John von Neumann Institute for Computing

Global Motions in the Nucleosome Explored Using a Coarse-Grained Model

K. Voltz, J. Trylska, V. Tozzini, V. Kurkal-Siebert,
J. C. Smith, J. Langowski

published in

From Computational Biophysics to Systems Biology (CBSB08),
Proceedings of the NIC Workshop 2008,
Ulrich H. E. Hansmann, Jan H. Meinke, Sandipan Mohanty,
Walter Nadler, Olav Zimmermann (Editors),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **40**, ISBN 978-3-9810843-6-8, pp. 141-144, 2008.

http://www.fz-juelich.de/nic-series/volume40

Global Motions in the Nucleosome Explored Using a Coarse-Grained Model

Karine Voltz[1,2], Joanna Trylska[3], Valentina Tozzini[4], Vandana Kurkal-Siebert[5], Jeremy C. Smith[2,6], and Joerg Langowski[1]

[1] Biophysics of Macromolecules, German Cancer Research Center, Heidelberg, Germany

[2] Computational Molecular Biophysics, IWR, University of Heidelberg, Germany
E-mail: Karine.Voltz@iwr.uni-heidelberg.de

[3] Interdisciplinary Center for Mathematical and Computational Modelling, University of Warsaw, Warsaw, Poland

[4] NEST-INFM-Scuola Normale Superiore, Pisa, Italy

[5] Molecular Modelling, Polymer Physics, BASF Aktiengesellschaft, Ludwigshafen, Germany

[6] Center for Molecular Biophysics, Oak Ridge National Laboratory, USA

The nucleosome is the basic compacting unit of chromatin, a complex structure that enables DNA to fit in the eucaryotic cell nucleus. Because many biological processes require free DNA as a substrate, the nucleosome has to undergo conformational transitions to allow the DNA target sites to be exposed. To obtain insight into the global dynamics of the nucleosome, microsecond timescale coarse-grained molecular dynamics is performed. Here we report a principal component analysis (PCA) realised on a 5μs coarse-grained molecular dynamics trajectory to identify the global motions obtained in the nucleosome.

1 Introduction

In the nucleosome, 147-bp DNA are wrapped almost twice around a protein core consisting of eight histone proteins, one tetramer H3-H4 and two dimers H2A-H2B. The histone protein core, composed mainly of α-helices, is very stable in contrast to the non-structured histone N-terminal tails, that pass between the DNA superhelix turns. Despite the high number of interactions between nucleosomal DNA and the surface of the protein core, nucleosomal DNA can become free to be processed by DNA binding proteins involved in DNA transcription, replication, recombination or repair. The mechanisms underlying DNA target site exposure are still under debate. One proposed mechanism, based on single-nucleosome FRET evidence, postulates that significant pieces of nucleosomal DNA (up to 70 bp) can transiently detach from the protein core[1,2]. These reversible conformational changes, believed to occur on the 50-250 ms timescale, could be responsible for the progressive invasion of the nucleosome by DNA-binding proteins.

To characterize the conformational transitions leading to DNA accessibility, it is important to understand nucleosome dynamics at equilibrium and on a long timescale. For this purpose, we performed coarse-grained (CG) molecular dynamics (MD) simulations using a model specifically developed for the nucleosome[3]. Classical all-atom MD of large systems such as the nucleosome is currently limited to about 100 nanosecond timescale, while with the present model of the nucleosome, a one-bead representation together with the absence of explicit water, MD simulations of the nucleosome can be carried out over

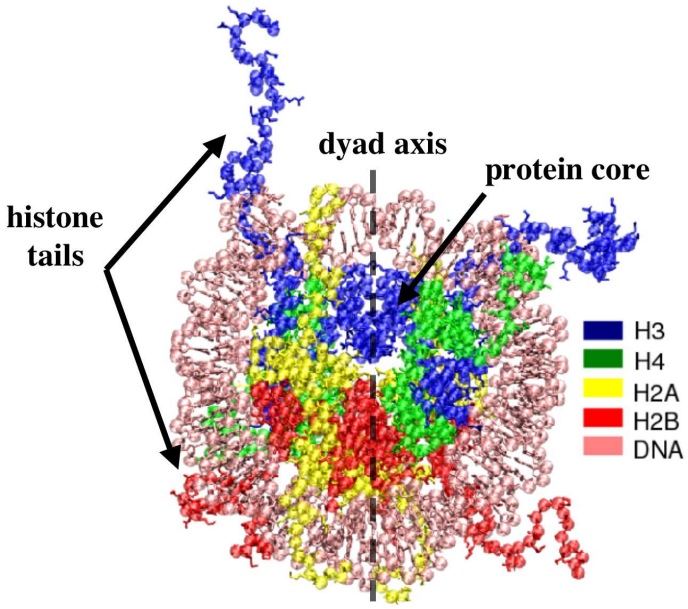

Figure 1. All-atom and coarse-grained representation of the nucleosome (structure 1KX5). The residues are represented by single spherical beads centered on α-carbon for amino acids and on phosphorus for nucleic acids. The histones and the DNA are represented in different colors. The nucleosome is symmetric with respect to the so-called dyad axis.

several microseconds. In the present proceedings, after briefly presenting the CG model, we report a principal component analysis (PCA) performed on a 5μs CG MD trajectory enabling the identification of slow collective motions in the nucleosome.

2 Method

A detailed description of the CG model and its force-field parameterization is reported in a previous study[3]. In this CG model, protein residues and DNA nucleotides are represented as single beads (Figure 1) interacting through harmonic (for neighboring) or Morse (for nonbonded) potentials which depend on the interbead distances. This model shares similarity with Gaussian network models (GNM) but, in contrast to GNM, the Morse description of nonbonded interactions allows realistic anharmonic dynamics of the system. Force-field parameters were estimated by Boltzmann inversion of the corresponding radial distribution functions computed from a reference 5-ns all-atom MD simulation and further refined to obtain agreement with all-atom MD root-mean-square fluctuations (RMSF).

Coarse-grained MD simulations were performed using the DL_POLY package[5]. The starting structure used was the 1KX5 structure[4] energy-minimized in solvent using the

Charmm[6]27 all-atom force-field. A Langevin bath was used to account for the frictional and stochastic effects of the solvent. The reduction of the system allowed us to apply a 20-fs timestep. The simulation was carried out at 300 K during 5 μs.

Principal component analysis, also called essential dynamics analysis[7], provides a way to identify the most significant directions of motions in the system. The method consists in diagonalizing the symmetric 3N$\times$3N covariance matrix derived from the MD trajectory, whose elements are defined as $C_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle$ where x_i and x_j are the Cartesian coordinates of the atom i and j. The resulting eigenvectors give the direction and their corresponding eigenvalues quantify the magnitude of the fluctuations. Eigenvectors with the highest eigenvalues are called the principal components or principal modes. Here, PCA was performed with the GROMACS package on a 5-μ CG MD trajectory. The tails were excluded from the analysis since their high mobility masked the motion of the DNA and the protein core.

2.1 Results and Discussion

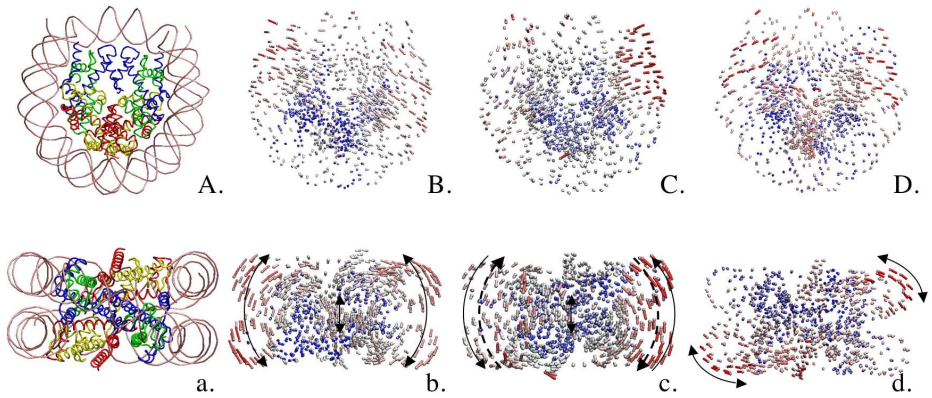

Figure 2. Global motions of the nucleosome. A and a: initial nucleosome 1KX5 structure. B–D and b–d: nucleosome conformations deformed along the first three modes. Each residue is colored with respect to its root-mean-squared fluctuation (RMSF) along the mode; blue and red correspond to low and high RMSFs, respectively.

Figure 2 shows the directions of the motions for the first three principal components. Residues are colored with respect to their RMSF along each eigenvector. The ten eigenvectors with the largest eigenvalues describe 22 % of the total protein motion, a smaller fraction than what is usually observed for other systems such as proteins. This result may be the consequence that important structural transitions in the nucleosome occur on a larger timescale than 5μs.

The motion along the first mode corresponds to a bending out of the plane of the nucleosome that mostly involves i) both extremities (ten basepairs) of the DNA superhelix; ii) the facing DNA stretches located on the respective opposite turn; and iii) amino acids of H3, H3', H4 and H4' interacting with these regions. Simultaneously to the bending of this

region, several residues centered around the dyad axis (i.e., ≈10-bp of DNA and residues from H3 and H3') translate in the opposite direction to the out-of-plane bending motion. The most rigid regions along this mode are residues that belong to the histones H2B and the DNA basepairs interacting with them.

The second principal component describes a rocking motion center of which is located on the dyad axis. As in the first mode, this mostly involves both extremities of the DNA superhelix, the facing DNA stretches located on the respective opposite turn, and amino acids interacting with these regions, while histones H2B and interacting DNA residues contribute less to this motion.

Finally, the third principal mode, represents a deformation of the nucleosome in the plane of the nucleosome. In this motion, the DNA superhelix extremities and amino acids interacting with these regions participate in the stretching of the nucleosome in a direction perpendicular to the dyad axis while H2B residues and the DNA part interacting with them contributes to stretching the nucleosome in the direction of the dyad axis.

3 Outlook

Further work is ongoing to investigate, in particular, the spontaneous unwrapping of nucleosomal DNA extremities from the nucleosome protein core. We also plan to study the interactions between several nucleosomes in a nucleosomal array within the context of the chromatin fiber.

References

1. G. Li, M. Levitus, C. Bustamante and J. Widom, *Rapid spontaneous accessibility of nucleosomal DNA* Nat. Struc. Mol. Biol. **12**, 46–53, 2005.
2. M. Tomschik, H. Zheng, K. van Holde, J. Zlatanova, S. H. and Leuba, *Fast, long-range, reversible conformational fluctuations in nucleosomes revealed by single-pair fluorescence resonance energy transfer* Proc. Natl. Acad. Sci. U.S.A. **102**, 3278–3283, 2005.
3. K. Voltz, J. Trylska, V. Tozzini, V. Kurkal-Siebert, J. Langowski and J. Smith, *Coarse-grained force field for the nucleosome from self-consistent multiscaling* J. Comput. Chem. **29**, 1429–1439, 2008.
4. C. A. Davey, D. F. Sargent, K. Luger, A. W. Maeder and T. J. Richmond, *Solvent mediated interactions in the structure of the nucleosome core particle at 1.9Å* J. Mol. Biol. **319**, 1097–1113, 2002.
5. W. Smith, T. R. Forester, *DLPOLY2.0: A general-purpose parallel molecular dynamics simulation package* J. Mol. Graph. **14**, 136–141, 1996.
6. Jr. A.D. MacKerell, N. Banavali, and N. Foloppe, *Development and current status of the CHARMM force field for nucleic acids* Biopolymers **56**, 257–265, 2001.
7. A. Amadei, A. B. M. Linssen, H. J. C. Berendsen *Essential dynamics of proteins* Proteins **17**, 412-425, 1993.