

Electrostatic mechanism of nucleosomal array folding revealed by computer simulation

Jian Sun, Qing Zhang, and Tamar Schlick*

Department of Chemistry and Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012

Edited by Gary Felsenfeld, National Institutes of Health, Bethesda, MD, and approved April 16, 2005 (received for review November 30, 2004)

Although numerous experiments indicate that the chromatin fiber displays salt-dependent conformations, the associated molecular mechanism remains unclear. Here, we apply an irregular Discrete Surface Charge Optimization (DiSCO) model of the nucleosome with all histone tails incorporated to describe by Monte Carlo simulations salt-dependent rearrangements of a nucleosomal array with 12 nucleosomes. The ensemble of nucleosomal array conformations display salt-dependent condensation in good agreement with hydrodynamic measurements and suggest that the array adopts highly irregular 3D zig-zag conformations at high (physiological) salt concentrations and transitions into the extended “beads-on-a-string” conformation at low salt. Energy analyses indicate that the repulsion among linker DNA leads to this extended form, whereas internucleosome attraction drives the folding at high salt. The balance between these two contributions determines the salt-dependent condensation. Importantly, the internucleosome and linker DNA–nucleosome attractions require histone tails; we find that the H3 tails, in particular, are crucial for stabilizing the moderately folded fiber at physiological monovalent salt.

chromatin modeling | irregular 3D zig-zag | Discrete Surface Charge Optimization model

Chromatin folding and the dynamic interplay between chromatin structures and various cellular factors directly regulate fundamental gene expression and silencing processes (1–4). Extensive studies have probed the structural and dynamic properties of chromatin and the molecular mechanisms that determine these properties (5–8). Nonetheless, questions remain concerning how a chain of nucleosomes connected by linker DNA folds into a condensed 30-nm chromatin fiber under physiological conditions (9) and what the exact arrangements of nucleosomes in the 30-nm fiber are (6). Two principal architectures of fiber arrangement proposed are the solenoid and zig-zag models. They differ significantly in how the linker DNA is arranged within the condensed chromatin fiber. In the solenoid model, the linker DNA is bent such that the consecutive nucleosomes interact with each other and form a helix (10–14); in the zig-zag model, the linker DNA remains straight and crosses the fiber axis so that the consecutive nucleosomes appear at the opposite side of the fiber axis (5, 15–18).

The structure of chromatin strongly depends on the salt environment, both in terms of the valence of the ions and their concentration in solution (1). At low salt concentrations, chromatin adopts the extended “beads-on-a-string” conformation, whereas at high salt concentrations it folds into compact forms. In the presence of linker histones (H1 or H5) under physiological salt conditions, chromatin forms the folded 30-nm fiber. The histone tails are known to be crucial for the folding of chromatin fibers (19–21; reviewed in ref. 22). The tail domains, accounting for ≈50% of the positive charges of the histone octamer, are located at the N-terminal portions of H2A, H2B, H3, and H4, and the C terminus of H2A. They are also targets for posttranslational modifications by many enzymes (reviewed in refs. 23 and 24) and are essential for the formation of highly ordered structures (25).

It has been established that the folding of chromatin is mostly attributed to the electrostatic free energy among linker DNA segments (26). However, the internal forces that govern the structure and dynamics of chromatin are not well understood. Only recently, with the advent of single molecule techniques, have these basic physical properties been measured (reviewed in ref. 27). Cui and Bustamante (28) estimated a $3.4 k_B T$ attractive energy per nucleosome at physiological ionic strength. Simulations showed that the nucleosome–nucleosome interaction played a major role in determining the mechanical properties of chromatin (29). A similar conclusion regarding the internucleosome attraction also was obtained in a theoretical analysis (30) of a two-angle model of chromatin (17).

Because of the size and complexity of chromatin, computational studies have been based on coarse-grained models (29, 31–34). For example, a dinucleosome was simulated on nucleosome-constrained circular DNA (31); strings of 4–24 nucleosomes were modeled by virtual bonds, and their structures were determined by geometric parameters (32); the nucleosome was represented by a spherical bead and the interaction between nucleosomes was modeled by a simple spherical isotropic step-like potential (29). A more realistic model considered nucleosomes as oblate ellipsoids, with internucleosome interactions described by the anisotropic Gay–Berne potential (34). Another approach (33) uses the Discrete Surface Charge Optimization (DiSCO) representation of the nucleosome (35). Based on the pioneering nucleosome crystal structure (36), the nucleosome is represented by a few hundred effective charges located on a regular disk and a cylinder, thereby capturing essential electrostatic and structural features of the nucleoprotein complex. Recently, the DiSCO model was further improved by an irregular-surface representation to include the fine surface details and reproduce the electrostatics interactions better (37) (see Fig. 1a).

To investigate the mechanism of the salt-dependent folding of the nucleosomal array in the absence of linker histones and dissect structural features of the array at different salt concentrations, we apply here the new irregular DiSCO model (37) to an array system in which all histone tails are incorporated as the model described in the recent crystal structure [Protein Data Bank (PDB) ID code 1KX5; Fig. 1a] by Davey *et al.* (38). Because salt effects are incorporated by means of a physical model, the nonlinear Poisson–Boltzmann equation solved at different salt concentrations to determine a set of nucleosome charges (35, 37), these analyses become possible. The effect of each histone tail on the stability, as well as salt-dependent folding and unfolding processes based on initial models (two solenoids and two zig-zags with different nucleosome orientations), are explored.

This paper was submitted directly (Track II) to the PNAS office.

Freely available online through the PNAS open access option.

Abbreviations: DiSCO, Discrete Surface Charge Optimization; MC, Monte Carlo; PDB, Protein Data Bank.

*To whom correspondence should be addressed. E-mail: schlick@nyu.edu.

© 2005 by The National Academy of Sciences of the USA

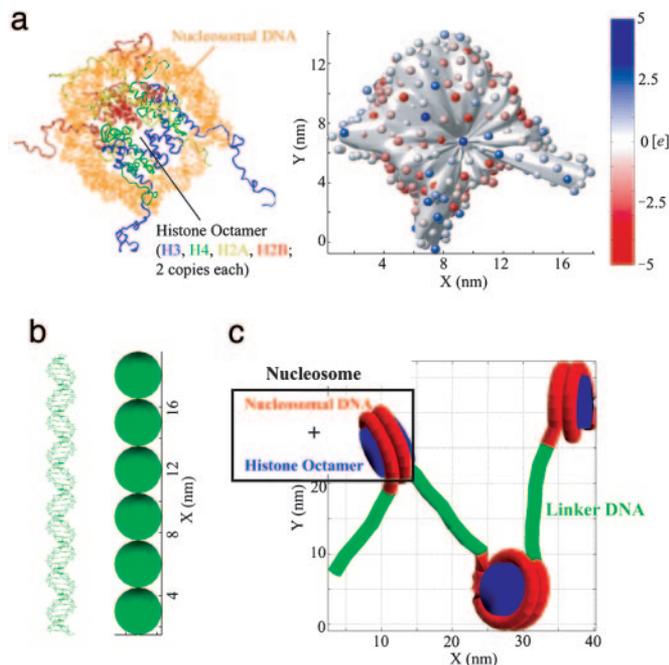


Fig. 1. Models of nucleosome and linker DNA. (a *Left*) Crystal structure of nucleosome with complete modeled histone tails (PDB ID code 1KX5). (a *Right*) DiSCO model of the nucleosome with 300 effective charges located on an irregular surface. The surface was reduced by 5 Å to show the positions of the charges. The color scale indicates the value of each charge in the unit of e . (b *Left*) Linker DNA that connects two nucleosomes. (b *Right*) DNA bead model with six beads of 3-nm diameter. (c) Simplified representation of a short nucleosomal array showing how linker DNA connects nucleosomes.

Our simulations reproduce the observed salt-dependent condensation and suggest that nucleosomal arrays adopt irregular 3D zig-zag conformations at high monovalent salt concentrations. The salt-dependent condensation is governed by electrostatic interactions, including the linker DNA–nucleosome and internucleosome attractions in addition to the repulsion between linker DNA units. Under physiological conditions, the major contribution to the folding and stability of the nucleosomal array arises from the internucleosome interactions. We also find that the histone tails, especially the H3 tails, are responsible for the internucleosome attraction and most of the linker DNA–nucleosome interactions.

Computational Methods

DiSCO Model. The details of constructing the DiSCO model with irregular surface are described in ref. 37. Briefly, 300 discrete charges were uniformly distributed on an irregular surface (Fig. 1*a*), which was built to represent the smoothed molecular surface of the nucleosome (PDB ID code 1KX5; with model of complete histone tails; Fig. 1*a*). The electrostatic potential of the nucleosome was obtained by solving the nonlinear Poisson–Boltzmann equation with the program QNIFTT 1.2 (39–41). The atomic charges and radii were set by the CHARMM27 force field (42). The discrete charges were optimized with the truncated Newton program TNPACK (43–45), integrated in the DISCO software package to reproduce with Debye–Hückel approximation the electric field in the region 0.5 nm away from the model surface (37).

These nucleosomes were linked together with six DNA beads of 3-nm diameter represented by the elastic DNA approach (46–48) (see Fig. 1*b*). The linker DNAs connected to nucleosomes at the entry and exit points of nucleosome DNA that wrap 1.75 turns on the core (33). The default local geometry of linker

DNA at the entry–exit point was such that these linker DNAs had the same orientation and phase as the nucleosomal DNA. The electrostatic and elastic properties of the DNA segments were as described in ref. 33. With this approach, the electrostatic interactions for both the internucleosome and the linker DNA–nucleosome were incorporated explicitly into the simulation of a 48-unit chromatin fiber (33), unlike other approaches.

The DiSCO model has been applied to Brownian dynamics simulations of dinucleosome, trinucleosome, and 48-mer chromatin fiber (33), as well as supercoiled DNA bound to Hin and Fis proteins (49, 50). In these applications, good agreement with experiments has been demonstrated.

Energy Terms. The total energy of the system includes stretching, twisting, bending energy, electrostatic interaction, and excluded volume effect (see *Supporting Appendix*, which is published as supporting information on the PNAS web site, and ref. 33 for details). All energy parameters are listed in Table 3, which is published as supporting information on the PNAS web site. Except for the excluded volume parameters (σ_1 and σ_2), which are reduced by 0.5 nm to accommodate the inclusion of the histone tails, all parameters were the same as in ref. 33.

Monte Carlo (MC) Simulation. We used the standard Metropolis method (51) in our MC sampling to generate the ensemble of nucleosomal array conformations in equilibrium at given salt concentrations. As in previous DNA simulations (46), two types of moves were used. For the local move, one bead (either a DNA bead or nucleosome core particle) was randomly selected and perturbed; the movement orientation was uniformly distributed in the space, and the amplitude was within [0, 0.05 nm]. For the global move, after a bead was randomly selected, the shorter part of the array was rotated along a random axis that passed through the selected bead; the orientation of the axis was uniformly distributed in space, and the rotation angle was uniformly distributed within [0, 5°] (or [0, 10°] at 10 and 30 mM). The acceptance ratio was ≈ 0.5 (0.6–0.7 for 10 and 30 mM). Each MC simulation involved at least 2 million steps with the last 1.5 million steps used to calculate system properties. The run test method was used to determine MC sample sizes and ensure uncorrelated samples, as applied in ref. 52 and implemented in program MMC (<http://inka.mssm.edu/~mezei/mmc>). We found that after 4,098 steps for the energy and 8,196 steps for the sedimentation coefficient, the data became independent; therefore we sampled every 5000th for energy and 10000th for sedimentation coefficient for analysis.

Sedimentation Coefficient Calculation. We calculated the sedimentation coefficient for a given conformation of the nucleosomal array with the method developed by Bloomfield *et al.* (53) following the Kirkwood approach (54), as applied to nucleosomal arrays in refs. 19 and 55. By neglecting the contribution of linker DNA, sedimentation coefficient $s_{20,w}$ can be approximated as

$$\frac{S_N}{S_1} = 1 + \frac{R_1}{N} \sum_i \sum_j \frac{1}{R_{ij}}, \quad [1]$$

where S_N represents $s_{20,w}$ for a rigid structure consisting of N nucleosomes of radius R_1 , R_{ij} is the distance between two nucleosomes, and S_1 is $s_{20,w}$ of a mononucleosome. This approach assumes spherical nucleosomes, a reasonable approximation. We use $R_1 = 5.5$ nm and $S_1 = 11.1$ Svedberg (S; 1 S = 10^{-13} sec) as in previous applications (19). We also used the program HYDRO (56) for the calculation with radius of nucleosome core particle and DNA bead as 5.0 and 1.5 nm (33). These two methods gave similar results ($<0.6S$ differences).

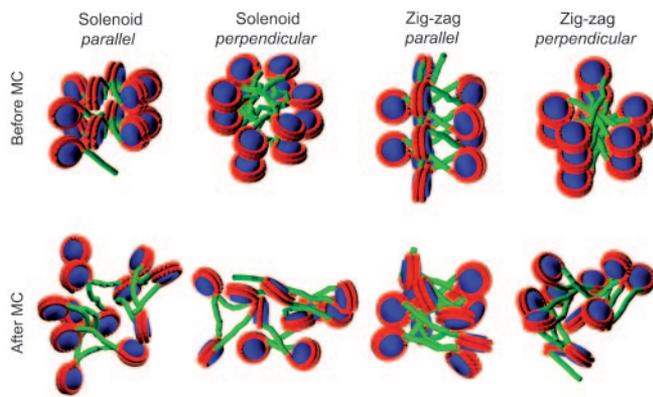


Fig. 2. Structures of nucleosomal arrays before and after MC simulations from different initial models. The initial models are constructed with nucleosome disk planes either parallel or perpendicular to the array helix axis before 2 million steps (6 million steps for the solenoid model with parallel nucleosomes) of MC simulations.

All computations were performed on an Origin 2000 processor at New York University. One million MC steps took ≈ 4 –5 days of central processing unit time.

Results

Pathway of Linker DNA Within Moderately Folded Nucleosomal Array.

Hydrodynamic analyses of the reconstituted homogeneous 208–12 nucleosomal array (12 core histone octamers and 12 tandem repeats of 208-bp DNA) revealed that the array forms a moderately folded structure, not the maximally condensed 30-nm fiber, at monovalent salt concentrations up to 200 mM (19, 55). Nonetheless, these experiments do not provide structural information regarding the pathway of linker DNA in the array. To determine whether the solenoid or the zig-zag model is favorable under these conditions, we constructed four initial configurations by using our basic chromatin model in which nucleosomes represented as effective charges are linked to one another by DNA represented by elastic beads (Fig. 1).

The solenoid model in refs. 10 and 13 has six nucleosomes per helix turn and an 11-nm helix pitch. For a less condensed form (without linker histones), we set six nucleosomes per helix turn but a 14-nm helix pitch. The distance from the center of a nucleosome to the helix axis was set to 13 nm, and the linker DNA connecting two consecutive nucleosomes was bent. This solenoid 12-nucleosomal array thus formed two turns of the helix, as shown in Fig. 2. Although it has been shown that the disk planes of nucleosomes in the condensed 30-nm fiber are roughly parallel to the fiber helix axis (11), we considered both parallel and perpendicular orientations (Fig. 2). The zig-zag model in ref. 5 has 5.9 nucleosomes per 11 nm at 80 mM monovalent salt concentration. For a less condensed form, we set five nucleosomes per helix turn and a 14-nm helix pitch. The distance from the center of a nucleosome to the helix axis was set to 13 nm, but the linker DNA was straight and connected two consecutive nucleosomes on the opposite sides of the array helix axis. In addition, both parallel and perpendicular orientations of nucleosomes were considered (Fig. 2).

From these four initial models, we performed MC simulations for 2 million steps to sample the favorable conformations at the 200 mM monovalent salt concentration. Their final conformations are shown in Fig. 2.

The linker DNA in the two zig-zag models remains straight (Fig. 2), and the bent linker DNA in the solenoid model with perpendicular nucleosomes becomes straight (Fig. 2). Thus, this solenoid structure converges to the zig-zag model. The solenoid model with parallel nucleosomes retains bent linker DNA

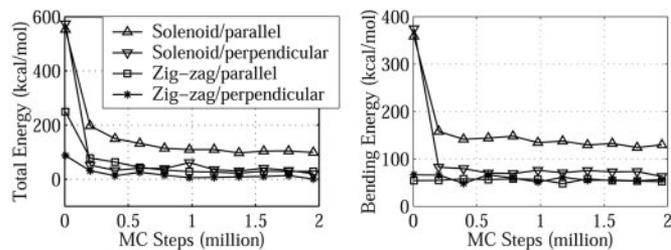


Fig. 3. Total and bending energies as a function of MC steps for different initial models at 200 mM. Four different initial structures, representing the solenoid and zig-zag models with both parallel and perpendicular nucleosome orientations, are shown in Fig. 2. Data points are collected every 5,000 MC steps but are plotted every 200,000 steps.

(within 2 million MC steps) but has not converged after 2 million steps of simulation (Fig. 3); only longer simulations (≈ 5 million steps) are required for its linker DNA to become straight (see Fig. 7, which is published as supporting information on the PNAS web site). These results suggest that at high monovalent salt concentrations, in the absence of linker histones, the linker DNA in the nucleosomal array would not adopt bent forms due to unfavorable bending energy; instead, the array forms irregular zig-zag conformations with the linker DNA mainly extended.

The two zig-zag models both converge to irregular structures with largely extended linker DNA and highly irregular nucleosome orientations (Fig. 2), regardless of the initial orientations. We chose the zig-zag model with perpendicular nucleosomes rather than parallel nucleosomes as the initial conformation for the remaining studies, because the former has a lower total energy (16 vs. 30 kcal/mol; Fig. 3) and reproduces experimental sedimentation coefficients better (e.g., 40 vs. 43 S at 200 mM; the former is closer to the experimental value in Fig. 4).

Salt-Dependent Condensation of the Nucleosomal Array.

We simulated the array's salt-dependent condensation by computing $s_{20,w}$ from MC trajectories at a series of salt concentrations of 10, 30, 50, 100, and 200 mM. Results, shown in Fig. 4 together with data from hydrodynamic measurements (55), reveal conformational changes from extended (27 S at 10 mM) to compact (40 S at 200 mM) forms with close reproduction of the experimental results. Discrepancies at high and low salt may arise from dissociation and ion heterogeneity factors, respectively. At high salt, some histone octamers may dissociate from the nucleosomal DNA (55, 57–59; reviewed in ref. 60), affecting the measurements (55). At low salt, experimentally measured values reflect a buffer with given NaCl concentrations but also other ions, whereas our calculations account only for monovalent salt; there are also charge effects on the measured values well known for polyelectrolytes at low salt.

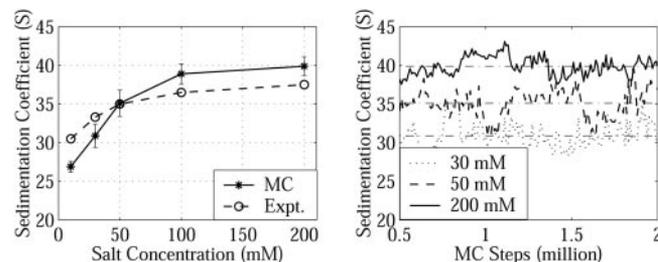


Fig. 4. Sedimentation coefficients $s_{20,w}$ as a function of salt concentration and MC steps. The experimental values (Expt.) are from ref. 55. Calculation of $s_{20,w}$ uses Eq. 1; every 1000th sample is used for the mean and standard deviation calculation.

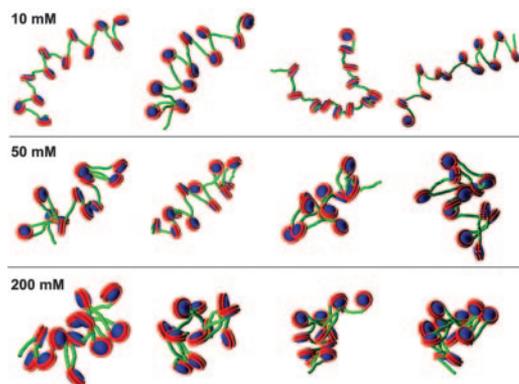


Fig. 5. Snapshots of the nucleosomal array at different salt concentrations. From left to right, the snapshots are taken after 0.5 million, 1 million, 1.5 million, and 2 million steps of the MC simulations.

We also note a wide distribution of $s_{20,w}$ in our simulations (Fig. 4), pointing to a heterogeneous conformational ensemble at a given salt concentration, in equilibrium between extended and compact forms. An increase of the salt concentration shifts the equilibrium to more compact forms, as shown in both Fig. 4 ($s_{20,w}$) and Fig. 5 (MC snapshots), but the ensembles are still heterogeneous because of entropic effects.

Structural Features of the Salt-Dependent Condensation. Detailed structural features of the salt-dependent condensation can be gleaned from averages and standard deviations of the radius of gyration (R_g), the angle (θ) formed by lines connecting three consecutive nucleosomes, and the dihedral angle (ϕ) formed by four consecutive nucleosomes (see Table 4, which is published as supporting information on the PNAS web site). An increase of the salt concentration from 10 to 200 mM decreases R_g from 41 to 19 nm, decreases θ from 64° to 41° , and decreases ϕ from 147° to 58° corresponding to the condensation of an extended beads-on-a-string conformation. These variables' distributions also change dramatically (see Fig. 8, which is published as supporting information on the PNAS web site). Both θ and ϕ are widely distributed at all salt concentrations; at high salt, this behavior indicates irregular zig-zag conformations with some structural heterogeneity.

Energetic Features of the Salt-Dependent Condensation. To interpret the driving force behind the salt-dependent condensation, we compared the energy components at different salt concentrations (see Table 5, which is published as supporting information on the PNAS web site). The total energy drops from 145 kcal/mol at 10 mM to 16 kcal/mol at 200 mM. The electrostatic energy change from 3 kcal/mol (10 mM) to -117 kcal/mol (200 mM) accounts for $>90\%$ of the total energy change.

We dissect the electrostatic energy into linker–linker, linker DNA–nucleosome, and internucleosome components in Table 1. The repulsion between linker DNA is 46 kcal/mol at low salt vs.

Table 1. Energy components of electrostatic interactions

Salt, mM	Linker–linker, kcal/mol	Linker–nucleosome, kcal/mol	Internucleosome, kcal/mol	Total, kcal/mol
10	46 ± 2	-44 ± 5	1 ± 0	3 ± 5
30	25 ± 3	-60 ± 6	-12 ± 6	-48 ± 8
50	19 ± 3	-62 ± 5	-43 ± 8	-86 ± 8
100	14 ± 3	-67 ± 6	-54 ± 8	-108 ± 10
200	9 ± 2	-71 ± 6	-55 ± 6	-117 ± 8

Values are \pm SD.

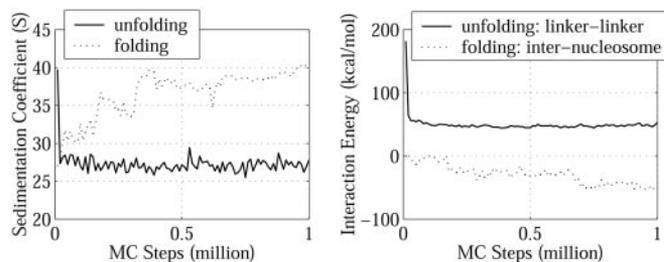


Fig. 6. Unfolding and folding simulations of nucleosomal array. (Left) $s_{20,w}$ during the unfolding simulation at 10 mM and the folding simulation at 200 mM. (Right) Linker–linker interaction energy during the unfolding simulation at 10 mM and internucleosome interaction energy during the folding simulation at 200 mM.

9 kcal/mol at high salt due to the stronger screening effect at high salt concentrations, even though the linker DNAs are closer to each other in more compact forms. Both the linker DNA–nucleosome and internucleosome attractions become much stronger at high salt (Table 1). Significantly, 80% of the internucleosome attraction increase stems from neighboring nucleosome interactions (nucleosomes i with $i - 2$ and $i + 2$ in the zig-zag model because nucleosomes $i - 1$ and $i + 1$ are both on the opposite side of the array helix axis). This attraction directly leads to the collapse of the angle θ and significantly contributes to the condensation of the array.

Folding/Unfolding of the Nucleosomal Array. To better understand the driving force for chromatin folding and unfolding, we analyzed two MC simulations (1 million steps each) that simulate unfolding at 10 mM starting from a compact structure obtained at 200 mM and folding at 200 mM starting from an extended form obtained at 10 mM.

During unfolding, the strong repulsion among the linker DNA in the compact form at low salt concentration drives the array into an extended form to relieve this unfavorable linker–linker interaction. After 10^5 MC steps, the repulsion drops from the initial 182 kcal/mol value to 46 kcal/mol, and the array reaches equilibrium around the extended 27S form (Fig. 6). In equilibrium, the repulsion is mostly compensated by the favorable linker DNA–nucleosome interaction, whereas the internucleosome interaction makes a negligible contribution (1 kcal/mol; Table 1).

During folding, the repulsion among linker DNA is only a few kilocalories per mole and thus does not determine the folding process. Instead, the internucleosome attraction dominates this process: the associated energy drops from 0 to -55 kcal/mol as $s_{20,w}$ increases from 27 to 40 S (Fig. 6). In this process, the linker DNA–nucleosome interaction is always favorable, as in the unfolding process, but here it changes little and exhibits no correlation to the array folding.

These results suggest that folding and unfolding have different driving forces. At low salt concentrations, it is the repulsion among linker DNA that drives the nucleosomal array to adopt an extended form; at high salt concentrations, the attraction between nucleosomes drives and stabilizes the folding of the array.

Role of Histone Tails in Stabilizing the Moderately Folded Array. How can the nucleosomes with a net charge each of approximately $-144 e$ induce strong attractive forces? The positive potential regions of the nucleosome are mainly located in the histone tail domains that protrude from the nucleosome core. Because most of the nucleosomes as well as the linker DNAs are negatively charged, these tails may drive the favorable linker DNA–nucleosome and internucleosome interactions. To estimate the contribution of histone tails to these interactions, we have

Table 2. Contribution of histone tail domains to the electrostatic interactions at 200 mM salt concentration

Tails neutralized	Linker–nucleosome, kcal/mol	Internucleosome, kcal/mol	Total electrostatics, kcal/mol
All	-9 ± 4	7 ± 4	7 ± 6
H2A	-36 ± 5	-14 ± 4	-41 ± 6
H2B	-38 ± 5	-13 ± 4	-43 ± 6
H3	-10 ± 4	1 ± 3	-1 ± 6
H4	-37 ± 5	-7 ± 5	-36 ± 6

The linker–linker interaction is 9 kcal/mol, unchanged from Table 1. Values are \pm SD.

constructed different DiSCO models in which we set the charges in some of the histone tails to neutral and recalculated the electrostatic interactions in the moderately folded array at 200 mM obtained above.

When all histone tails are set to neutral, the linker DNA–nucleosome interaction changes from -71 kcal/mol (Table 1) to -9 kcal/mol (Table 2), the internucleosome interaction changes from favorable -55 kcal/mol (Table 1) to unfavorable 7 kcal/mol (Table 2), and the total electrostatic energy changes from -117 kcal/mol (Table 1) to 7 kcal/mol (Table 2). Therefore, the histone tails dominate the favorable electrostatic interactions responsible for the stability of the folded forms of the array at high salt concentrations, modulating the linker DNA–nucleosome and internucleosome interactions.

To further dissect the contribution of different tails, we set the tails of each type of histone to neutral while keeping the rest of the tails intact. With H3 tails neutral, the linker DNA–nucleosome interaction is reduced to -10 kcal/mol (close to -9 kcal/mol when all tails are neutralized), the internucleosome interaction becomes unfavorably 1 kcal/mol, and the total electrostatic energy is -1 kcal/mol only (Table 2). Although neutralization of each of the H4, H2A, and H2B tails also reduces the favorable interactions, these interactions remain attractive, and the total electrostatic interactions remain favorable (Table 2). These calculations suggest that the most important contribution to the favorable electrostatic interactions of compact forms at high salt comes from the H3 tail domains, which determine the overall favorable linker DNA–nucleosome and internucleosome interactions.

The above results consider neutralized histone tails without resimulating the array at 200 mM. To examine the effect of histone tails on stability, we simulated trypsinized nucleosomal arrays in which given histone tails were excised. As expected, the array with tailless histones unfolds at 200 mM, and its average $s_{20,w}$ is 31 S. Importantly, we find that the array missing only H3 tails cannot achieve the compact form 40 S either at 200 mM; instead, it adopts extended forms with an average $s_{20,w}$ of 32 S. However, removal of any other histone tail has no effect on the stability of the array at 200 mM, and all of the arrays remain in the compact 40S form during the simulations. The unique features of the H3 tail domains in our DiSCO model (longest extensions from the core and proximity to both the entry and the exit of linker DNA) may explain why H3 tails are so important for the folding and stability of chromatin.

Therefore, the internucleosome attraction comes from the interactions of one nucleosome's positively charged histone tails with another's negatively charged nucleosome cores (particularly the nucleosomal DNA). The attraction increases at high salt because the repulsion between linker DNA is screened, allowing closer approach between the positive tails and negative DNA.

Discussion

The strong salt dependence of chromatin structures indicates that electrostatic interactions play a major role in chromatin folding. Our computational study of the 12-unit nucleosomal array has quantitatively analyzed this behavior and compared it with available experimental data to provide further insights into its structural and energetic features. The change of $s_{20,w}$ with the salt from our simulations is in good agreement with the hydrodynamic measurements (55).

Our results also suggest that a highly irregular 3D zig-zag conformation with linker DNA mainly extended is preferred by the nucleosomal array at high monovalent salt concentrations. At low salt, the array becomes extended. The structural changes during the reversible folding process are mirrored by reduction of both the angle θ and the dihedral angle ϕ . We propose that, in addition to the collapse of θ (5, 7, 9), the substantial decrease of ϕ also contributes significantly to the condensation of the array. Of course, all simulations are subject to the well known sampling limitations, although MC trajectories sample a wide range. Because all our simulations lead to irregular zig-zag conformations at high salt, they are likely representative.

Through disulfide cross-links between nucleosome particles, Dorigo *et al.* (61) demonstrated the compact arrays formed in the presence of divalent ions adopted a two-start helix with straight linker DNA, not the solenoid model. Our predicted zig-zag model is consistent with their topologies in the zig-zag class and represents an intermediate folded state formed with monovalent ions on the pathway from an extended form to the final most condensed 30-nm form.

In a theoretical analysis of the electrostatic mechanism of chromatin folding, Clark and Kimura (26) concluded that the mechanism of chromatin folding was essentially electrostatic in nature and proposed the major role of repulsion between linker DNA in determining the stability of chromatin fiber. Our results suggest that the repulsion among linker DNA dominates at low salt, but favorable internucleosome interactions dominate at high salt. Thus, the present studies demonstrate the importance of the internucleosome interaction for the folding and stability of chromatin fiber at physiological salt concentrations, as suggested by van Holde, Zlatanova, and coworkers (7, 9) and indicated by the single molecular experiments on the chromatin fiber (28).

We also traced the origin of the favorable internucleosome interaction to the histone tail domains, especially of H3 (Table 2). At high salt, the repulsion between linker DNA is screened, allowing stronger attraction between one nucleosome's positively charged histone tails and another's negatively charged nucleosome cores. In an experimental study of a nucleosome solution without linker DNA (62), it was found that at low salt concentrations (<50 mM) the histone tails collapsed to interact with its own nucleosomal DNA and the internucleosome interactions were unfavorable, whereas at high salt concentrations (>50 mM) the histone tails extended from the nucleosome to form favorable internucleosome interactions. The origin of the observed internucleosome attraction was further attributed to the histone tail domains by comparing the intact nucleosome solution and the tailless nucleosome solution (63).

The present approach, which includes all histone tails modeled, allowed us to investigate these forces because the effective charges in the DiSCO model are derived directly from the nucleosome's electrostatic potential, and thus salt effects are naturally incorporated (35, 37). Prior simulations used simpler approximations. For example, in the simulations of pulling chromatin fibers, the interactions involving nucleosomes included nucleosome–DNA excluded volume effects and short-range internucleosome attraction (29); this short-range attraction was modeled by an isotropic hard-core potential in which the

depth of the attractive well was adjusted to increase with salt concentration. Thus, in the absence of imposed stretching forces, the fiber condensation at high salt concentration was induced by adjustment of a parameter, the decrease of the effective nucleosome diameter. In the study of chromatin folding by Langowski's group, the internucleosome interaction was described by a Gay-Berne potential, a generalization of Lennard-Jones potential for objects with ellipsoidal symmetry (34). However, their focus was on modeling chromatin at near-physiological ionic conditions; thus the effects of the ionic strength on the nucleosomal interactions and on the geometrical parameters were not modeled; it was suggested that reparameterizing the Gay-Berne potential according to the experimental salt-dependence of the linear mass density of the chromatin fiber is required to explore salt-dependent behavior.

Our present chromatin model treats nucleosomes, including the histone tails, as rigid bodies; these bodies can translate/rotate as a whole; only the linker DNA is fully flexible, within our resolution (Fig. 1). Clearly, the histone tails are flexible and rearrange to different locations depending on the context in which they are isolated or assembled (64, 65; reviewed in refs. 22 and 60). Still, we expect that the flexible histone tails would adjust their conformations in chromatin to form favorable linker

DNA-nucleosome interactions as well as internucleosome and intranucleosome interactions. In fact, we have tested different orientations of histone tails in our DiSCO model and found that the exact positions of the tails do not matter. The array stays in moderate folded forms at high salt as long as the tails protrude from nucleosome cores; the tails prefer protrusion at high salt as observed in experiments (62). Another approximation in the present model is our use of Debye-Hückel approximation, which accounts for neither the nonlinear effect from the polyelectrolyte feature of both the linker DNA and the nucleosomal DNA nor the contribution of desolvation terms. Although this macroscopic model of the nucleosome is reasonable, as demonstrated by a good agreement with experiments, and allows systematic exploration of general features, more realistic and refined models of the nucleosome are required to address questions such as the effects of posttranslational modifications (most on the histone tails) and histone variants on the nucleosomal array and chromatin fiber.

We thank Dr. Gary Felsenfeld for stimulating discussions related to this work, Drs. Daniel A. Beard and Jing Huang for help with the MC simulation program, Dr. Mihaly Mezei for help on the uncorrelated analysis, and two reviewers for critical comments on our manuscript. This work is supported by National Institutes of Health Grant R01 GM55164.

- van Holde, K. E. (1989) *Chromatin* (Springer, New York).
- Wolffe, A. (1995) *Chromatin: Structure and Function* (Academic, London).
- Wolffe, A. P. & Guschin, D. (2000) *J. Struct. Biol.* **129**, 102–122.
- Felsenfeld, G. & Groudine, M. (2003) *Nature* **421**, 448–453.
- Bednar, J., Horowitz, R. A., Grigoryev, S. A., Carruthers, L. M., Hansen, J. C., Koster, A. J. & Woodcock, C. L. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 14173–14178.
- Widom, J. (1998) *Annu. Rev. Biophys. Biomol. Struct.* **27**, 285–327.
- Zlatanova, J., Leuba, S. H. & van Holde, K. (1998) *Biophys. J.* **74**, 2554–2566.
- Luger, K. (2003) *Curr. Opin. Genet. Dev.* **13**, 127–135.
- van Holde, K. & Zlatanova, J. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 10548–10555.
- Finch, J. T. & Klug, A. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 1897–1901.
- McGhee, J. D., Nickol, J. M., Felsenfeld, G. & Rau, D. C. (1983) *Cell* **33**, 831–841.
- Butler, P. J. (1984) *EMBO J.* **3**, 2599–2604.
- Widom, J. & Klug, A. (1985) *Cell* **43**, 207–213.
- Bartolome, S., Bermudez, A. & Daban, J. R. (1994) *J. Cell Sci.* **107**, 2983–2992.
- Worcel, A., Strogatz, S. & Riley, D. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 1461–1465.
- Bordas, J., Perezgrau, L., Koch, M. H. J., Vega, M. C. & Nave, C. (1986) *Eur. Biophys. J. Biophys. Lett.* **13**, 175–185.
- Woodcock, C. L., Grigoryev, S. A., Horowitz, R. A. & Whitaker, N. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 9021–9025.
- Horowitz, R. A., Koster, A. J., Walz, J. & Woodcock, C. L. (1997) *J. Struct. Biol.* **120**, 353–362.
- Garcia-Ramirez, M., Dong, F. & Ausio, J. (1992) *J. Biol. Chem.* **267**, 19587–19595.
- Moore, S. C. & Ausio, J. (1997) *Biochem. Biophys. Res. Commun.* **230**, 136–139.
- Hansen, J. C., Tse, C. & Wolffe, A. P. (1998) *Biochemistry* **37**, 17637–17641.
- Zheng, C. & Hayes, J. J. (2003) *Biopolymers* **68**, 539–546.
- Turner, B. M. (2002) *Cell* **111**, 285–291.
- Khorasanizadeh, S. (2004) *Cell* **116**, 259–272.
- Allan, J., Harborne, N., Rau, D. C. & Gould, H. (1982) *J. Cell Biol.* **93**, 285–297.
- Clark, D. J. & Kimura, T. (1990) *J. Mol. Biol.* **211**, 883–896.
- Zlatanova, J. & Leuba, S. H. (2003) *J. Mol. Biol.* **331**, 1–19.
- Cui, Y. & Bustamante, C. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 127–132.
- Katritch, V., Bustamante, C. & Olson, W. K. (2000) *J. Mol. Biol.* **295**, 29–40.
- Schiessel, H., Gelbart, W. M. & Bruinsma, R. (2001) *Biophys. J.* **80**, 1940–1956.
- Martino, J. A. & Olson, W. K. (1998) *Biophys. J.* **74**, 2491–2500.
- Martino, J. A., Katritch, V. & Olson, W. K. (1999) *Struct. Fold. Design* **7**, 1009–1022.
- Beard, D. A. & Schlick, T. (2001) *Structure (London)* **9**, 105–114.
- Wedemann, G. & Langowski, J. (2002) *Biophys. J.* **82**, 2847–2859.
- Beard, D. A. & Schlick, T. (2001) *Biopolymers* **58**, 106–115.
- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997) *Nature* **389**, 251–260.
- Zhang, Q., Beard, D. A. & Schlick, T. (2003) *J. Comput. Chem.* **24**, 2063–2074.
- Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. (2002) *J. Mol. Biol.* **319**, 1097–1113.
- Gilson, M. K., Sharp, K. A. & Honig, B. H. (1988) *J. Comput. Chem.* **9**, 327–335.
- Sharp, K. A. & Honig, B. (1990) *Annu. Rev. Biophys. Biophys. Chem.* **19**, 301–332.
- Sharp, K. A. & Honig, B. (1990) *J. Phys. Chem.* **94**, 7684–7692.
- Foloppe, N. & MacKerell, A. D. (2000) *J. Comput. Chem.* **21**, 86–104.
- Schlick, T. & Fogelson, A. (1992) *ACM Trans. Math. Software* **18**, 46–70.
- Schlick, T. & Fogelson, A. (1992) *ACM Trans. Math. Software* **18**, 71–111.
- Xie, D. & Schlick, T. (1999) *ACM Trans. Math. Software* **25**, 108–122.
- Jian, H. M., Vologodskii, A. V. & Schlick, T. (1997) *J. Comput. Phys.* **136**, 168–179.
- Jian, H., Schlick, T. & Vologodskii, A. (1998) *J. Mol. Biol.* **284**, 287–296.
- Huang, J., Schlick, T. & Vologodskii, A. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 968–973.
- Huang, J. & Schlick, T. (2002) *J. Chem. Phys.* **117**, 8573–8586.
- Huang, J., Zhang, Q. & Schlick, T. (2003) *Biophys. J.* **85**, 804–817.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. (1953) *J. Chem. Phys.* **21**, 1087–1092.
- White, S. H. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 407–439.
- Bloomfield, V., Dalton, W. O. & van Holde, K. E. (1967) *Biopolymers* **5**, 135–148.
- Kirkwood, J. G. (1954) *J. Polymer Sci.* **12**, 1–14.
- Hansen, J. C., Ausio, J., Stanik, V. H. & van Holde, K. E. (1989) *Biochemistry* **28**, 9129–9136.
- Garcia, d. I. T., Navarro, S., Lopez Martinez, M. C., Diaz, F. G. & Lopez Cascales, J. J. (1994) *Biophys. J.* **67**, 530–531.
- Yager, T. D. & van Holde, K. E. (1984) *J. Biol. Chem.* **259**, 4212–4222.
- Ausio, J., Seger, D. & Eisenberg, H. (1984) *J. Mol. Biol.* **176**, 77–104.
- Hansen, J. C., van Holde, K. E. & Lohr, D. (1991) *J. Biol. Chem.* **266**, 4276–4282.
- Hansen, J. C. (2002) *Annu. Rev. Biophys. Biomol. Struct.* **31**, 361–392.
- Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R. R. & Richmond, T. J. (2004) *Science* **306**, 1571–1573.
- Mangenot, S., Leforestier, A., Vachette, P., Durand, D. & Livolant, F. (2002) *Biophys. J.* **82**, 345–356.
- Bertin, A., Leforestier, A., Durand, D. & Livolant, F. (2004) *Biochemistry* **43**, 4773–4780.
- Usachenko, S. I., Bavykin, S. G., Gavin, I. M. & Bradbury, E. M. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 6845–6849.
- Angelov, D., Vitolo, J. M., Mutskov, V., Dimitrov, S. & Hayes, J. J. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 6599–6604.