cations		Average Spm ⁴⁺ , mM	Bulk Spm ⁴⁺ , μM	Ionic strength, mM	Sedimentation coefficient, S	Radius of gyration, nm	Core-Core e	RDF maximum, rel. units	
K ⁺	Spm ⁴⁺							ExtTail-core	IntTail-core
3296*	0	0.0	0.0	1.97	31.2±0.3	34.3±1.2	2.88	1.19	74.0
2472	206	0.20	0.0505	3.16	32.3±0.4	32.8±1.3	3.44	1.08	66.7
1648	412	0.40	0.1441	4.34	34.1±0.4	28.4±0.6	4.33	1.24	61.1
824	618	0.59	3.114	5.53	54.2±0.8	14.0±0.5	10.75	5.70	51.6
0	824	0.79	209.3	6.72	58.8±0.8	12.3±0.3	11.74	7.26	51.5
No tails, 0	1088	1.05	198.9	8.75	54.3±1.3	13.6±0.6	11.11	--	--

7 Tailless NCP Simulation with The Sphere-Bead Model

The tailless NCP contains tail-truncated form of histones. It is also called globular NCP, denoted as g-NCP. In experiment, individual truncated histone constructs are prepared to forms tailless histones and subsequently recombined with DNA to tailless NCP. The truncation positions are wt-H2A(1-129aa)→g-H2A(13-118aa), wt-H2B(1-122aa)→g-H2B(22-122aa), wt-H3(1-135aa)→g-H3(27-135aa) and wt-H4(1-102aa)→g-H4(20-122aa), thus including all N-terminals and H2A C-terminal deletion as well. In modeling, tailless NCP is achieved with turning off both long- and short-range interaction for tails with other particles, described in “3.2.3 Tailless NCP”. To offset the loss of +88e from positively charged tails, we make them up from monovalent ions K^+ , which would not promote aggregation even with small concentrations.

7.1 Simulation Results

Figure 7-1 Core-Core RDFs and snapshots from simulation of wild-type NCP and g-NCP with the Sphere-Bead model in Mg^{2+} and $CoHex^{3+}$

We show Core-Core RDFs of five representative systems in **Figure 7-1**. Three are for tailless NCP (two with $CoHex^{3+}$, one with Mg^{2+}), and two for wt-NCP (with Mg^{2+} and

CoHex³⁺ respectively). The system with excessive CoHex³⁺ for tailless NCP system is in dashed line; other systems are in solid lines.

wt-NCPs form a single dense aggregate with CoHex³⁺ (blue) and smaller aggregates with Mg²⁺ (green). Under same conditions, regardless of Mg²⁺ (red) and CoHex³⁺ (black), g-NCPs rarely form long-lived large aggregates and their Core-Core RDFs are rather flat. The dash line representing excessive CoHex³⁺ system with 2.75 times of CoHex³⁺ exhibits a smaller peak. Short-lived formed aggregates contribute to this peak because RDF is not temporal to manifest the “life” time of aggregates. There are no large aggregates observed from the snapshot, shown in the right-lower corner in **Figure 7-1**.

To give a clearer picture, we performed other simulations with g-NCP systems with different cation concentration. The configuration and results are tabulated in **Table 7-1**, together with results from wt-NCP system for comparison. In **Figure 7-2**, we present the Core-Core contact numbers, Internal tail-Core RDFs and External tail-Core RDFs in a titration curve manner with concentration as x-axis.

Table 7-1 Simulation result summary with the Sphere-Bead model for wt- and tailless NCP. (Systems shown in Figure 7-1 are denoted by *)

(A) wt-NCP with Mg²⁺

Number of ions			Average conc., mM		Ionic strength, mM	Sedimentation coefficient, S	Core-Core contact number	RDF maximum, rel. units	
K ⁺	Mg ²⁺	Cl ⁻	Mg ²⁺	Cl ⁻				ExtTail-core	IntTail-core
1800	0	320	0.00	8.30	27.50	38.36±0.25	0.008	0.001	3.44
450	675	320	17.51	8.30	45.02	40.65±0.96	0.907	6.54	2.66
*0	900	320	23.35	8.30	50.85	44.31±1.93	2.025	15.45	2.53
0	1060	640	27.50	16.61	63.31	47.32±1.58	2.800	20.81	2.48

(B) wt-NCP with CoHex³⁺

Number of ions			Average conc., mM		Ionic strength, mM	Sedimentation coefficient, S	Core-Core contact number	RDF maximum, rel. units	
K ⁺	CoHex ³⁺	Cl ⁻	CoHex ³⁺	Cl ⁻				ExtTail-core	IntTail-core
1350	150	320	3.89	8.30	39.18	38.48±0.27	0.036	0.023	3.09
*0	600	320	15.57	8.30	74.21	53.33±1.54	4.145	32.13	2.34

(C) Tailless NCP with Mg^{2+}

Number of ions			Average conc., mM		Ionic strength, mM	Sedimentation coefficient, S	Core-Core contact number
K^+	Mg^{2+}	Cl^-	Mg^{2+}	Cl^-			
2680	0	320	0	8.30	38.92	38.36 ± 0.24	0.012
2010	335	320	8.69	8.30	47.61	38.37 ± 0.25	0.010
1340	670	320	17.38	8.30	56.30	38.59 ± 0.29	0.034
880	900	320	23.35	8.30	62.27	38.66 ± 0.33	0.037
670	1005	320	26.08	8.30	64.99	38.85 ± 0.37	0.085
0	1340	320	34.77	8.30	73.69	39.78 ± 0.55	0.304
0	1420	480	36.84	12.45	79.91	39.74 ± 0.56	0.365
0	1500	640	38.92	16.61	86.14	39.72 ± 0.64	0.350

(D) Tailless NCP with CoHex^{3+}

Number of ions			Average conc., mM		Ionic strength, mM	Sedimentation coefficient, S	Core-Core contact number
K^+	CoHex^{3+}	Cl^-	CoHex^{3+}	Cl^-			
2680	0	320	0	8.30	38.92	38.36 ± 0.24	0.012
2011	223	320	5.79	8.30	56.28	38.31 ± 0.16	0.005
1342	446	320	11.57	8.30	73.63	38.53 ± 0.20	0.019
880	600	320	15.57	8.30	85.62	38.94 ± 0.45	0.090
538	714	320	18.53	8.30	94.50	39.72 ± 0.54	0.431
268	804	320	20.86	8.30	101.50	41.40 ± 1.05	0.976
1	893	320	23.17	8.30	108.43	41.55 ± 0.70	1.178
0	1000	640	25.95	16.61	125.06	41.89 ± 1.13	1.170

Figure 7-2 (a) Core-Core contact number and magnitudes of maximum in (b) External tail-Core RDFs for titrations of systems of ten g-NCPs (upper), or wt-NCPs (lower), with Mg^{2+} (red, square) and CoHex^{3+} (black, hollow circle).

wt-NCP systems show increasing aggregation with the increase of cations. CoHex³⁺ is able to induce greater aggregation with a smaller amount than Mg²⁺. For tailless systems, both CoHex³⁺ and Mg²⁺ cannot induce aggregation of more than 2-3 NCPs as observed in the Core-Core contact number curve. This is different with array's results where Mg²⁺ is able to induce about 50% compaction (folding) of wt-NCP for tailless NCP, and CoHex³⁺ can induce almost 100% of compaction with ~125% CoHex³⁺ (1717 v.s. 1365)

7.2 Discussion

7.2.1 Experimental Results

In recent experiment reported in (Liu, Lu et al. 2011), the authors used precipitation assay (PA¹) and static light scattering (SLS²) to study the self-association of NCPs induced by the addition of cations, i.e. Mg²⁺, Ca²⁺, Cobalt(III)-hexamine³⁺, Spermidine³⁺ and

¹ **Static light scattering (SLS)** Static light scattering (SLS) experiments is performed as such: NCP solution (1 mL) in a cuvette was titrated by step-wise addition of high concentration stock solution containing chloride salts of the cations, Mg²⁺, Ca²⁺, Spd³⁺, CoHex³⁺ and Spm⁴⁺ until the maximal level of NCP aggregation was achieved as detected from the static light scattering (SLS) signal intensity increase. The effectiveness of a cation in inducing condensation was quantified by the molar concentration of the ligand needed for a 50% increase in the static light scattering intensity (EC₅₀). Values of EC₅₀ were calculated from the titration curves displaying the scattering intensity versus ligand concentration using a sigmoidal fitting.

² **Precipitation assay (PA)** Precipitation of NCPs induced by addition of oligocations was studied by a series of solutions (typically 20-24 aliquots of 20 μ L each) prepared containing fixed concentration of the NCP (0.22 mg/mL or 0.552 μ M NCP or C_p=162 μ M; absorbance at 260 nm A²⁶⁰ = 1) and increasing concentration of the chloride salts of the cations (Mg²⁺, Ca²⁺, Spd³⁺, CoHex³⁺ and Spm⁴⁺). After 20 min incubation at room temperature, NCP aggregates were precipitated by centrifugation (20,000 g $\times$ 10 min). Values of A²⁶⁰ were measured from the supernatant. The fraction of NCP remaining in the solution at various cation concentrations was derived by dividing the A₂₆₀ absorbance of each aliquots by that of the standard. Results of the precipitation assay were plotted as fraction of NCP remaining in the solution versus concentration of added ligand and fitted by sigmoidal function similar to the light scattering data. C_{50%}, which reflects the cation concentration at 50% precipitation of NCPs, was used to quantify the efficiency of all cations in these systems.

Spermine⁴⁺, for NCPs reconstituted from wild-type histones (wt-NCP), or N-terminal truncated histones (g-NCP), or H4 histone tail having acetylation or K→Q mutations at positions K5, K8, K12 and/or at position K16. The experimental results regarding wild-type and globular NCPs are summarized below first.

EC₅₀, derived from SLS, quantifies the effectiveness of a cation in inducing condensation by the molar concentration of the ligand needed for a 50% increase in the static light scattering intensity, caused by NCP aggregation. In PA, it is C_{50%}, which reflects the cation concentration at 50% precipitation of NCPs. Removal of the histone tails decreases the tendency of the g-NCP to aggregate and leads to higher C_{50%} (PA) and EC₅₀ (SLS). **Table 7-2** tabulates both max precipitation and C_{50%} value to reach such precipitation in PA experiments.

Table 7-2 Summary of precipitation assay results. Cation concentrations at 50% NCP precipitation (C_{50%}) from at least 3 independent measurements are shown, with maximal degree of NCP precipitation (Max) also indicated, adapted from Table 2 of (Liu, Lu et al. 2011).

Cations Conditions		Mg ²⁺		Ca ²⁺		CoHex ³⁺	
		C _{50%} (mM)	Max	C _{50%} (mM)	Max	C _{50%} (μM)	Max
wt-NCP	Tris	1.4±0.2	98%	2.0±0.01	97%	37.2±1.9	97%
	Tris + KCl	2.5±0.1	96%	2.1±0.1	95%	59.5±15.9	91%
g-NCP	Tris	20.9±4.4	50.5%	13.0±1.3	60%	66.6±1.9	100%
	Tris + KCl	25.0 (C _{39.5%})	39.5%	24.0±6.9	55%	127±5.9	99%
147 bp DNA	Tris	44.7 (C _{23.6%})	23.6%	81.5 (C _{23%})	23%	94.1±0.7	99%

Remarkably, the addition of the divalent cations Mg²⁺ and Ca²⁺ still leads to 50-60% precipitation of the g-NCP and to the formation of aggregates as observed in PA, although with lower scattering intensity than wt-NCP. Precipitation of g-NCPs induced by the trivalent CoHex³⁺ is the same as the wt-NCP, even slightly greater, with doubled value of C_{50%}. It therefore appears that NCP-NCP stacking in itself may favorably contribute to g-NCP attraction in the presence of divalent cations (perhaps with some small contribution from remaining stretches of the histone tails).

7.2.2 Comparison of Experiment and Simulation

There are considerable differences between conditions in the simulation and in the test tube. In simulations, the NCP concentration C_p (the DNA phosphate concentration) is 77 mM and the constant small amount of added monovalent salt about 5 mM, which corresponds to a ratio $C_{\text{salt}}/C_p = 0.06$. In the experiments, the NCP phosphate concentration is much lower ($C_p = 0.162$ mM and 0.010-0.050 mM for the PA and SLS methods respectively). Additionally, there is a huge excess of added salt relative to the net NCP concentration. For solutions with 10 mM KCl, the ratio C_{salt}/C_p is approximately equal to 60 for the PA and up to 1000 for the SLS experiments (at $C_p = 10$ μ M). Even in the absence of KCl, the ionic strength contributed from Tris buffer is large compared to the NCP concentration (approximately 8 mM Tris⁺ cations in the 10 mM Tris·HCl buffer at pH 7.5).

It is not possible to perform simulations with large box size and large number of particles that would correspond to these conditions. Consequently, it is not possible to compare the absolute cation concentrations that lead to pronounced attraction in the simulations, with $C_{50\%}$ concentrations characterizing the precipitated state in the experiments. Therefore, the result of the simulations should be analyzed on their own and only the general qualitative trends can be compared with the experimental results.

In g-NCP systems, the charge density of DNA wrapped around the histone octamer core, is reduced as compared to naked DNA. DNA has a higher negative charge compared to the tailless NCP; where the net positive charge (+92e) of the tailless octamer effectively reduces a substantial fraction of the net charge of the 147 bp DNA (-294e). In spite of lower charge density to DNA, based on polyelectrolyte behavior, the tailless g-NCP system is not expected to display any significant attraction even in the presence of divalent ions that would otherwise lead to aggregation.

The simulations of rNCP exhibit typical polyelectrolyte behavior as manifested first by the inability of monovalent cations to induce aggregation and the considerably stronger attractions between NCPs in the presence of trivalent CoHex³⁺ than divalent ions like Mg²⁺. And second, as shown in the simulation results for g-NCP, the much-reduced attraction in the absence of positively charged tails was expected. The first behavior is consistent with experimental results while the second one is in contrast to the experiment,

which shows g-NCP aggregation to the same level as rNCP under the presence of divalent and trivalent ions.

The absence of tails nullifies the contribution from the attractive tail bridging to the interaction between NCPs. Its effects shall be extrapolating from previous results on a fully acetylated NCP, described in “4.2.1 Modification of Histone Tails Weakens NCP-NCP Interaction”. The discrepancy between simulation and experiment has exposed the possibility and capability of another kind of short ranged specific non-tail-related nucleosome-nucleosome interactions, contributing to the aggregation of NCPs, and perhaps also resulting in stacked columnar phases observed by Livolant and co-workers for wt-NCP, see (Leforestier and Livolant 1997; Leforestier, Dubochet et al. 2001).

For example, most NCP crystals report a structure of the NCP1 domain of the domain H4 K16-R23 located between the surfaces of the two nucleosomes and interacting with the H2A acidic patch of the NCP2. The present coarse-grained model lacks the detailed description of the geometry and charge distribution in the NCP and as such cannot account for a number of specific effects such as charge-charge, hydrogen bonding and hydrophobic interactions.

7.2.3 Comparison of Simulations: Tailless Nucleosome Array versus g-NCP

Simulations of tailless nucleosome array with Mg^{2+} achieve 50-60% in the core-core contact number of the wt-nucleosome array, and almost the same contact number with $CoHex^{3+}$, tabulated in **Table 6-5(B)(C)** and **(D)(E)**; naturally, more cations are needed for tailless array to achieve such aggregation. The difference in array and NCP aggregation behavior could be due to linker DNA (discussed above in “6.3.5 Cation-DNA RDF Cumulative Number”). Linker DNA also imposes maximum inter-core distance of 10nm, which is the sum of five DNA beads diameters. These beads are highly negatively charged and attractive to cations. This could have made ion-ion correlations play a major role in inducing array self-association in the substitution role of tails. In contrast, tailless NCPs are not subject to these constraints and conditions; therefore, they, in simulation, follow polyelectrolyte theory to form no aggregates.

7.3 Conclusion

Tailless NCPs in the Sphere-Bead model cannot form aggregates even with multivalent cations because of the absence of tail for tail bridging mechanism and charge neutralization. This is an expected behavior following polyelectrolyte theory. Comparably, tailless nucleosome array can form aggregates because of the presence of negatively charged linker DNA.

Contrary to our simulation results and expectation, experiments (Liu, Lu et al. 2011), using well-defined nucleosome core particles prepared by recombinant methods, revealed that g-NCPs still can aggregate in the presence of oligocations. The tail-independent nucleosome-nucleosome attraction and DNA-DNA attraction became conspicuous and effective to nucleosome aggregation.

The discrepancy between experiment and simulation in the aggregation of tailless NCP for multivalent cations highlights the need for a more refined NCP model. The importance of detailed description of geometry and charge distribution to aggregation becomes more significant when tails are absent. Yet, the major roles of unspecific long-range electrostatics from multivalent cations and histone tail mediation in NCP aggregation processes are confirmed again.

8 Summary and Conclusion

In this work, we employed two models of the nucleosome array in dielectric continuum to simulate two systems: NCPs and single nucleosome array, with a variety of conditions relevant to the experiments. The purpose is to elucidate how salts of different valences and concentration, and tails with different modification, can influence the NCP aggregation or nucleosome array folding behaviors.

Results from simulations with the Sphere model for NCPs, presented in “4 NCP Simulation with The Sphere Model”, showed the better aggregation capability of divalent cations over monovalent cations, and how tail modification towards less positive charges on them reduces the aggregation. The artificial acetylation locations are not effective as the *in vivo* locations. We also elucidated that the key to tail-bridging mechanism is, on the contrary to intuition, the strong attraction between tail and core and how entropy drives the process of migration to the foreign core for tails and this only happens when two nucleosomes are in close distance due to ion fluctuation attraction.

In “6 Array Simulation with The Sphere-Bead Model”, we simulated nucleosome array with the Sphere-Bead model in the presence of monovalent (like K^+ or Na^+), divalent (Mg^{2+} , Ca^{2+}), sphere-like trivalent ($Co(NH_3)_6^{3+}$) as well as tri- and tetra-cationic polyamine-like particles (Spermidine³⁺ and Spermine⁴⁺). We mainly explored the different behavior of array under different types of salt and aimed to establish the ionic conditions that lead to a compacted state. Monovalent cation is not able to induce array folding while multivalence cations are able to. And tailless nucleosome array can induce folding with higher concentration than that of wild-type nucleosome array. Good agreement between calculated sedimentation velocity coefficients and the experimental results was obtained. In agreement with our findings with NCP, ion fluctuation attraction and tail bridging are still key to the array folding mechanisms. The inclusion of linker DNA has made ion fluctuation attraction more pronounced. Therefore, an important conclusion is that the theoretical description of array folding caused by multivalent cations must use a model including ion correlation. This is achieved by incorporating explicit mobile ions, while simplifications of the ionic environment by using a Debye-Hückel treatment is inadequate and may lead to a misunderstanding of the nature of the electrostatic mechanism that drives this important structural transition.

It the last part of the results “7 Tailless NCP Simulation with The Sphere-Bead Model ”, we attempted to match simulations of tailless NCP with experiments. g-NCP in the Sphere-Bead model with multivalent ions cannot form aggregates due to insufficient screening following typical polyelectrolyte behavior and lack of tail bridging mechanism. The unexpected mismatch with experiments is due to pronounced tail-independent specific nucleosome-nucleosome interaction and DNA-DNA attraction in the presence of cations. Non-electrostatic short-range interactions, i.e. nucleosome-nucleosome stacking, hydrogen bonding, water release etc., and very specific interactions with electrostatic origin, i.e. H4 to H2A acidic patch are actually not within the scope of modeling here. Simulation results re-confirm the role of unspecific long-range electrostatics and tail bridging assisted by multivalent ions fluctuation attraction for aggregation. The comparison between tailless gNCP and tailless array has confirmed that the view of how linker DNA helps tailless array fold.

Pertaining to the two points concerning salt and tail raised in the beginning “1.6.2 Our Approach”, our simple models and exhaustive simulations have successfully been able to reproduce experimental observations, especially the compaction with multivalent cations for both NCP and nucleosome array, and the reduced aggregation arising from tail modifications which reduces positive charges on them. We are also able to provide insight to the mechanisms from mobile cation distribution, and the influence of entropically “tail-bridging” mechanism.

We study the aggregation of nucleosome array not by simulating the aggregation of several arrays because it is very computationally demanding; instead, aggregation of the NCP system, while being of interest in its own right, can shed light on chromatin inter-array fiber formation since the aggregation of isolated (linker free) NCPs bear considerable resemblance to the inter-array aggregation process.

In order to include tail-independent NCP-NCP interaction, model shall include adequate details in shape and charge. This will also lead to fiber formation simulations. These could be areas for further research and development.

Appendix A: A Primer to Molecular Dynamics Simulation

In recent years, computational methods have been widely applied to solve problems in structural biology. Molecular Dynamics (“MD”) is a simulation technique to investigate the motions of atoms and molecules subject to well-defined force and equation of motion with coarse-grained or full atomic model. MD’s application in biology includes peptide/protein folding, inter-molecular interaction, briefly reviewed in (Karplus and McCammon 2002).

In many ways, performing MD simulation is like performing real experiment. When we perform real experiment, we prepare a sample of the material to be studied and subject them to certain conditions and use instruments to measure their properties during a certain time interval. If statistical noise interferes the measurement, the longer we average, the more accurate our measurement becomes. In molecular dynamics simulation, we follow the same approach: first, we prepare the “sample”, a system composed of N particles based on certain model, and define how they interact with each other; then we solve for Newton’s equations of motion for this system until the properties of the system no longer change with time, implying reaching convergence of equilibrium. Then we take the measurement of properties during equilibrium phase.

In theoretical terms, MD generates an ensemble of the system under certain thermodynamic conditions by sampling on the free energy surface. The motion of the particles is recorded in the trajectory file. Given sufficient long simulation time, based on Ergodic hypothesis, the trajectory obtained represents the whole phase and we can extract the “stable” or “native” conformation from the trajectory.

MD refers to the simulation method using Newton's equation of motion to numerically solve for the trajectories of molecules and atoms. Langevin dynamics (or, “Stochastic Dynamics” because of the use of stochastic noise in the equation of motion) uses another form of equation of motion approach to mathematically model the dynamics of molecular systems, originally developed by French physicist Paul Langevin.

We use SD to omit explicit water molecules surrounding the molecule to be studied because water-water interaction and water-molecule interaction make up for a large portion of calculation in full atomic MD simulations. Langevin dynamics attempts to extend molecular dynamics to mimics the viscous aspect of a solvent. Also, Langevin dynamics allows controlling the temperature like a thermostat, thus approximating the canonical ensemble. SD does not normally account for the electrostatic screening and also not for the hydrophobic effect. Note that in our work, we include mobile explicit ions and water relative permittivity to count for screening effects,

Here is a brief introduction of MD and SD adapted from manual (van der Spoel, Lindahl et al. 2005b) of GROMACS (Lindahl, Hess et al. 2001; van der Spoel, Lindahl et al. 2005a). GROMACS is a very fast open source MD software package, comprising force fields, and dynamics calculation engine and analysis tools, most suitable for full atomic simulation. We used ESPResSo (Limbach, Arnold et al. 2006) as dynamics calculation engine for our coarse-grain models; and analysis tools from GROMACS in this work.

A.1 Potential Energy Equation

The force acting on particle i is the gradient of the potential energy:

$$F_i = -\frac{\partial}{\partial r_i} V(r_1, \dots, r_N)$$

where r_i is the position of the particle m_i . F_i is the force depending on the position of all (N) particles in the system, and V is the potential energy function by which the system is described.

The potential energy function in Class I molecular mechanics force field ("MMFF")¹ is generally described as

¹ Class I FF employs a simple form of potential energy function, which limits transferability. Class II FF employs extended potential energy function, including cross terms, increases transferability. Class II FF designed to be a transferable force field for pharmaceutical compounds that accurately treats conformational energetics and non-bonded interactions but with increased difficulty during parameterization. Transferability

$$V(r_i) = V(r_1, \dots, r_N) = \sum V_{\text{bonded}}(r) + \sum V_{\text{nonbonded}}(r)$$

This can be expanded as below, defined in the CHARMM force field (a Class I MMFF).

$$V(r) = \left\{ \sum \frac{k_{\text{bond}}}{2} (r - r_0)^2 + \sum \frac{k_{\theta}}{2} (\theta - \theta_0)^2 + \dots \right\} V_{\text{bonded}} \\ + \left\{ \sum_{i < j} \left(4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{r_{ij}} \right) \right\} V_{\text{nonbonded}}$$

The first term in the function is the bond stretching energy (harmonic potential) along bond r . Vibration constant k_{bond} and minimum energy bond length r_{bond} of two bonded atoms are supplied by the force field. The second term is the bond angle (θ) bending energy and has a similar form. There are more terms to describe four-body or n-body interactions. These terms make up for bonded potential energy V_{bonded} . The last term in the equation is non-bonded potential energy V_{unbonded} over all pairs of atoms. It is composed of van der Waals interaction and Coulomb electrostatic interaction between atom i and j with charges q_i and q_j separated by a distance r_{ij} . Van der Waals interactions' decay along reaction distance is much faster than electrostatic interaction and thus much easier to be dealt in the computation with a neighborhood list with spherical cut-off at a certain distance (usually at 10Å). Because of its long-range effectiveness, electrostatic interaction is hard to be dealt with and is also computationally expensive. Achieving accurate electrostatic interaction representation in simulation is important and even critical, such as for highly charged molecules, e.g. nucleic acid, in bio-molecular simulations. Among many algorithms to deal with electrostatic interaction, Ewald summation and its more efficient implementations such as P3M (Particle-Particle Particle-Mesh) and PME (Particle Mesh Ewald) are able to accurately compute electrostatic field with infinite boundary conditions.

A.2 Molecular Dynamics/Stochastic Dynamics

The “governing” equation in MD is Newton’s equation of motion,

is the ability of applying empirical force field parameters to molecules not explicitly included during the parameter optimization.

$$m_i \frac{d^2 r_i(t)}{dt^2} = F_i(r_1, \dots, r_N)$$

Langevin dynamics adds a friction term and a random noise term,

$$m_i \frac{d^2 r_i(t)}{dt^2} = F_i(r_1, \dots, r_N) - m_i g_i \frac{dr_i(t)}{dt} + \sqrt{2m_i g_i k_B T} \overset{\circ}{r}_i$$

where $g([1/ps])$ is the friction constant and $\overset{\circ}{r}_i(t)$ is the noise process satisfying $\langle \overset{\circ}{r}_i(t) \rangle = 0$ and $\langle \overset{\circ}{r}_i(t) \overset{\circ}{r}_i(t+s) \rangle = \delta(t-t')$. When $1/g$ is large compared to the time scales present in the system, one could see stochastic dynamics as molecular dynamics with stochastic temperature coupling. If simulating a system in vacuum, there is additional advantage that there is no accumulation of errors for the overall translational and rotational degrees of freedom. When $1/g$ is small compared to the time scales present in the system The dynamics will be completely different from MD, but the sampling is still correct.

Following flowchart is to show how these equations cascades with each other to form a loop. MD/SD has a repetitive process that, during each loop, forces are calculated and integrated to generate new velocities so as to update the positions of atoms after one time step; updated position and velocity information are fed into the next loop. Time proceeds with step-wise additions of time steps. For every hundreds of time steps (equivalent to $1ps$ normally), positions of the atoms are stored in a *frame*. Frames along the time axis form a *trajectory*.

I. Input

Positions $\mathbf{r}$ and velocity $\mathbf{v}$ of all atoms in the system

Interaction potential V as a function of atom positions

II. Forces

The force on any atom

$$F_i = -\frac{\partial V}{\partial r_i} = -\frac{\partial}{\partial r_i} \left(\sum V_{\text{bonded}}(r) + \sum V_{\text{nonbonded}}(r) \right)$$

can be computed by calculating the force between non-bonded atom pairs.

$$F_i = \sum F_{ij} = \sum \left(4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{r_{ij}} \right)$$

plus the forces due to bonded interactions.

Potential/kinetic energies and the pressure are computed.

III. Update

The movement of the atoms is simulated by numerically solving

Newton's equation of motion

$$\frac{d^2 r_i}{dt^2} = \frac{F_i}{m_i}, \text{ or}$$

Langevin Molecular Dynamics equation of motion

$$m_i \frac{d^2 r_i(t)}{dt^2} = F_i(r_1, \dots, r_N) - m_i g_i \frac{d^2 r_i(t)}{dt^2} + r_i$$

Velocity/positions are scaled according to pressure/temperature conditions.

IV. Output

Positions, Velocities, Energies, Temperature, Pressure, etc.

↑ Go back to II with updated positions and velocities.

Repeat this for each time step until total time step has been reached.

Assumption

Set of configurations (ensemble) is statistically representative for the whole phase space

Statistical Mechanics

(Microscopic behavior) ➔ (Macroscopic properties)

Statics: e.g., positions, distances, orientations, correlations

Dynamics: e.g. diffusion, frequencies, relaxation times

Thermodynamics: e.g. density, pressure, viscosity, energy

Reference

1. Allahverdi, A., R. Yang, N. Korolev, Y. Fan, C. A. Davey, C. F. Liu and L. Nordenskiöld (2011). "The Effects of Histone H4 Tail Acetylations on Cation-Induced Chromatin Folding and Self-Association." Nucleic Acids Research **39**(5): 1680-1691.
2. Arents, G. and E. N. Moudrianakis (1993). "Topography of the Histone Octamer Surface: Repeating Structural Motifs Utilized in the Docking of Nucleosomal DNA." Proceedings of the National Academy of Sciences of the United States of America **90**: 10489-10493.
3. Arents, G. and E. N. Moudrianakis (1995). "The Histone Fold: A Ubiquitous Architectural Motif Utilized in DNA Compaction and Protein Dimerization." Proceedings of the National Academy of Sciences of the United States of America **92**(24): 11170-11174.
4. Arya, G. and T. Schlick (2006). "Role of Histone Tails in Chromatin Folding Revealed by a Mesoscopic Oligonucleosome Model." Proceedings of the National Academy of Sciences of the United States of America **103**(44): 16236-16241.
5. Arya, G. and T. Schlick (2009). "A Tale of Tails: How Histone Tails Mediate Chromatin Compaction in Different Salt and Linker Histone Environments." Journal of Physical Chemistry A **113**(16): 4045-4059.
6. Arya, G., Q. Zhang and T. Schlick (2006). "Flexible Histone Tails in a New Mesoscopic Oligonucleosome Model." Biophysical journal **91**: 133-150.
7. Ausio, J. (2000). "Analytical Ultracentrifugation and the Characterization of Chromatin Structure." Biophysical Chemistry **86**: 141-153.
8. Baker, N. A., D. Sept, S. Joseph, M. J. Holst and J. A. McCammon (2001). "Electrostatics of Nanosystems: Application to Microtubules and the Ribosome." Proceedings of the National Academy of Sciences of the United States of America **98**(18): 10037-10041.
9. Barbera, A. J., J. V. Chodaparambil, B. Kelley-Clarke, V. Joukov, J. C. Walter, K. Luger and K. M. Kaye (2006). "The Nucleosomal Surface as a Docking Station for Kaposi's Sarcoma Herpesvirus Lana." Science **311**: 856-861.
10. Beard, D. A. and T. Schlick (2001). "Computational Modeling Predicts the Structure and Dynamics of Chromatin Fiber." Structure **9**: 105-114.
11. Bertin, A., D. Durand, M. Renouard, F. Livolant and S. Mangenot (2007). "H2a and H2b Tails Are Essential to Properly Reconstitute Nucleosome Core Particles." European Biophysics Journal **36**(8): 1083-1094.
12. Bertin, A., A. Leforestier, D. Durand and F. Livolant (2004). "Role of Histone Tails in the Conformation and Interaction of Nucleosome Core Particles." Biochemistry **43**: 4773-4780.
13. Bertin, A., S. Mangenot, M. Renouard, D. Durand and F. Livolant (2007). "Structure and Phase Diagram of Nucleosome Core Particles Aggregated by Multivalent Cations." Biophysical journal **93**: 3652-3663.

14. Bertin, A., M. Renouard, J. S. Pedersen, F. Livolant and D. Durand (2007). "H3 and H4 Histone Tails Play a Central Role in the Interactions of Recombinant Ncps." Biophysical journal **92**: 2633-2645.
15. Bloomfield, V., W. O. Dalton and K. E. Van Holde (1967). "Frictional Coefficients of Multisubunit Structures. I. Theory." Biopolymers **5**(2): 135-148.
16. Bloomfield, V. A. (1997). "DNA Condensation by Multivalent Cations." Biopolymers **44**: 269-282.
17. Boeger, H., J. Griesenbeck, J. S. Strattan and R. D. Kornberg (2003). "Nucleosomes Unfold Completely at a Transcriptionally Active Promoter." Molecular Cell **11**: 1587-1598.
18. Calestagne-Morelli, A. and J. Ausio (2006). "Long-Range Histone Acetylation: Biological Significance, Structural Implications, and Mechanisms." Biochemistry and Cell Biology **84**(4): 518-527.
19. Cherstvy, A. G. (2009). "Positively Charged Residues in DNA-Binding Domains of Structural Proteins Follow Sequence-Specific Positions of DNA Phosphate Groups." Journal of Physical Chemistry B **113**(13): 4242-4247.
20. Cheung, P., C. D. Allis and P. Sassone-Corsi (2000). "Signaling to Chromatin through Histone Modification." Cell **103**: 263-271.
21. Chodaparambil, J. V., A. J. Barbera, X. Lu, K. M. Kaye, J. C. Hansen and K. Luger (2007). "A Charged and Contoured Surface on the Nucleosome Regulates Chromatin Compaction." Nature Structural & Molecular Biology **14**(11): 1105-1107.
22. Claesson, P. M., E. Poptoshev, E. Blomberg and A. Dedinaite (2005). "Polyelectrolyte-Mediated Surface Interactions." Advances in Colloid Interface Science **114-115**: 173-187.
23. Clark, D. J. and T. Kimura (1990). "Electrostatic Mechanism of Chromatin Folding." Journal of Molecular Biology **211**: 883-896.
24. Davey, C. A., D. F. Sargent, K. Luger, A. W. Maeder and T. J. Richmond (2002). "Solvent Mediated Interactions in the Structure of Nucleosome Core Particle at 1.9 Å Resolution." Journal of molecular biology **319**: 1097-1113.
25. de Frutos, M., E. Raspaud, A. Leforestier and F. Livolant (2001). "Aggregation of Nucleosomes by Divalent Cations." Biophysical journal **81**: 1127-1132.
26. Deserno, M. and C. Holm (1998a). "How to Mesh up Ewald Sums. I. A Theoretical and Numerical Comparison of Various Particle Mesh Routines." Journal of Chemical Physics **109**(18): 7678-7693.
27. Deserno, M. and C. Holm (1998b). "How to Mesh up Ewald Sums. Ii. An Accurate Error Estimate for the Particle-Particle-Particle-Mesh Algorithm." Journal of Chemical Physics **109**(18): 7694-7701.
28. Dorigo, B., T. Schalch, K. Bystricky and T. J. Richmond (2003). "Chromatin Fiber Folding: Requirement for the Histone H4 N-Terminal Tail." Journal of molecular biology **327**: 85-96.
29. Dorigo, B., T. Schalch, A. Kulangara, S. Duda, R. R. Schroeder and T. J. Richmond (2004). "Nucleosome Arrays Reveal the Two-Start Organization of the Chromatin Fiber." Science **306**: 1571-1573.

30. Dou, Y. and M. A. Gorovsky (2000). "Phosphorylation of Linker Histone H1 Regulates Gene Expression in Vivo by Creating a Charge Patch." Molecular Cell **6**: 225-231.
31. Dou, Y. and M. A. Gorovsky (2002). "Regulation of Transcription by H1 Phosphorylation in Tetrahymena Is Position Independent and Required Clustered Sites." Proceedings of the National Academy of Sciences of the United States of America **99**(9): 6142-6146.
32. Dyer, P. N., R. S. Edayathumangalam, C. L. White, Y. Bao, S. Chakravarthy, U. M. Muthurajan and K. Luger (2004). "Reconstitution of Nucleosome Core Particles from Recombinant Histones and DNA." Methods in Enzymology **375**: 23-44.
33. Finch, J. T., R. S. Brown, T. Richmond, B. Rushton, L. C. Lutter and A. Klug (1981). "X-Ray Diffraction Study of a New Crystal Form of the Nucleosome Core Showing Higher Resolution." Journal of Molecular Biology **145**(4): 757-769.
34. Finch, J. T. and A. Klug (1976). "Solenoidal Model for Superstructure in Chromatin." Proceedings of the National Academy of Sciences of the United States of America **73**(6): 1897-1901.
35. Finch, J. T., L. C. Lutter, D. Rhodes, R. S. Brown, B. Rushton, M. Levitt and A. Klug (1977). "Structure of Nucleosome Core Particles of Chromatin." Nature **269**(5623): 29-36.
36. Garcia-Ramirez, M., F. Dong and J. Ausio (1992). "Role of the Histone "Tails" in the Folding of Oligonucleosomes Depleted of Histone H1." Journal of Biological Chemistry **267**(27): 19587-19595.
37. Gelbart, W. M., R. F. Bruinsma, P. A. Pincus and V. A. Parsegian (2000). "DNA-Inspired Electrostatics." Physics Today **53**(9): 38-44.
38. Gordon, F., K. Luger and J. C. Hansen (2005). "The Core Histone N-Terminal Tail Domains Function Independently and Additively During Salt-Dependent Oligomerization of Nucleosomal Arrays." Journal of Biological Chemistry **280**(40): 33701-33706.
39. Grest, G. S. and K. Kremer (1986). "Molecular Dynamics Simulation for Polymers in the Presence of a Heat Bath." Physical Review A **33**(5): 3628-3631.
40. Grigoryev, S. A., G. Arya, S. Correll, C. L. Woodcock and T. Schlick (2009). "Evidence for Heteromorphic Chromatin Fibers from Analysis of Nucleosome Interactions." Proceedings of the National Academy of Sciences of the United States of America **106**(32): 13317-13322.
41. Hansen, J. C. (2002). "Conformational Dynamics of the Chromatin Fiber in Solution: Determinants, Mechanisms, and Functions." Annual Review of Biophysics and Biomolecular Structure **31**: 361-392.
42. Hansen, J. C., J. Ausio, V. H. Stanik and K. E. van Holde (1989). "Homogeneous Reconstituted Oligonucleosomes, Evidence for Salt-Dependent Folding in the Absence of Histone H1." Biochemistry **28**: 9129 - 9136.
43. Hansen, J. C., C. Tse and A. P. Wolffe (1998). "Structure and Function of the Core Histone N-Terminal: More Than Meets the Eye." Biochemistry **37**(51): 17637-17641.
44. Harp, J. M., B. L. Hanson, D. E. Timm and G. J. Bunick (2000). "Asymmetries in the Nucleosome Core Particle at 2.5 Å Resolution." Acta Crystallographica Section D **56**: 1513-1534.

45. Horn, P. J. and C. L. Peterson (2002). "Chromatin Higher Order Folding: Wrapping up Transcription." Science **297**: 1824-1827.
46. Huynh, V. A. T., P. J. J. Robinson and D. Rhodes (2005). "A Method for the in Vitro Reconstitution of a Defined "30 Nm" Chromatin Fibre Containing Stoichiometric Amounts of the Linker Histone." Journal of molecular biology **345**: 957-968.
47. Iwaki, T., T. Saito and K. Yoshikawa (2007). "How Are Small Ions Involved in the Compaction of DNA Molecules?" Colloids and Surfaces B: Biointerfaces **56**(1-2): 126-133.
48. Jiang, L., J. N. Smith, S. L. Anderson, P. Ma, C. A. Mizzen and N. L. Kelleher (2007). "Global Assessment of Combinatorial Post-Translational Modification of Core Histones in Yeast Using Contemporary Mass Spectrometry. Lys4 Trimethylation Correlates with Degree of Acetylation on the Same H3 Tail." Journal of Biological Chemistry **282**(38): 27923-27934.
49. Kan, P.-Y., T. L. Caterino and J. J. Hayes (2009). "The H4 Tail Domain Participates in Intra- and Internucleosome Interactions with Protein and DNA During Folding and Oligomerization of Nucleosome Arrays." Molecular and Cellular Biochemistry **29**(2): 538-546.
50. Kan, P.-Y. and J. J. Hayes (2007). "Detection of Interactions between Nucleosome Arrays Mediated by Specific Core Histone Tail Domains." Methods **41**(3): 278-285.
51. Kan, P.-Y., X. Lu, J. C. Hansen and J. J. Hayes (2007). "The H3 Tail Domain Participates in Multiple Interactions During Folding and Self-Association of Nucleosome Arrays." Molecular and Cellular Biology **27**(6): 2084-2091.
52. Karplus, M. and J. A. McCammon (2002). "Molecular Dynamics Simulations of Biomolecules." Nature structural biology **9**(9): 646-652.
53. Kepper, N., D. Foethke, R. Stehr, G. Wedemann and K. Rippe (2008). "Nucleosome Geometry and Internucleosomal Interactions Control the Chromatin Fiber Conformation." Biophysical journal **95**: 3692-3705.
54. Khorasanizadeh, S. (2004). "The Nucleosome: From Genomic Organization to Genomic Regulation." Cell **116**: 259-272.
55. Kirkwood, J. G. (1954). "The General Theory of Irreversible Processes in Solutions of Macromolecules." Journal of Polymer Science **12**(1): 14.
56. Korolev, N., A. Allahverdi, Y. Yang, Y. Fan, A. P. Lyubartsev and L. Nordenskiöld (2010). "Electrostatic Origin of Salt-Induced Nucleosome Array Compaction." Biophysical journal **99**(6): 1896-1905.
57. Korolev, N., N. V. Berezhnoy, K. D. Eom, J. P. Tam and L. Nordenskiöld (2009). "A Universal Description for the Experimental Behavior of Salt-(in)Dependent Oligocation-Induced DNA Condensation." Nucleic Acids Research **37**(21): 7137-7150.
58. Korolev, N., A. P. Lyubartsev and L. Nordenskiöld (2006). "Computer Modeling Demonstrates That Electrostatic Attraction of Nucleosomal DNA Is Mediated by Histone Tails." Biophysical journal **90**: 4305-4316.
59. Korolev, N., A. P. Lyubartsev, L. Nordenskiöld and A. Laaksonen (2001a). "Spermine: An "Invisible" Component in the Crystals of B-DNA. A Grand Canonical Monte Carlo and Molecular Dynamics Simulation Study." Journal of molecular biology **308**: 907-917.

60. Korolev, N., A. P. Lyubartsev, A. Rupprecht and L. Nordenskiöld (1999). "Competitive Binding of Mg^{2+} , Ca^{2+} , Na^+ , and K^+ to DNA in Oriented DNA Fibers: Experimental and Monte Carlo Simulation Results." Biophysical journal **77**(5): 2736-2749.
61. Korolev, N., A. P. Lyubartsev, A. Rupprecht and L. Nordenskiöld (2001b). "Competitive Substitution of Hexamine Cobalt(III) for Na^+ and K^+ Ions in Oriented DNA Fibers." Biopolymers **58**: 268-278.
62. Kurdistan, S. K., S. Tavazole and M. Grunstein (2004). "Mapping Global Histone Acetylation Patterns to Gene Expression." Cell **117**: 721-733.
63. Langowski, J. and D. W. Heermann (2007). "Computational Modeling of the Chromatin Fiber." Seminars in Cell & Developmental Biology **18**(5): 659-667.
64. Leforestier, A., J. Dubochet and F. Livolant (2001). "Bilayers of Nucleosome Core Particles." Biophysical journal **81**: 2114-2421.
65. Leforestier, A. and F. Livolant (1997). "Liquid Crystalline Ordering of Nucleosome Core Particles under Macromolecular Crowding Conditions: Evidence for a Discotic Columnar Hexagonal Phase." Biophysical journal **73**(4): 1771-1776.
66. Lewis, P. N., J. G. Guilleminette and S. Chan (1988). "Histone Accessibility Determined by Lysine-Specific Acetylation in Chicken Erythrocyte Nuclei." European Journal of Biochemistry **172**: 135-145.
67. Limbach, H. J., A. Arnold, B. A. Mann and C. Holm (2006). "Espresso - an Extensible Simulation Package for Research on Soft Matter Systems." Computer Physics Communications **174**(9): 704-727.
68. Limbach, H. J. and C. Holm (2001). "End Effects of Strongly Charged Polyelectrolytes: A Molecular Dynamics Study." Journal of Chemical Physics **114**: 9674-9682.
69. Lindahl, E., B. Hess and D. van der Spoel (2001). "Gromacs 3.0: A Package for Molecular Simulation and Trajectory Analysis." Journal of Molecular Modeling **7**: 306-317.
70. Liu, Y., C. Lu, Y. Yang, Y. Fan, R. Yang, C. F. Liu, N. Korolev and L. Nordenskiöld (2011). "Influence of Histone Tails and H4 Tail Acetylations on Nucleosome-Nucleosome Interactions." Journal of molecular biology **414**(5): 749-764.
71. Lowary, P. T. and J. Widom (1998). "New DNA Sequence Rules for High Affinity Binding to Histone Octamer and Sequence-Directed Nucleosome Positioning." Journal of molecular biology **276**: 19-42.
72. Luger, K. (2003). "Structure and Dynamic Behavior of Nucleosomes." Current Opinion in Genetics & Development **13**: 127-135.
73. Luger, K. and J. C. Hansen (2005). "Nucleosome and Chromatin Fiber Dynamics." Current Opinion in Structural Biology **15**: 188-196.
74. Luger, K., A. W. Mader, R. K. Richmond, D. F. Sargent and T. J. Richmond (1997). "Crystal Structure of the Nucleosome Core Particle at 2.8 Å Resolution." Nature **389**(6648): 251-260.
75. Luger, K., T. J. Rechsteiner and T. J. Richmond (1999). "Preparation of Nucleosome Core Particle from Recombinant Histones." Methods in Enzymology **304**: 3-19.

76. Luger, K. and T. J. Richmond (1998). "The Histone Tails of the Nucleosome." Current Opinion in Genetics & Development **8**: 140-146.
77. Lyubartsev, A. P. (2005). "Multiscale Modeling of Lipids and Lipid Bilayers." European Biophysics Journal **35**: 53-61.
78. Lyubartsev, A. P. and L. Nordenskiöld (1995). "Monte Carlo Simulation Study of Ion Distribution and Osmotic Pressure in Hexagonally Oriented DNA." Journal of Physical Chemistry **99**(25): 10373-10382.
79. Lyubartsev, A. P. and L. Nordenskiöld (1997). "Monte Carlo Simulation Study of DNA Polyelectrolyte Properties in the Presence of Multivalent Polyamine Ions." Journal of Physical Chemistry B **101**(21): 4335-4342.
80. Mangelot, S., A. Leforestier, D. Durand and F. Livolant (2003a). "Phase Diagram of Nucleosome Core Particles." Journal of molecular biology **333**: 907-916.
81. Mangelot, S., A. Leforestier, D. Durand and F. Livolant (2003b). "X-Ray Diffraction Characterization of the Dense Phases Formed by Nucleosome Core Particles." Biophysical journal **84**: 2570-2584.
82. Mangelot, S., A. Leforestier, P. Vachette, D. Durand and F. Livolant (2002). "Salt-Induced Conformation and Interaction Changes of Nucleosome Core Particles." Biophysical journal **82**: 345-356.
83. Mangelot, S., E. Raspaud, C. Tribet, L. Belloni and F. Livolant (2002). "Interactions between Isolated Nucleosome Core Particles. A Tail Bridging Effect?" European Physics Journal E **7**: 221-231.
84. McBryant, S. J., J. Klonoski, T. C. Sorensen, S. S. Norskog, S. Williams, M. G. Resch, J. A. Toombs, 3rd, S. E. Hobdey and J. C. Hansen (2009). "Determinants of Histone H4 N-Terminal Domain Function During Nucleosomal Array Oligomerization: Roles of Amino Acid Sequence, Domain Length, and Charge Density." Journal of Biological Chemistry **284**(25): 16716-16722.
85. Miklavic, S. J., C. E. Woodward, B. Jönsson and T. Åkesson (1990). "Interaction of Charged Surfaces with Grafted Polyelectrolytes: A Poisson-Boltzmann and Monte Carlo Study." Macromolecules **23**: 4149-4157.
86. Mizzen, C., M.-H. Kuo, E. Smith, J. Brownell, J. Zhou, R. Ohba, Y. Wei, L. Monaco, P. Sassone-Corsi and C. D. Allis (1998). "Signaling to Chromatin through Histone Modifications: How Clear Is the Signal?" Cold Spring Harbor Symposia on Quantitative Biology **58**: 469-481.
87. Montoro, J. C. G. and J. L. F. Abascal (1995). "Ionic Distribution around Simple DNA Models. I. Cylindrically Averaged Properties." Journal of Chemical Physics **103**(18): 8273-8284.
88. Muhlbacher, F. (2004). Tail Bridging Vs. Patchiness: Models for the Interactions of Nucleosomes, Johannes Gutenberg-Universität, Mainz. **PhD**.
89. Muhlbacher, F., C. Holm and H. Schiessel (2006). "Controlled DNA Compaction within Chromatin: The Tail-Bridging Effect." Europhysics Letters **73**: 135-141.
90. Muhlbacher, F., H. Schiessel and C. Holm (2006). "Tail-Induced Attraction between Nucleosome Core Particles." Physical Review E **74**: 031919.
91. Noll, M. (1974). "Subunit Structure of Chromatin." Nature **251**: 249-251.

92. Olins, D. E. and A. L. Olins (2003). "Chromatin History: Our View from the Bridge." Nature Reviews Molecular Cell Biology **4**(10): 809-814.
93. Oosawa, F. (1968). "Interaction between Parallel Rodlike Macroions." Biopolymers **6**: 1633-1647.
94. Pelta, J., D. Durand, J. Doucet and F. Livolant (1996). "DNA Mesophases Induced by Spermidine: Structural Properties and Biological Implications." Biophysical Journal **71**(1): 48-63.
95. Perry, M. and R. Chalkley (1981). "The Effect of Histone Hyperacetylation on the Nuclease Sensitivity and the Solubility of Chromatin." Journal of Biological Chemistry **256**(7): 3313-3318.
96. Perry, M. and R. Chalkley (1982). "Histone Acetylation Increases the Solubility of Chromatin and Occurs Sequentially over Most of the Chromatin. A Novel Model for the Biological Role of Histone Acetylation." Journal of Biological Chemistry **257**(13): 7336-7347.
97. Pesavento, J. J., C. R. Bullock, R. D. LeDuc, C. A. Mizzen and N. L. Kelleher (2008). "Combinatorial Modification of Human Histone H4 Quantitated by Two-Dimensional Liquid Chromatography Coupled with Top Down Mass Spectrometry." Journal of Biological Chemistry **283**(22): 14927-14937.
98. Podgornik, R. (2003). "Two-Body Polyelectrolyte-Mediated Bridging Interactions." Journal of Chemical Physics **118**: 11286-11296.
99. Podgornik, R. (2004). "Polyelectrolyte-Mediated Bridging Interactions." Journal of Polymer Science Part B **42**: 3539-3556.
100. Podgornik, R. and M. Licer (2006). "Polyelectrolyte Bridging Interactions between Charged Macromolecules." Current Opinion in Colloid & Interface Science **11**: 273-279.
101. Pollard, K. J., M. L. Samuels, K. A. Crowley, J. C. Hansen and C. L. Peterson (1999). "Functional Interaction between Gcn5 and Polyamines: A New Role for Core Histone Acetylation." EMBO Journal **18**(20): 5622-5633.
102. Raspaud, E., I. Chaperon, A. Leforestier and F. Livolant (1999). "Spermine-Induced Aggregation of DNA, Nucleosome, and Chromatin." Biophysical journal **77**(3): 1547-1555.
103. Raspaud, E., B. Pitard, D. Durand, O. Aquerre-Chariol, J. Pelta, G. Byk, D. Scherman and F. Livolant (2001). "Polymorphism of DNA/Multi-Cationic Lipid Complexes Driven by Temperature and Salt." Journal of Physical Chemistry B **105**: 5291-5297.
104. Ray, J. and G. S. Manning (1997). "Effect of Counterion Valence and Polymer Charge Density on the Pair Potential of Two Polyions." Macromolecules **30**(19): 5739-5744.
105. Razin, S. V., O. V. Iarovaia, N. Sjakste, T. Sjakste, L. Bagdoniene, A. V. Rynditch, E. R. Eivazova, M. Lipinski and Y. S. Vassetzky (2007). "Chromatin Domains and Regulation of Transcription." Journal of molecular biology **369**: 597-607.
106. Ren, Q. and M. A. Gorovsky (2001). "Histone H2a.Z Acetylation Modulates an Essential Charge Patch." Molecular Cell **7**: 1329-1335.
107. Ren, Q. and M. A. Gorovsky (2003). "The Nonessential H2a N-Terminal Tail Can Function as an Essential Charge Patch on the H2a.Z Variant N-Terminal Tail." Molecular Biology of the Cell **23**(8): 2778-2789.

108. Rescic, J. and P. Linse (2000). "Charged Colloidal Solutions with Short Flexible Counterions." Journal of Physical Chemistry B **104**: 7852-7857.
109. Richmond, T. J., J. T. Finch, B. Rushton, D. Rhodes and A. Klug (1984). "Structure of the Nucleosome Core Particle at 7 Å Resolution." Nature **311**: 532-537.
110. Robinson, P. J., L. Fairall, V. A. Huynh and D. Rhodes (2006). "Em Measurements Define the Dimensions of the "30-Nm" Chromatin Fiber: Evidence for a Compact, Interdigitated Structure." Proceedings of the National Academy of Sciences of the United States of America **103**(17): 6506-6511.
111. Robinson, P. J. J., W. An, A. Routh, F. Martino, L. Chapman, R. G. Roeder and D. Rhodes (2008). "30 Nm Chromatin Fibre Decompaction Requires Both H4-K16 Acetylation and Linker Histone Eviction." Journal of molecular biology **381**: 816-825.
112. Routh, A., S. Sandin and D. Rhodes (2008). "Nucleosome Repeat Length and Linker Histone Stoichiometry Determine Chromatin Fiber Structure." Proceedings of the National Academy of Sciences of the United States of America **105**(26): 8872-8877.
113. Schalch, T. (2004). The 30-Nm Chromatin Fiber: In Vitro Reconstitution and Structural Analysis. Swiss Federal Institute of Technology, Zurich. Zurich, ETH. **PhD**: 179.
114. Schalch, T., S. Duda, D. F. Sargent and T. J. Richmond (2005). "X-Ray Structure of a Tetranucleosome and Its Implications for the Chromatin Fibre." Nature **436**: 138-141.
115. Schrodinger, L. (2010). The Pymol Molecular Graphics System, Version 1.3r1.
116. Shogren-Knaak, M. A., H. Ishii, J.-M. Sun, M. Pazin, J. R. Davie and C. L. Peterson (2006). "Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions." Science **311**: 844-847.
117. Stehr, R., N. Kepper, K. Rippe and G. Wedemann (2008). "The Effect of Internucleosomal Interaction on Folding of the Chromatin Fiber." Biophysical journal **95**(8): 3677-3691.
118. Stevens, M. J. (2004). "Coarse Grained Simulations of Lipid Bilayers." Journal of Chemical Physics **121**: 11942-11948.
119. Stevens, M. J. and K. Kreker (1995). "The Nature of Flexible Linear Polyelectrolytes in Salt Free Solution: A Molecular Dynamics Study." Journal of Chemical Physics **103**: 1669-1690.
120. Strahl, B. D. and C. D. Allis (2000). "The Language of Covalent Histone Modifications." Nature **403**: 41-45.
121. Sun, J., Q. Zhang and T. Schlick (2005). "Electrostatic Mechanism of Nucleosomal Array Folding Revealed by Computer Simulation." Proceedings of the National Academy of Sciences of the United States of America **102**(23): 8180-8185.
122. Szerlong, H. J., J. E. Prenti, J. K. Nyborg and J. C. Hansen (2010). "Activator-Dependent P300 Acetylation of Chromatin in Vitro: Enhancement of Transcription by Disruption of Repressive Nucleosome-Nucleosome Interactions." Journal of Biological Chemistry **285**(42): 31954-31964.
123. Thoma, F., T. Koller and A. Klug (1979). "Involvement of Histone H1 in the Organization of the Nucleosome and of the Salt-Dependent Superstructures of Chromatin." Journal of Cell Biology **83**(2 Pt 1): 403-427.

124. Tse, C., T. Sera, A. P. Wolffe and J. C. Hansen (1998). "Disruption of Higher-Order Folding by Core Histone Acetylation Dramatically Enhances Transcription of Nucleosomal Arrays by Rna Polymerase Iii." Molecular and Cellular Biology **18**(8): 4629-4638.
125. Tumbar, T., G. Sudlow and A. S. Belmont (1999). "Large-Scale Chromatin Unfolding and Remodeling Induced by Vp16 Acidic Activation Domain." Journal of Cell Biology **145**(7): 1341-1354.
126. Turner, B. M. (1991). "Histone Acetylation and Control of Gene Expression." Journal of Cell Science **99**: 13-20.
127. Turner, B. M. (2002). "Cellular Memory and the Histone Code." Cell **111**: 285-291.
128. van der Spoel, D., E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen (2005a). "Gromacs: Fast, Flexible and Free." Journal of computational chemistry **26**: 1701-1718.
129. van der Spoel, D., E. Lindahl, B. Hess, A. R. van Buuren, E. Apol, P. J. Meulenhoff, D. P. Tieleman, A. L. T. M. Sijbers, K. A. Feenstra, R. van Drunen and H. J. C. Berendsen (2005b) Gromacs User Manual Version 3.3.
130. van Holde, K. and J. Zlatanova (1996). "What Determine the Folding of the Chromatin Fiber?" Proceedings of the National Academy of Sciences of the United States of America **93**(20): 10548-10555.
131. van Holde, K. E. (1988). Chromatin. New York, Springer-Verlag.
132. Walker, I. O. (1984). "Differential Dissociation of Histone Tails from Core Chromatin." Biochemistry **23**: 5622-5628.
133. Wang, X. and J. J. Hayes (2007). "Site-Specific Binding Affinities within the H2b Tail Domain Indicate Specific Effects of Lysine Acetylation." Journal of Biological Chemistry **282**(45): 32867-32876.
134. Wang, X. and J. J. Hayes (2008). "Acetylation Mimics within Individual Core Histone Tail Domains Indicate Distinct Roles in Regulating Stability of Higher-Order Chromatin Structure." Molecular and Cellular Biology **28**(1): 227-236.
135. Wang, X., C. He, S. C. Moore and J. Ausio (2001). "Effects of Histone Acetylation on the Solubility and Folding of the Chromatin Fiber." Journal of Biological Chemistry **276**(16): 12764-11268.
136. Waterborg, J. H. (2002). "Dynamics of Histone Acetylation in Vivo. A Function for Acetylation Turnover?" Biochemistry and Cell Biology **80**: 363-378.
137. Wedemann, G. and J. Langowski (2002). "Computer Simulation of the 30-Nanometer Chromatin Fiber." Biophysical journal **82**(6): 2847-2859.
138. Wegel, E. and P. Shaw (2005). "Gene Activation and Deactivation Related Changes in the Three-Dimensional Structure of Chromatin." Chromosoma **114**(5): 331-337.
139. Weidemann, T., M. Wachsmuth, T. Knoch, G. Muller, W. Waldeck and J. Langowski (2003). "Counting Nucleosomes in Living Cells with a Combination of Fluorescent Correlation Spectroscopy and Confocal Imaging." Journal of molecular biology **334**: 229-240.

140. Widom, J. (1986). "Physicochemical Studies of the Folding of the 100 a Nucleosome Filament into the 300 a Filament. Cation Dependence." Journal of Molecular Biology **190**(3): 411-424.
141. Widom, J. (1989). "Toward a Unified Model of Chromatin Folding." Annual Review of Biophysics and Biophysical Chemistry **18**: 365-395.
142. Wolffe, A. P. (1998). Chromatin: Structure and Function. San Diego, CA, Academic Press.
143. Wolffe, A. P. and J. J. Hayes (1999). "Chromatin Disruption and Modification." Nucleic Acids Research **27**(3): 711-720.
144. Woodcock, C. L. and S. Dimitrov (2001). "Higher-Order Structure of Chromatin and Chromosomes." Current Opinion in Genetics & Development **11**(130): 130-135.
145. Wu, J. and M. Grunstein (2000). "25 Years after the Nucleosome Model: Chromatin Modification." Trends in Biochemical Sciences **25**: 619-623.
146. Yang, Y., A. P. Lyubartsev, N. Korolev and L. Nordenskiöld (2009). "Computer Modeling Reveals That Modifications of the Histone Tail Charges Define Salt-Dependent Interaction of the Nucleosome Core Particles." Biophysical journal **96**(6): 2082-2094.
147. Zhang, Q., D. A. Beard and T. Schlick (2003). "Constructing Irregular Surfaces to Enclose Macromolecular Complexes for Mesoscale Modeling Using the Discrete Surface Charge Optimization (Disco) Algorithm." Journal of computational chemistry **24**: 2063-2074.
148. Zheng, C. and J. J. Hayes (2003a). "Intra- and Inter-Nucleosomal Protein-DNA Interactions of the Core Histone Tail Domains in a Model System." Journal of Biological Chemistry **278**(26): 24217-24124.
149. Zheng, C. and J. J. Hayes (2003b). "Structures and Interactions of the Core Histone Tail Domains." Biopolymers **68**(4): 539-546.
150. Zheng, C., X. Lu, J. C. Hansen and J. J. Hayes (2005). "Salt-Dependent Intra- and Inter-Nucleosomal Interactions of the H3 Tail Domain in a Model Oligonucleosomal Array." Journal of Biological Chemistry **280**: 33552-33557.
151. Zhou, J., J. Y. Fan, D. Rangasamy and D. J. Tremethick (2007). "The Nucleosome Surface Regulates Chromatin Compaction and Couples It with Transcriptional Repression." Nature Structural & Molecular Biology **14**(11): 1070-1076.