Effect of A-tract-mediated linker histone orientation on trinucleosome compaction

Madhura De¹²³, Martin Würtz¹⁴, Gabriele Müller¹, Katalin Tóth¹⁵*, Rebecca C. Wade²³⁴⁶*

¹Division of Biophysics of Macromolecules, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany
²Molecular and Cellular Modelling Group, Heidelberg Institute for Theoretical Studies (HITS), 69118 Heidelberg, Germany
³Faculty of Biosciences, Heidelberg University, 69120 Heidelberg, Germany
⁴Center for Molecular Biology (ZMBH), DKFZ-ZMBH Alliance, Heidelberg University, 69120 Heidelberg, Germany
⁵Present address: Department of Biophysics and Cell Biology, Faculty of Medicine, University of Debrecen, Hungary
⁶Interdisciplinary Center for Scientific Computing (IWR), Heidelberg University, 69120 Heidelberg, Germany

*To whom correspondence should be addressed. Emails: Katalin Tóth: dr.toth.katalin@med.unideb.hu, Rebecca C. Wade: rebecca.wade@h-its.org

Abstract

Linker histones (LH) have been shown to preferentially bind to AT-rich DNA, particularly A-tracts, contiguous stretches of adenines. Using spFRET (single pair Förster/Fluorescence Resonance Energy Transfer), we recently found that the globular domain (gH) of Xenopus laevis H1.0b LH orients towards A-tracts on the linker-DNA (L-DNA) while binding on-dyad in LH:mononucleosome complexes. Here, we investigate the impact of this A-tract-mediated orientation of the gH on the compaction of higher-order structures by studying trinucleosomes as minimal models for chromatin. Two 600 bp DNA sequences were constructed, each containing three Widom 601 core sequences connected by about 40 bp L-DNA but differing in the positioning of A-tracks on either the outer or the inner L-DNAs flanking the first and the third Widom 601 sequences. The two inner L-DNAs were fluorescently labelled at their midpoints. Trinucleosomes were reconstituted using the doubly labelled 600 bp DNA, core histone octamers and the full-length H1.0b LH. SpFRET was performed for a range of NaCl concentrations. While the LH compacted the trinucleosomes, the extent of compaction was similar for the two DNA sequences. Modeling constrained by the measured FRET efficiency suggests that the trinucleosomes adopt a zig-zagged two-helical start arrangement with the first and third nucleosomes stacked on top of each other, but with the first and third LHS insufficiently close to interact and affect compaction. Thus, despite differences in the positioning of the A-tracks in the sequences studied, LH binding compacts the corresponding trinucleosomes similarly.
Highlights

1. Does the A-tract-mediated orientation of the gH affect chromatin structure?
2. Trinucleosomes were reconstituted with A-tracts on the outer or inner linker DNAs.
3. spFRET-based modeling shows LH compacts trinucleosomes to a zig-zag conformation.

Keywords: chromatin structure, linker histone, nucleosome, single-pair FRET, A-tract DNA

Abbreviations: NRL: Nucleosome Repeat Length, spFRET: single pair Förster/Fluorescence Resonance Energy Transfer, L-DNA: linker DNA, LH: Linker histone, gH: globular (head) domain of the LH, CTD: C-terminal domain of the LH, AFM: Atomic Force Microscopy

The binding of the highly positively charged linker histone (LH) H1 to the nucleosome plays an important role in the compaction of chromatin [1]. The LH associates with the nucleosome in the region bounded by the linker-DNA (L-DNA) arms entering and exiting the nucleosome core particle ([2–6]. The LH has a tripartite structure composed of a conserved globular domain (gH), a 100 residue long, intrinsically disordered C-terminal domain (CTD) and a shorter N-terminal domain [7]. The CTD has been observed to associate with one [5,8] or both L-DNA arms [6] experimentally, and in a salt-dependent manner with one (low salt) or both (high salt) L-DNAs in coarse-grained simulations [9]. The association of the CTD with the L-DNAs has also been suggested to have an effect on the charge distributions of nucleosomal arrays [5]. The gH has been observed to bind the nucleosome in an on-dyad [3–6] or an off-dyad [8,10] position, or a dynamic combination of the two [11]. The positioning of the gH has been suggested, from both experimental [7,9] and computational studies [13], to affect the compaction of nucleosomal arrays. In this work, we investigate whether the orientation in which the gH binds on-dyad to the nucleosome affects chromatin compaction.

To address this question, the gH of LHs should be made to associate with the nucleosomes in a specific manner, either by mutating the gH domain residues that determine on- or off-dyad binding [10], or by exploiting the preference of the LH gH for AT-rich DNA or A-tracts (a series of contiguous adenes) [6,14–25]. Here, we built on our earlier study of mononucleosome-LH complexes by spFRET [24] that showed that the gH domain of a full-length LH recognizes an 11 bp A-tract on the L-DNA by binding on-dyad in a specific orientation (Figure 1B ii). We therefore here addressed the question of whether such an A-tract mediated orientation of the gH domain affects the compaction of nucleosomal arrays. For this purpose, we constructed two trinucleosomes. Each trinucleosome had a 600 bp DNA sequence with three strongly nucleosome positioning Widom 601 sequences [26] situated about 40 bp apart from each other, thereby resulting in a nucleosome repeat length (NRL) of 187 bp. The flank sequences of nucleosomes (Nuc) 1 and 3 were either A-tract (11 contiguous adenes complementary to thymines) or purely GC-tract. Depending upon the location of the A-tracts (Table 1), the trinucleosome constructs were named A-far (A-tracts on outer L-DNA and GC-tracts on inner L-DNA) or A-near (A-tracts on inner L-DNA and GC-tracts on outer L-DNA). Both the constructs were doubly labelled at the mid-points of the two inner L-DNAs with donor and acceptor fluorophores (Figure 1A).
Figure 1: A. Schematic illustration of the trinucleosome systems studied. The three nucleosomes, Nuc 1, 2, and 3 are shown with the outer and inner L-DNAs labelled. The LH is not shown. The outer (blue) and inner (grey) flanks of the L-DNA arms are shown by cylinders. Fluorophores were attached to the methyl group of the thymine residue at the mid-point of each of the inner L-DNAs and are represented by green (Alexa 488 or donor) and red (Alexa 594 or acceptor) pentagons. B. (i) Cartoon representation of the globular domain (gH) of the LH (PDB ID 5NL0 [5]) showing three DNA interacting zones, A, B and C [24]. Zone B can interact with and mediate recognition of A-tracts. The start of the C-terminal domain (CTD) is represented by a purple circle. The start of the N-terminal domain is labelled as NTD. (ii) Binding of the LH H1.0b gH to a nucleosome with an A-tract flank on one of the L-DNA arms. The on-dyad gH is oriented towards the A-tract (dA in orange, dT in blue) minor groove on the L-DNA, with interactions mediated by two arginines in zone B [24]. The gHs that bind to Nuc 1 and Nuc 3 in the trinucleosome constructs are expected to be similarly oriented to the A-tract flanking sequences.

Table 1: Flank sequences of the outer and inner L-DNA arms for the A-near and A-far trinucleosomes

<table>
<thead>
<tr>
<th>Trinucleosome construct</th>
<th>Inner L-DNA flank sequence</th>
<th>Outer L-DNA flank sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-near</td>
<td>5' - AAAAAAAAAAAA -3'</td>
<td>5' - GGGCGGCCGGCG -3'</td>
</tr>
<tr>
<td></td>
<td>3' - TTTTTTTTTTTTTTTTTTTTT</td>
<td>3' - CCCGCCGGCGCGCGCGCGCG</td>
</tr>
<tr>
<td>A-far</td>
<td>5' - GGGCGGCCGGCG -3'</td>
<td>5' - AAAAAAAAAAAA -3'</td>
</tr>
<tr>
<td></td>
<td>3' - CCCGCCGGCGCGCGCGCGCGCG</td>
<td>3' - TTTTTTTTTTTTTTTTTTT</td>
</tr>
</tbody>
</table>

The 600 bp constructs were constructed using a modified Golden-Gate assembly [27–29] in vitro that ensured that the three DNA fragments corresponding to the three
nucleosomes were correctly oriented with respect to each other (see Supplementary Information section 1, Preparation of 600 bp DNA). After gel extraction and purification of the A-far and A-near DNA sequences, reconstitution of the trinucleosomes was performed with and without LH and verified by 1% agarose gel electrophoresis and atomic force microscopy (AFM), see Supplementary Information section 2 (Reconstitution of trinucleosomes) and references [24] and [29] for details.

Single-laser excitation spFRET measurements [24,30–36] were carried out for the freely diffusing trinucleosomes as described previously. To study the stability and observe how the trinucleosomes dissociate with an increase in salt concentration, and how the presence of the LH affects such dissociation [37–39], experiments were performed at NaCl concentrations ranging from 25 to 700 mM NaCl, with 10 mM Tris and 0.1 mM EDTA (ethylene diamine tetraacetic acid) at pH 7.5. From spFRET measurements, we obtained the proximity ratio, P, which is directly proportional to the FRET efficiency and computed by dividing the number of photons detected in the acceptor channel (Na) by the total number of photons detected in the acceptor and donor (Nd) channels, i.e. \( P = N_a/(N_a+N_d) \) [31,33,34]. For each system, proximity ratio histograms were fitted to multiple Gaussian peaks, see Supplementary Information section 3. Then structural models of the gH-bound trinucleosomes, corresponding to the major high-P peak, were built using the Chimera [40] and FRET Positioning and Screening (FPS) [41,42] software. The results of the spFRET experiments and the corresponding models are presented below.

**The linker histone H1.0b compacts trinucleosomes similarly irrespective of the placement of A-tract flanking sequences**

The proximity ratio histograms obtained for the systems studied at seven different salt concentrations are shown in Figure 2 and Supplementary Information section 3B, (Figures S4-S10). The change in the proximity ratio histograms in both the LH containing trinucleosomes (A-near and A-far) is similar at all salt concentrations, as shown in Figure 2A and B. In the absence of LH, the mean of the highest proximity ratio peak varies from 0.34 (low salt) to about 0.68 at 250 mM NaCl concentration (Figure 2B iii). With further increasing salt concentration, the mean proximity ratio of this peak falls to about 0.24 (700 mM), suggesting that dissociation of the trinucleosome occurs. In the presence of LH (Figure 2Bi and ii) and at NaCl concentrations from 25 mM to 300 mM, the highest proximity ratio peak is fit by two Gaussians (Supplementary figures S4-10B) with similar populations, that we designate as populations I and II (with II being the Gaussian with the higher proximity ratio). Population II (black trace in Figure 2Bi-i-ii) maintains a stable mean proximity ratio between 0.85 and 0.93 with a small peak width, indicating a highly compacted trinucleosome with restricted movement of the inner L-DNA. While the LH is present in population I as well, as suggested by its high mean proximity ratio values at low salt concentration, this population could represent trinucleosomes having a somewhat relaxed compaction with the larger peak width indicating greater diversity in the positioning of the fluorophores on the inner L-DNA arms. The possibility that population I contains less than three LHS per trinucleosome at low salt concentration cannot be discounted. Population II diminishes at 300 mM salt concentration, and at salt concentrations of 550 and 700 mM, trinucleosomes with and without LH all show only one population with a peak with a similar mean proximity ratio in the presence and absence of LH. This suggests that the LH starts to dissociate at salt concentrations of 300 mM and completely dissociates at salt concentrations of 550 mM and 700 mM.
For the trinucleosomes with LH (Figure 2.A and Bi and ii), the mean proximity ratios of the selected peaks are higher than those of the selected peaks for the trinucleosomes without LH (Figure 2.A and B iii), at salt concentrations from 25 mM to 250 mM. This difference confirms that the LH compacts the trinucleosomes, bringing the inner L-DNA arms closer to each other. Shielding of the negative charges on the DNA backbone by the positively charged sodium ions also brings the two inner L-DNA arms closer to each other in trinucleosomes lacking LH (Figure 2.Biii). This increase in compaction as a result of charge shielding by the salt is more pronounced in trinucleosomes lacking LH (Figure 2C iii and iv) than trinucleosomes associated with LHS (Figure 2C i and ii).

Figure 2: spFRET measurements of the compaction of trinucleosomes, A-far and A-near, with and without LH. A. Heat map representations of proximity ratio histograms measured by spFRET under seven NaCl concentrations ranging from 25 to 700 mM for (i) A-far with LH, (ii) A-near with LH, and (iii) A-near without LH. The colour intensity represents the normalised frequency of the proximity ratio value. B. Plots of mean proximity ratio of populations I and II of the high proximity ratio major population peaks (see Supplementary Information, section 3B, 1-7) against salt concentration. The marker size and label show the % population of the peak. (i) A-far with LH, (ii) A-near with LH, (iii) A-near without LH. The proximity ratio histograms were similar for A-far without LH to those for A-near without LH (see Supplementary Information, section 3B). C. Proximity ratio (P) histograms measured at 25 mM (solid line) and 150 mM (bars) NaCl concentration for (i) A-far with LH, (ii) A-near with LH, (iii) A-near without LH, and (iv) A-far without LH. (Histograms measured for other salt concentrations are given in see Supplementary Information, section 3B).
Our results indicate that the LH compacts the A-near and A-far trinucleosomes similarly at 25 mM salt. The proximity ratio histograms for the other salt concentrations show that this similarity between A-near and A-far trinucleosomes holds over the range of salt concentrations studied.

Next, we focused on population II, which contains the Gaussian peaks with the higher proximity ratios and lower widths, to investigate what three-dimensional structure it corresponds to and why, despite the different positions of the A-tracts, LH binding does not appear to affect the relative extent of compaction of the A-far and A-near trinucleosomes. The mean proximity ratio of population II only varies from 0.8 to 0.9, indicating little variation in the interfluorophore distance over the population [43]. Thus, for each of the two DNA sequences studied, we built a single representative trinucleosome model that has a computed FRET efficiency between the inner L-DNA arm fluorophores of 0.8.

**The modeled structures of LH-bound trinucleosomes with a high FRET efficiency have a 2-start helical arrangement**

To model the trinucleosomes, the experimentally determined starting structure of the mononucleosome-linker histone complex, PDB 5NL0 [5], was mutated to the sequence of Nuc 2 using Chimera [40]. This structure of Nuc 2 was then subjected to elastic network normal mode analysis using the elNémo webserver [44] to generate a range of arm-opening conformations [22–24]. Accessible volume simulated fluorophores were attached to the generated ensemble on the labelling positions, using the FPS software [41,42], with dye parameters described previously [24,41,42,45]. A model that showed a computed FRET efficiency proximity ratio of 0.8 was selected from the ensemble. Then, two LH-mononucleosome structures (two more copies of the PDB structure (PDB 5NL0)) were added to the structure of Nuc 2, and the resulting DNA was renumbered and mutated to that of the A-near and A-far sequences using the Chimera ‘swapna’ command [40]. The gH domain was positioned on Nuc 2 in a canonical on-dyad fashion as observed in the structures: 4QLC [3], 5WCU [4], 5NL0 [5], and 7K5X [6]. The gH domains on Nuc 1 and Nuc 3 were positioned according to the AG and GA models described in our earlier work [24], i.e., on-dyad, with Zone B proximal to the A-tract minor groove (as in Figure 1B ii).

While building the models, apart from the single spFRET restraint, we also ensured that the L-DNAs were not kinked or bent at an angle. Thus, as the inner L-DNAs are straight, Nuc 1 and 3 stack on top of each other (Figure 3), resembling the canonical two-start helical arrangement [46] as predicted and observed for 187 NRL nucleosomes [47–49]. The stacking is optimal as confirmed by aligning the modelled structure (for A-near and A-far) to the structures 6HKT [48] and 6L49 [50]. The stacking plane between Nuc 1 and 3 (Figure 3Bi in green) was used to calculate the screw angle between Nuc 2 and the stacked Nuc 1 and 3, which was 73° (Figure 3Bi i). The dyad axes of Nuc 1 and 3 were rotated by 28° with respect to each other (Figure 3Bi iii).

To understand why, despite having A-tracts at different positions (either on the inner or on the outer L-DNA arms of Nuc 1 and Nuc 3), the two trinucleosomes were compacted similarly by the LH, we focused on the interactions of the gHs on Nuc 1 and Nuc 3.
**Figure 3:** Modelling of a representative structure of population II of the A-near and A-far LH-bound trinucleosomes. **A.** Two perpendicular views of a model with a computed FRET efficiency between the labels on the inner L-DNA arms of 0.8. The outer L-DNAs are truncated for clarity. The green arrows represent energy transfer between the fluorophore labels (which are not shown). Nuc 1 and nuc 3 stack. The trinucleosome was built using the three copies of a LH-monomonucleosome structure, hence the discontinuities in the inner L-DNA arms. **B.** Geometric parameters defining the structure of the model computed similarly to Korolev et al. (2018) [51] using Chimera [40]. (i) Three octamer planes shown in pink with the stacking plane between Nuc 1 and Nuc 3 shown in green. The stacking plane was defined in Chimera [40] by choosing the side chain carbon atoms of the residues on the core histone octamers of Nuc 1 and Nuc 3 facing each other: Arg 134 from the H3 histone, Val 21 from the H4, Asn 68 from H2A, and Ser 120 from H2B. (ii) The angle between the stacking plane and Nuc 2 is 73° and between the octamer planes of Nuc 1 and Nuc 3 is 20°. (iii) The angle between the dyad axes of Nuc 1 and Nuc 3 is 28°. The dyad axes were defined in Chimera [40] on each nucleosome by the vector between the mid-point of the two dyad bases and Arg 134 of the two H3 proteins in the core. Axes were defined using the backbone atoms of the selected residues [40]. **C.** Stacking of LH-bound Nuc 1 and Nuc 3 showing the different relative positioning of the gHs. Left: A-near and Right: A-far. Each LH is shown with its gH positioned on-dyad and oriented towards the A-tracts [24]. The LH gH is colour coded as in Figure 1. The core histone octamers are shown as grey surfaces, the DNA is pale green. The A-tracts are shown in stick representation with adenosines in orange and thymines in blue. The tips of the L-DNA containing the GC-tract are coloured cyan and deep blue.
The orientations of the linker histones on stacking nucleosomes do not influence trinucleosome compaction

Although we did not directly observe the location of the full-length LH in our experiments, we assume that the gHs (of LH of subtype H1.0b) associate in an on-dyad position [5,6,52] in an orientation like we observed for A-tract-containing mononucleosomes [24], shown in Figure 1Bii. The models for A-near and A-far with the LHs positioned in this way are shown in Figure 3C. In A-near (Figure 3C left), the A-tracts are on the inner L-DNA and in A-far, they are on the outer L-DNA (Figure 3C right). For the gH to affect compaction of a nucleosomal array, it would be expected to self-associate [8] or form salt-bridges. In our models, the gH domains on Nuc 1 and Nuc 3 are not close enough to each other to self-associate. These models suggest that the orientation of the gH domains of the LH H1.0b does not play a role in the compaction of the trinucleosomes.

We next considered the role of the LH CTDs. The purple circles in Figure 3C represent the C-terminus of the gH domain, corresponding to the start of the 100-residue long intrinsically disordered CTD tail. If the CTDs associate with the L-DNA arm to which the C-terminus of the gH (purple circle) is closest, then the CTD would associate with the inner L-DNA arm of A-near and with the outer L-DNA arm of A-far. In this case, we would expect that the highly positively charged CTD would bring the inner L-DNAs closer together in A-near than A-far, resulting in a higher FRET efficiency between the fluorophores.

In our previous work on mononucleosomes, we have observed that position 101 (6 residues into the CTD from the gH) is nearly equidistant from both L-DNA arms [24]. Zhou et al. (2021) [6] observed that the CTD of LH H1.0b (PDB 7K5X) associates with both L-DNA arms. But coarse grained simulations have shown that the CTD can associate with one L-DNA at low salt (5 and 15 mM), and with both L-DNA arms at high salt (80 and 150 mM) concentrations [9]. If this is the case, then we would expect that, at low salt concentrations, A-near would be more compact compared to A-far. In Figure 2B, the mean proximity ratio peak of Population I of A-near with LH is slightly higher than A-far with LH at salt concentrations of 25, 150 and 250 mM (difference, ΔPmean, of about 0.1, Supplementary Information Section 3B), whereas the more compact Population II shows similar mean proximity ratios (ΔPmean ~ 0.05) for both A-near and A-far. This difference in proximity ratio would only correspond to a small difference in interfluorophore distance. Thus, these data provide little or no support for a strong association of the CTD with the inner L-DNA arms of A-near.

Overall, considering both populations I and II, the mean proximity ratios of the two trinucleosome sequences remain similar as the salt concentration is raised, suggesting that the LH CTDs associate with both inner and outer L-DNA arms of Nuc 1 and Nuc 3 and, therefore, similarly contribute to the compaction of the two trinucleosomes, despite the different locations of the A-tracts.

Conclusions

Our study confirms that the presence of LHS compacts trinucleosomes (Figure 2). We find that the different locations of the A-tracts in the DNA sequences of the two trinucleosomes studied, do not affect LH-mediated trinucleosome compaction (Figure 2), despite the expected A-tract-mediated orientation of the LH gH domains [24] (Figure 1B ii). Modeling constrained by the experimental observations (Figure 3) indicates that this is
because the gH domains of the stacked nucleosomes in the trinucleosomes are too far away to associate. Furthermore, analysis of the spFRET data at different salt concentrations, indicates that the LH CTDs contribute to trinucleosome compaction but that their interaction with the L-DNA arms of the nucleosomes is rather insensitive to the locations of the inserted A-tracts. These results thus point to a robustness of chromatin structure to heterogeneity at the level of DNA sequence and individual nucleosomes. However, further study would be necessary to assess the generality of our observations for other DNA and linker histone sequences.

Here, we studied the LH H1.0b isoform, whose gH has been observed to associate in an energetically favourable on-dyad fashion [5,6,11,52]. A different LH isoform with an off-dyad gH binding mode or an altered affinity, due to differences in sequence or post-translational modification in the gH or CTD, may show a difference sensitivity to L-DNA sequence and A-tract insertions, as regards its ability to compact chromatin. Likewise, less strongly nucleosome positioning sequences than the Widom sequences studied here may result in altered sensitivity of LH-mediated compaction to A-tract insertions in the L-DNA. Furthermore, we studied two sequences with differently positioned A-tracts but clearly, further variations on DNA position and sequence could be studied, and additional positions for fluorophore labelling could be used to provide more structural constraints. Nevertheless, this study shows how the preference of a LH for A-tracts can be exploited to probe the determinants of chromatin compaction, as well as to understand to what extent LHs – as ubiquitous proteins – regulate chromatin structure, and thus transcription.

Acknowledgements The authors are very grateful to Prof. Jeffrey Hayes and Dr. Amber Cutter for providing us with the plasmid for wild-type linker histone. The authors are grateful to Nathalie Schwarz for purifying the linker histone. This work was supported by the Helmholtz International Graduate School for Cancer Research (DKFZ) and the Klaus Tschira Foundation.

Author contributions MD, KT and RCW conceived the research. MD, GM, KT designed the DNA constructs. MD and GM prepared and purified the DNA. MW acquired the AFM images. MD performed the spFRET measurements with suggestions from KT, and the modeling, with suggestions from RCW, and wrote the paper with suggestions from all the authors. All the authors contributed to the finalisation of the manuscript.

Data availability Data is available in this repository: https://doi.org/10.5281/zenodo.5090768. This repository includes the models of A-far and A-near (.pdb) and raw data from the spFRET experiments (.t3r) in different salt concentrations, along with the donor-only and A-near DNA sample.

References


