Multiscale modelling of nucleosome core particle aggregation

Alexander P Lyubartsev1, Nikolay Korolev2, Yanping Fan2 and Lars Nordenskiöld2

1 Division of Physical Chemistry, Department of Materials and Environmental Chemistry, Arrhenius Laboratory, Stockholm University, 106 91 Stockholm, Sweden
2 School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, 637551 Singapore
E-mail: alexander.lyubartsev@mmk.su.se and LarsNor@ntu.edu.sg

Received 16 July 2014, revised 19 September 2014
Accepted for publication 26 September 2014
Published 7 January 2015

Abstract
The nucleosome core particle (NCP) is the basic building block of chromatin. Under the influence of multivalent cations, isolated mononucleosomes exhibit a rich phase behaviour forming various columnar phases with characteristic NCP–NCP stacking. NCP stacking is also a regular element of chromatin structure in vivo. Understanding the mechanism of nucleosome stacking and the conditions leading to self-assembly of NCPs is still incomplete. Due to the complexity of the system and the need to describe electrostatics properly by including the explicit mobile ions, novel modelling approaches based on coarse-grained (CG) methods at the multiscale level becomes a necessity. In this work we present a multiscale CG computer simulation approach to modelling interactions and self-assembly of solutions of NCPs induced by the presence of multivalent cations. Starting from continuum simulations including explicit three-valent cobalt(III)hexammine (CoHex3+) counterions and 20 NCPs, based on a previously developed advanced CG NCP model with one bead per amino acid and five beads per two DNA base pair unit (Fan et al 2013 PLoS One 8 e54228), we use the inverse Monte Carlo method to calculate effective interaction potentials for a ‘super-CG’ NCP model consisting of seven beads for each NCP. These interaction potentials are used in large-scale simulations of up to 5000 NCPs, modelling self-assembly induced by CoHex3+. The systems of ‘super-CG’ NCPs form a single large cluster of stacked NCPs without long-range order in agreement with experimental data for NCPs precipitated by the three-valent polyamine, spermidine3+.

Keywords: nucleosome core particles, chromatin compaction, multivalent ions, coarse-grained models, multiscale simulations, inverse Monte Carlo

(Some figures may appear in colour only in the online journal)

1. Introduction
In the nucleus of the eukaryotic cell, the genetic material of DNA is packed into chromatin as linear domains of a uniform DNA–histone complex, the nucleosomes. The regular central part of the nucleosome is the nucleosome core particle (NCP) (figure 1(a)) and it consists of 147 base pairs (bp) of DNA that is wrapped as a 1.75-turn superhelix around the histone octamer (HO) protein complex, comprising two copies of each of the four histone proteins H2A, H2B, H3 and H4. The core histones have flexible and positively charged N-terminal domains, so-called ‘histone tails’. These eight tails each have a net charge of +9e−14e (due to varying number of positively charged lysine and arginine residues). Linker DNA with a length in the range 10–70 bp connects the NCPs into a nucleosome array that can compact to a 30 nm chromatin fibre. The histone tails protrude out through the NCP core and interact with the DNA of its own and other NCPs as well as...
Figure 1. Multiscale coarse-graining of the nucleosome core particle. The top row shows the view from the top; the bottom row displays a side view. (a) All-atom structure of the NCP with the histone tails shown in stretched conformation. Charged phosphate groups of the DNA shown in colour; atoms of the positively (Lys, Arg) and negatively (Glu, Asp) charged amino acids are shown respectively in blue and red. (b) CG model of the NCP [34] with a resolution of one particle per amino acid and five-particle per two bp of the DNA. Colouring is the same as in (a). (c) SCG model of the NCP consisting of one central (green) and six outer (red) particles. The SCG model is superimposed on CG NCP shown in (b). See text for details.

with linker DNA of the chromatin fibre. These tails are needed for secondary as well as tertiary condensation of chromatin by mediating intra- and inter-array interactions and for regulation of transcription and replication [1, 2]. Recent data indicate that folding and aggregation of chromatin is facilitated and regulated by tail bridging between adjacent NCPs [3–6]. This condensation is reduced by acetylations of the tail domains, which neutralizes the lysine charges. Details of the physics of this regulation are still poorly understood [6, 7]. In higher organisms, another histone protein, the linker histone H1, is also present. The linker histones contribute to the chromatin folding but it has been shown that the fibres can compact even in their absence.

Nucleosome arrays and isolated NCPs (mononucleosomes) exhibit a salt-induced aggregation (self-association) behaviour [5, 8, 9]. The seminal work of Livolant and co-workers investigated the phase behaviour of NCPs in solution [3, 4, 8, 10–15]. It may be noted that the cation-induced NCP self-association is similar to the behaviour exhibited by other polyelectrolyte systems such as DNA [16, 17]. The aggregation of NCPs is dependent on the charge and nature of the cations present and histone tails mediate NCP–NCP contacts in a salt-dependent way.

A necessary condition for condensed mononucleosomes and for compact folded chromatin is the nucleosome stacking, i.e. a close NCP–NCP contact between the flat surfaces of the HO core on both sides of the cylindrical wedge-shaped NCP. This stacking has been observed in NCP crystals [18–22], in NCP liquid crystalline phases [8, 11, 13, 15, 23], in the crystal of the tetranucleosome [24], in folded nucleosome arrays [25–27], and in cryo-microscopy images of frozen isolated native chromatin [28, 29]. Theoretical investigations and development of computational approaches aimed at studying the multivalent induced NCP phase separation and the mechanism of nucleosome–nucleosome stacking will therefore contribute to our understanding of chromatin condensation.

Recently, computer simulation approaches to model chromatin structure and dynamics have become important tools to address questions related to the physics of chromatin [30, 31]. Computer modelling has the advantage that it is possible to investigate different physical mechanisms and assess their importance to chromatin interactions and condensation behaviour. An atomistic approach to modelling a chromatin fibre or a solution of many NCPs that includes water molecules and mobile ions would result in a system with such a large number of particles that convergent simulations cannot be accomplished, even using the most powerful computers at present. Consequently, coarse-grained (CG) approaches are needed. Ideally, it would be desirable to perform a systematic coarse-graining based on bridging a fine-grained description of the system (i.e. based on all-atom models) to the
CG model where the CG interaction potentials were obtained from a rigorous mapping of the fine-grained description to the CG model, using e.g. the inverse Monte Carlo (IMC) method [32, 33]. However, the CG approaches presently used in chromatin modelling are not yet at such a rigorous level and to the best of our knowledge, have not used a systematic CG approach based on a fully atomistic description. Continuum models are usually employed and in most cases using a mean field Poisson–Boltzmann (PB) or Debye–Hückel (DH) description of the electrostatic interactions, which neglects ion–ion correlations. Development of systematic CG approaches to the study of chromatin condensation and NCP interactions is therefore highly desirable.

We have recently developed a CG model of DNA that was used to construct a CG NCP model. This new model, which is illustrated in figure 1(b), was benchmarked and used in continuum model simulations with explicit presence of mobile ions to study cation-induced NCP self-association [34]. The model is based on the crystal structure of the NCP with appropriate modelling of the geometry and charge distribution in the NCP (figures 1(a) and (b)). In a subsequent work, the parameters of the DNA model that determines the flexibility of the free DNA were further optimized, starting from atomistic simulations and using the IMC method [35]. In agreement with experiments [3, 36], in simulations of a system with a rather small number of NCPs (ten particles) and with only neutralizing monovalent counterions present (no added salt) [34, 37], NCP–NCP interactions are repulsive. However, the addition of the multivalent Co(NH3)63+ resulted in aggregation into a single cluster of NCPs. The results showed that this model reproduced the expected salt-dependent self-assembly of NCPs induced by multivalent cations. Remarkably, for such a small test system as one consisting of ten particles, nucleosome–nucleosome stacking is observed. We analysed the structural details and found very good agreement with the resulting NCP–NCP stacking parameters observed experimentally [34]. The model introduced in [34] contained about 1350 CG sites, which, together with the necessity to include explicit mobile ions, restrict the number of simulated NCPs to a few tens. This is not enough to investigate interactions and coordination of a large amount NCPs which might be relevant for large-scale features of chromatin folding, or to model results of x-ray scattering experiments on aggregation of NCPs in the presence of multivalent ions.

We therefore, in the present paper use the IMC method to compute effective NCP–NCP potentials for a ‘super-coarse-grained’ (SCG) model, describing the NCP with one central (HO) and six surrounding (DNA) beads (figure 1(c)). The potentials are obtained from simulation in a system of 20 NCPs in the presence of three-valent counterions (representing cobalt(III)hexamine, CoHex3+) using the CG model illustrated in figure 1(b), by calculating radial distribution functions (RDFs) mapped to the effective sites corresponding to the beads of the SCG model. Using these effective potentials derived from the advanced and detailed NCP model, we performed multiscale simulations of 500 and 5000 NCPs in 100–200 nm cubic boxes to study phase separation and self-assembly at the mesoscopic scale. The purpose is to test the IMC approach for the study of NCP phase separation and NCP–NCP stacking in a large-scale simulation with a highly coarse-grained model of the NCP, but based on underlying simulations at a much more fine-grained level (albeit not atomistic). In the presence of the three-valent CoHex3+, these multiscale simulations result in phase separation of the NCP solution into a separate large cluster of aggregated NCPs, which exhibit internal structure very similar to experimental small angle x-ray scattering (SAXS) data obtained for NCPs in the presence of the three-valent spermidine3+ ions [11]. The approach used in this work can be further developed by making refinements in the fine-grained NCP model, which in itself can be subject to systematic coarse-graining using the IMC method to derive more rigorous effective CG potentials based on atomistic NCP simulations.

2. Computational methods

2.1. High-resolution NCP simulations

The underlying high-resolution simulation was carried out using the NCP model described in the previous paper [34]. Briefly, this model includes a CG model of DNA (with one CG site per each phosphate group and one central site per each two bp), a CG model of the HO core and histone tails (with one site per each amino acid) and hydrated ions in continuum solvent. The DNA consisting of 148 bp is wrapped around the histone core making about 1.75 turns. The internal DNA structure was maintained by harmonic bonds and angular interactions between neighbouring CG sites. The histone core monomers are bound using an elastic network scheme [38] according to the high-resolution crystal structure (1KX5.pdb entry [21]), while histone tails are completely flexible. The total number of CG sites on each NCP was 1350. The short-range interactions (excluded volume) were described by a short-range $r^{-12}$ potential with specific effective radii for each CG site type. All the charges of DNA, histone core and histone tails as well as of ions were treated explicitly by a Coulombic potential scaled by the dielectric permittivity of water $\varepsilon = 78$. The net charge of each NCP was $-150e$, with contributions $-296e$, $+52e$ and $+94e$ from DNA, histone core and histone tails, respectively. Further details on this ‘high-resolution’ CG NCP model, together with the values of parameters are given in [34].

In the present work, molecular dynamics simulations using a Langevin thermostat have been carried out using the ESPResSO package [39] for a system consisting of 20 high-resolution NCPs and 1000 neutralizing three-valent counterions representing CoHex3+. No added salt ions were present. The approximation of ‘no added salt’ is a good mimic of the real experimental systems where the background concentration of monovalent cations is in the millimolar range and cations like Na+ or K+ are in practice completely expelled from the vicinity of DNA or NCP as shown in experiments on DNA and chromatin condensation in mixed salt solutions [16, 40, 41]. As in [34], the effective CoHex3+ radius was set to 3.5 Å (which was validated in our previous
work [42]). The NCPs with counterions were put in a cubic periodic box of size 500 Å. The electrostatic interactions were treated by the P3M method [43]. Masses of beads were given in reduced units (with unit energy defined as $k_B T$) according to the same scheme as in the previous work [34]. The time step, in the same reduced units, was set to 0.04 (which correspond to about 30 fs in real units). Simulations, which started from an initial random configuration, were run for $2 \times 10^7$ time units (equivalent to 15 $\mu$s of real time) of which the first $10^7$ time units were disregarded in analysis. The obtained trajectories were used for computations of RDFs which were further used to determine interaction potentials for the SCG model following the IMC method.

### 2.2. The SCG NCP model

In the ‘super-coarse-grained’ (SCG) model, an NCP is represented by seven sites, with one central core representing the octamer core and six sites on a ring around the central one representing both turns of DNA wrapped around it (see figure 1(c)). The molecular dynamics trajectory obtained in the high-resolution NCP simulations was mapped onto a CG trajectory, in which coordinates of the central bead in each NCP (labelled as ‘O’) was determined by the centre of mass of the histone core and the coordinates of the ring beads (labelled as ‘P’) were determined from the DNA beads that occurred within each 60$^\circ$ sectors of the DNA ring. In the present model we did not distinguish beads occurring near the positions of DNA entry/exit points of the NCP, though taking this into account can be introduced in future models.

The SCG trajectory was used to compute RDFs between ‘O’ and ‘P’ sites of the SCG model, as well as for P–O and two types of P–P intramolecular bond lengths distributions. The P–P bond distributions were computed separately for P–P bonds along the ring and for pairs of opposite P beads in the ring, the later were included in order to keep the NCP structure nearly planar. The structural information acquired from the detailed simulations was used in the IMC procedure in which effective potentials for interactions between O and P sites of the SCG model were determined. The IMC calculations were carried out using the MagiC package [44] for the same number of NCPs (20) and in a box of the same size as the high-resolution NCP simulations. At the start, the trial intermolecular potential was set to zero, and 10 iterations within the iterative Boltzmann inversion have been carried out. After that, the IMC scheme of updating potentials, which takes into account cross-correlations between different site types, was started. After 20 additional iterations the IMC calculation converged to the potentials which, within the SCG model, reproduce the same RDFs and distributions of intramolecular bond lengths as those computed in the high-resolution NCP simulation.

The SCG model does not include histone tails or mobile ions (and hence no dielectric continuum); instead, the effect of those is incorporated into the effective SCG potentials. The effective potentials obtained in this way depend on the ionic conditions and on the characteristics of the histone tails and can be used only for conditions used in the high-resolution simulations (wild type NCPs in the presence of three-valent counterions). Nevertheless, since the obtained SCG model contains much less degrees of freedom but reproduces the same structural properties as the detailed model, it can be used in larger scale simulations to study aggregation and phase behaviour of a large number of NCPs under those conditions.

### 2.3. Large-scale simulations of the SCG NCP model

The SCG NCP model, with the tabulated interaction potentials obtained by the IMC method from high-resolution simulations, has been used in Langevin MD simulations of a large number of NCPs. As in the case of the high-resolution simulations, the ESPResSO package [39] was used. The masses of the beads were taken in reduced units to be 50 and 10 for ‘O’ and ‘P’ sites respectively, the time step was set to 0.05 (being equivalent to about 150 fs) and the frictional parameter $\gamma$ was set to 0.1. Two types of simulations have been carried out: (a) self-assembly of 500 and 5000 dissolved NCPs from an initial random configuration and (b) simulation of a condensed state in a periodic system for a constant-pressure (NPT) ensemble with an osmotic pressure of 0.1 bar.

### 3. Results and discussion

#### 3.1. Simulations of aggregation in a system of 20 NCPs and CoHex$^{3+}$ ions within the advanced CG model

Similarly to our previous work [34], the high-resolution NCPs in presence of three-valent counterions aggregated into a single cluster after about $2 \times 10^3$ of Langevin MD time units. A typical snapshot is shown in figure 2(a). One can clearly see a number of NCPs in stacking ‘columnar’ conformations. Some other NCP pairs were in ‘perpendicular’ conformation with DNA strands of one NCP contacting the HO core of another NCP. The mutual coordination of the NCP was changed dynamically during the simulation with both (stacking and perpendicular) types of contacts. The resulting RDF between NCP HO core centres of mass is shown in figure 2(b). It exhibits three clear maxima. The first one, at about 60 Å, corresponds to stacking NCP coordination while the second, at about 80 Å, corresponds to ‘perpendicular’ coordination as depicted by the inserts near the corresponding RDF maxima. The third maximum is a combination of ‘side-by-side’ coordination of NCPs and second neighbours in ‘stacking’ coordination. This simulation with 20 high-resolution CG NCPs confirms all essential features of NCP aggregation behaviour described in details in our previous work [34] for 10 NCPs in the presence of CoHex$^{3+}$ counterions.

#### 3.2. Extraction of interaction potentials from the advanced CG NCP simulations

The RDFs between all pairs of sites of the SCG NCP model extracted from the simulation with the high-resolution model are shown in figures 3(a)–(c). The maximum of the ‘core-P’ RDF in the range 70–80 Å correspond to ‘stacking’ NCP coordination. There is still a noticeable contribution to the core-P RDF at shorter distances starting from about 35 Å which
Figure 2. Molecular dynamics simulations of a system consisting of 20 CG NCPs with three-valent cationic CoHex$^{3+}$ counterions. (a) A snapshot from the simulations demonstrating NCP–NCP aggregation. The DNA backbone is shown in gold; particles of the globular part of the HO are shown in light blue; particles of the phosphate groups, histone tails and CoHex$^{3+}$ are not displayed. (b) NCP–NCP RDF calculated from the simulations of the system shown in (a). The centre of mass of the globular part of the HO was used to define coordinates of the NCPs in the simulated system. Cartoons near the peaks show NCP orientation for the most populated NCP–NCP contacts.

Figure 3. Radial distribution (a)–(c) and effective CG potential in kJ mol$^{-1}$ (d)–(f) functions for inter-nucleosome contacts of the 7-bead SCG model of the NCP. RDFs were calculated from data of simulations in a system of 20 CG NCPs with CoHex$^{3+}$ counterions. Potential functions were determined using the IMC method [44]. In calculations of RDFs and potentials, the CG model of the NCP (shown in figure 1(b)) was coarse-grained with two types of particles, one central ‘core’ and six outer ‘P’ beads as shown in figure 1(c). (a), (d) Core–core, (b), (e) core–P, and (c), (f) P–P radial distribution and potential functions.

come from the perpendicular NCP coordination. The P–P RDF has a broad distribution in the range 50–120 Å which comes from distances between different phosphate beads of the SCG NCP.

The intermolecular potentials between the core and P sites of the SCG NCP model are shown in figures 3(d)–(f). The IMC-derived potentials provide for the SCG NCP model RDFs and distributions of intramolecular bond distances which coincide (within at least 0.02 of the RDF magnitude) with the corresponding distributions of the high-resolution model if the simulations are carried out for the same number of NCPs (20) in a cubic periodic box of the same size (500 Å). The core–core and P–P potentials are attractive while the core–P potential is repulsive. The overall interaction is attractive and results in self-assembly of the SCG NCPs with site–site interactions described by these potentials.

3.3. Simulations within the SCG model

Several large-scale Langevin MD simulations of the SCG NCP model have been carried out. Two simulations were carried out for 500 and 5000 NCP in periodic cubic boxes of fixed sizes 1600 Å and 2000 Å respectively. The starting coordinates of the NCPs were chosen randomly. In the both simulations, the NCPs were gathered in a single cluster, formed after about 5 µs and 20 µs for 500 and 5000 NCP systems respectively. After
assembly to a single cluster, the systems were simulated for 9 µs additionally (3 million time units). The final snapshots are displayed in figure 4. One can see that the aggregated cluster of the 500 NCP system has a nearly spherical shape while shape of the bigger 5000 NCP cluster is elongated. The shape was however not fixed and fluctuated during the production stage of the simulation.

We have also carried out simulation of a condensed NCP system where 5000 NCPs, being in the aggregated state, fill the whole volume of a periodic cubic cell. This simulation was done in order to investigate the structure factor of the condensed phase excluding the influence of the border region effect, which was substantial (see discussion below) even for the large 5000 NCP system. The initial configuration for this system was obtained by taking one of the latest configurations of the 5000 NCP system and simulating it with gradually decreased volume until the aggregated cluster begin to interact through periodic boundary conditions. Then the constant-pressure (Parrinello–Rahman NPT) algorithm was switched on at a pressure 0.1 bar (formally corresponding to an osmotic pressure of the NCP solution), and the system was simulated until the average box size was stabilized (at about 1020 Å) and the NCPs continuously filled the whole available volume. After that the system was simulated additional 4 µs.

The RDFs of the simulated systems are plotted in figure 5. The first two systems are phase-separated (with aggregated NCPs and empty space representing pure solvent), that is why their RDFs have high values at small distances and tend to zero at large distances. In order to compare RDFs of cluster systems of different average density and RDF of a homogeneous system on an equal footing, the RDFs of the 500 and 5000 cluster system were scaled by factors 0.05 and 0.15 respectively, which brings their values in the range of the first 2–3 maxima to about the same level. Comparison of these RDFs, as well as core–core RDF of the high-resolution model with 20 NCP (figure 3(a)) shows some interesting features. First, all systems show RDF maxima at about 60 and 110 Å, relative heights of which are similar in all systems. However, the second maximum of the high-resolution NCP model (figures 2(b) and 3(a)) becomes small in the 500 NCP cluster system, practically disappears in the 5000 NCP system and is completely absent in the continuous NCP system. Note, that the core–core RDF of the SCG system coincides with the core–core RDF of the high-resolution system if simulations are done at the same conditions, that is 20 NCPs in a 500 Å box. The reason for the disappearance of the second RDF maxima at around 80 Å (which correspond to perpendicular orientations of the NCPs) is most likely that this maxima comes from NCPs at the surface of the cluster: in the 20 NCP cluster practically all NCPs are at the surface, while the fraction of surface NCPs is substantially less in a 500 system and even less in the 5000 cluster system and they are completely absent in the continuous system.

The structure factor of the aggregated NCPs was computed for the 5000 cluster and for the continuous systems. The standard way of computing the structure factor from the RDF is using following expression:

$$S(q) = 1 + \frac{\rho}{q} \int_0^\infty r (g(r) - 1) \sin(qr) \, dr$$  \hspace{1cm} (1)$$

where \(\rho\) is the system density (in terms of number of particles per unit volume), \(q\) is the scattering vector (\(q = 4\pi \sin(\theta)/\lambda\); \(\theta\) is the scattering angle, \(\lambda\) is the x-ray wavelength), and \(g(r)\) is the RDF. This formula works well for an isotropic system, but
is not suitable for an aggregated cluster system since its RDF is decreasing and do not tend to 1 at large distances. In order to resolve this issue, the 5000 NCP cluster RDF was fitted by a function $g_0(r) = 6.74 - 0.008r$ in the range of distances 300–500 Å (pink dotted line in figure 5 after scaling by factor 0.15), and the RDF of the 500 cluster system by function $g_0(r) = 15 - 0.03r$, and these functions were subtracted from $g(r)$ instead of 1 in equation (1). The resulting structure factors are plotted in figure 6; together with experimental structure factor from [11] obtained for NCP solutions with the three-valent natural polyanime, spermidine$^{3+}$ in a concentration range 7–40 mM. To the best of our knowledge, experimental SAXS data for precipitated NCP phases in the presence of CoHex$^{3+}$ have not been published. However data from precipitation assay studies of recombinant highly homogeneous NCPs [5] as well as our preliminary SAXS results on precipitated NCP systems in the presence of multivalent cations, show that the phase behaviour in the presence of spermidine$^{3+}$ and CoHex$^{3+}$ is very similar (Liu et al unpublished observations). In figure 6 one can see a remarkable similarity of the simulated and experimental structure factors, with the main peak at about 0.12 Å$^{-1}$, broader maxima at 0.18 and 0.23 Å$^{-1}$ and a shoulder at 0.07 Å$^{-1}$ though the latter is expressed well only in the ‘cluster’ 500 NCP system. The experimentally observed scattering factor was already divided by the form factor of a single NCP, thus the simulated scattering factor computed from core–core RDF allows for direct comparison of the simulated and experimental results and validation of the developed SCG NCP model.

We have investigated a few other structural properties for the aggregated NCP system modelled within the SCG model. We introduced a distance-dependent order parameter which is defined by the expression

$$S_{ord}(r) = \frac{3\cos^2\theta_r - 1}{2}$$

where $\theta_r$ is the angle between the normal vectors to planes of the two NCPs whose centres are separated from each other by a distance $r$. The order parameter defined by equation (2) is typically used to describe orientational ordering in liquid-crystal like systems. An order parameter equal to one means perfectly orientationally ordered NCPs which are all parallel to each other. Perpendicular orientations of NCPs contribute with $-1/2$ to the average value of the order parameter, and completely random orientation results in an order parameter equal to 0.

The distance-dependent order parameter computed from the simulated SCG NCP systems is displayed in figure 7(a). One can see that at contact distances 50–60 Å, the order parameter is high with values 0.4–0.6, indicating a strong correlation of neighbouring NCP orientation. The corresponding angular distribution for NCP separated by $60 \pm 1$ Å distance (normalized by an angular factor $\sin \theta$) has a well-expressed maximum at 0° corresponding to parallel orientations of NCPs with typical deviations within 30°, figure 7(b). The angular distribution at 100 Å distance between NCPs is more shallow but still, with a well-expressed maximum at 0. The orientational correlation between NCPs largely disappears at distances above 150 Å. One can conclude from these data that in the aggregated state, the NCPs have a local orientational order close to that of a hexagonal columnar phase with spread to distances within 150 Å (2–3 NCP units), but without a long-range order.

4. Conclusions

The present approach to multiscale CG simulations of NCP solutions has shown that starting from a detailed and fine-grained CG model of the NCP (1350 CG sites), we can use the IMC method on the trajectories of simulations of a rather small system of 10–20 such NCPs and map the result to a highly coarse-grained NCP model (7 CG sites), and calculate effective potentials within the ‘super-CG’ model. Simulations for 5000 (or more) NCPs using these effective potentials are highly feasible and when performed for potentials obtained for the presence of the three-valent CoHex$^{3+}$, results in precipitation to a single phase. The results should be an accurate representation
of the system under the conditions of the model and force field used for the advanced fine-grained CG model, which forms the basis for the IMC calculation of effective potentials between sites in the ‘super-CG’ model. Hence, the results are subject to the approximations within the advanced CG model, namely the dielectric continuum description of the solvent water and the absence of specific interactions such as hydrogen bonding and hydrophobic interactions.

The results are in qualitative (near quantitative) agreement with SAXS data for NCP precipitation in the presence of spermidine$^{3+}$ ions, which have shown formation of stacked NCPs in precipitates lacking long-range order [11]. The formation of the ordered hexagonal columnar phase of NCPs was observed only in a limited range of concentration of added spermidine$^{3+}$ salt. But it should be noted that the SAXS experiments detecting ordered NCP phases were generally performed after very long sample equilibration (months up to a year) and it is still not clear whether the stacked disordered NCP phases represents the true equilibrium structure of the system. In fact, the same caveat should be mentioned regarding the large-scale simulations. We performed preliminary simulations for the 4875 NCP system starting from a periodic columnar hexagonally oriented ‘super-CG’ NCPs (rather than starting from a random configuration). This simulation seems to maintain the ordered hexagonal structure over considerable simulation times. The issue of stability of each forms should be the subject of further studies, e.g. using some enhanced sampling procedures suitable for studies of phase transition, but preferably future advances should start from a refinement of the advanced CG NCP model itself (see below), which forms the basis for the IMC calculation of effective potentials within the ‘super-CG’ model.

In fact, one of the great advantages of the present approach is that further refinements can, in principle, be made in a straightforward way. The most obvious improvement of the approach seems to be to improve the interaction potentials of the advanced NCP model by performing all-atom simulations for a system of one or more NCPs and use the IMC method to calculate effective interaction potentials to be used in the advanced CG NCP model. Another important improvement would be considering P beads at the DNA entering points as a separate CG site type, in order to include effect related to axial asymmetry of the NCPs. Introduction of the ‘asymmetric’ super-CG NCP model will allow to capture mutual orientation of the dyad axes in the condensed NCP phases to reproduce dyad–dyad either ‘head-to-head’ or ‘head-to-tail’ positioning observed for aggregated NCPs experimentally [8, 45], analysed theoretically [46] and obtained in the simulations with the advanced CG NCP [34].

Acknowledgments
This work has been supported by the Singapore Agency for Science Technology and Research (A*STAR) through the Biomedical Research Council (LN), by the Singapore Ministry of Education Academic Research Fund (AcRF) through a Tier 1 and a Tier 3 Grant (LN) and by the Swedish Research Council VR (to APL).

References
[1] Hansen J C 2002 Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions Annu. Rev. Biophys. Biomol. Struct. 31 361–92
[2] Wolffe A P and Hayes J J 1999 Chromatin disruption and modification Nucleic Acids Res. 27 711–20
[3] Bertin A, Leforestier A, Durand D and Livolant F 2004 Role of histone tails in the conformation and interaction of nucleosome core particles Biochemistry 43 4773–80
[4] Mangenot S, Raspaud E, Tribet C, Belloni L and Livolant F 2002 Interactions between isolated nucleosome core particles. A tail bridging effect? Eur. Phys. J. E 7 221–31
[5] Liu Y, Lu C, Yang Y, Fan Y, Yang R, Liu C-F, Korolev N and Nordenskiöld L 2011 Influence of histone tails and H4 tail acetylations on nucleosome–nucleosome interactions J. Mol. Biol. 414 749–64
[6] Podgornik R 2003 Two-body polyelectrolyte-mediated bridging interactions J. Chem. Phys. 118 11286–96
chromatin fiber: evidence for a compact, interdigitated structure \textit{Proc. Natl Acad. Sci. USA} \textbf{103} 6506–11

[27] Song F, Chen P, Sun D, Wang M, Dong L, Liang D, Xu K-M, Zhu P and Li C 2014 Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units \textit{Science} \textbf{344} 376–80

[28] Scheffer M P, Eltsov M, Bednar J and Frangakis A S 2012 Nucleosomes stacked with aligned dyad axes are found in native compact chromatin \textit{in vitro} \textit{J. Struct. Biol.} \textbf{178} 207–14

[29] Castro-Hartmann P, Milla M and Daban J-R 2010 Irregular orientation of nucleosomes in the well-defined chromatin plates of metaphase chromosomes \textit{Biochemistry} \textbf{49} 4043–50

[30] Rippe K, Stehr R and Wedemann G 2012 Monte Carlo simulations of nucleosome chains to identify factors that control DNA condensation and access \textit{Innovations in Biomolecular Modeling and Simulations} ed T Schlick (London: Royal Society of Chemistry) pp 198–235

[31] Korolev N, Fan Y, Lyubarstev A P and Nordenskiöld L 2012 Modelling chromatin structure and dynamics: status and prospects \textit{Curr. Opin. Struct. Biol.} \textbf{22} 151–9

[32] Lyubarstev A P and Laaksonen A 1995 Calculation of effective interaction potentials from radial distribution functions: a reverse Monte Carlo approach \textit{Phys. Rev. E} \textbf{52} 3730–7

[33] Lyubarstev A P, Mirzaee A, Chen L-J and Laaksonen A 2010 Systematic coarse-graining of molecular models by the Newton inversion method \textit{Faraday Discuss.} \textbf{144} 43–56

[34] Korolev N, Fan Y, Lyubarstev A P and Nordenskiöld L 2013 An advanced coarse-grained nucleosome core particle model for computer simulations of nucleosome–nucleosome interactions under varying ionic conditions \textit{PLoS One} \textbf{8} e54228

[35] Castro-Hartmann P, Milla M and Daban J-R 2010 Systematic coarse-graining of molecular models by the Newton inversion method \textit{Faraday Discuss.} \textbf{144} 43–56

[36] Mangenot S, Leforestier A, Vachette P, Durand D and Livolant F 2002 Salt-induced conformation and interaction changes of nucleosome core particles \textit{J. Mol. Biol.} \textbf{324} 817–32

[37] Mangenot S, Leforestier A, Vachette P, Durand D and Livolant F 2003 X-ray diffraction characterization of the dense phases formed by nucleosome core particles \textit{Biophys. J.} \textbf{84} 2570–84

[38] de Frutos M, Raspaud E, Leforestier A and Livolant F 2001 Aggregation of nucleosomes by divalent cations \textit{Biophys. J.} \textbf{81} 1127–32

[39] Leforestier A, Dubochet J and Livolant F 2001 Bilayers of nucleosome core particles \textit{Biophys. J.} \textbf{81} 2414–42

[40] Korolev N, Berezhnoy N V, Eom K D, Tam J P and Nordenskiöld L 2012 A universal description for the experimental behavior of salt-(in)dependent oligocation-induced DNA condensation \textit{Nucleic Acids Res.} \textbf{40} 2808–21

[41] Korolev N, Allahverdi A, Lyubarstev A P and Nordenskiöld L 2012 The polyelectrolyte properties of chromatin \textit{Soft Matter} \textbf{8} 9322–33

[42] Finch J T, Lutter L C, Rhodes D, Brown R S, Rushton B, Levitt M and Klug A 1977 Structure of nucleosome core particles of chromatin \textit{Nature} \textbf{269} 29–36

[43] Finch J T, Brown R S, Richmond T J, Rushton B, Lutter L C and Klug A 1981 X-ray diffraction study of a new crystal form of the nucleosome core showing higher resolution \textit{J. Mol. Biol.} \textbf{145} 757–69

[44] Leforestier A and Livolant F 1997 Crystal structure of the nucleosome particles of chromatin \textit{Nature} \textbf{389} 251–60

[45] Davey K A, Sargent D F, Luger K, Maeder A W and Richmond T J 2002 Solvent mediated interactions in the structure of nucleosome core particle at 1.9 Å resolution \textit{J. Mol. Biol.} \textbf{319} 1097–113

[46] Vasudevan D, Chua E Y and Davey C A 2010 Crystal structures of nucleosome core particles containing the ‘601’ strong positioning sequence \textit{J. Mol. Biol.} \textbf{403} 1–10

[47] Leforestier A and Livolant F 1997 Liquid crystalline ordering of nucleosome core particles under macromolecular crowding conditions: evidence for a discotic columnar hexagonal phase \textit{Biophys. J.} \textbf{73} 1771–6

[48] Schalch T, Duda S, Sargent D F and Richmond T J 2005 X-ray structure of a tetranucleosome and its implications for the chromatin fibre \textit{Nature} \textbf{436} 138–41

[49] Dorigo B, Schalch T, Kulangara A, Duda S, Schroeder R R and Richmond T J 2004 Nucleosome arrays reveal the two-start organization of the chromatin fiber \textit{Science} \textbf{306} 1571–3

[50] Robinson P J J, Fairall L, Huynh V AT and Rhodes D 2006 EM measurements define the dimensions of the ‘30-nm’
[45] Livolant F, Mangenot S, Leforestier A, Bertin A, de Frutos M, Raspaud E and Durand D 2007 Are liquid crystalline properties of nucleosomes involved in chromosome structure and dynamics? Phil. Trans. R. Soc. 364 2615–33

[46] Lorman V, Podgornik R and Žekš B 2005 Correlated and decorrelated positional and orientational order in the nucleosomal core particle mesophases Europhys. Lett. 69 1017–23