Please cite this article in press as: Öztürk et al., Toward an Ensemble View of Chromatosome Structure: A Paradigm Shift from One to Many, Structure (2018), https://doi.org/10.1016/j.str.2018.05.009

Structure
Perspective

Toward an Ensemble View of Chromatosome Structure: A Paradigm Shift from One to Many

Mehmet Ali Öztürk,[1,2] Vlad Cojocaru,[3,4] and Rebecca C. Wade[1,5,6,*]
[1]Molecular and Cellular Modeling Group, Heidelberg Institute for Theoretical Studies (HITS), 69118 Heidelberg, Germany
[2]The Hartmut Hoffmann-Berling International Graduate School of Molecular and Cellular Biology (HBIGS), Heidelberg University, 69120 Heidelberg, Germany
[3]Computational Structural Biology Laboratory, Department of Cellular and Developmental Biology, Max Planck Institute for Molecular Biomedicine, 48149 Münster, Germany
[4]Center for Multiscale Theory and Computation, Westfälische Wilhelms University, 48149 Münster, Germany
[5]Center for Molecular Biology (ZMBH), DKFZ-ZMBH Alliance, Heidelberg University, 69120 Heidelberg, Germany
[6]Interdisciplinary Center for Scientific Computing (IWR), 69120 Heidelberg, Germany
*Correspondence: rebecca.wade@h-its.org
https://doi.org/10.1016/j.str.2018.05.009

There is renewed interest in linker histone (LH)—nucleosome binding and how LHs influence eukaryotic DNA compaction. For a long time, the goal was to uncover "the structure of the chromatosome," but recent studies of LH-nucleosome complexes have revealed an ensemble of structures. Notably, the reconstituted LH-nucleosome complexes used in experiments rarely correspond to the sequence combinations present in organisms. For a full understanding of the determinants of the distribution of the chromatosome structural ensemble, studies must include a complete description of the sequences and experimental conditions used, and be designed to enable systematic evaluation of sequence and environmental effects.

Introduction

In eukaryotes, DNA is wrapped around core histone protein oligomers to form chromatin (Kornberg, 1974). For cell function, it is crucial to dynamically compact the genetic material—which is about 2 m long in humans—in such a way that specific genes for transcription can be accessed when required (Taube and Barton, 2006). Despite more than 30 years of research, the mechanism of higher-order chromatin compaction is not fully resolved (van Holde and Zlatanova, 2007; Grigoryev and Woodcock, 2012). The repeating unit of chromatin, the nucleosome, is composed of a nucleosome core flanked by two linker DNA (L-DNA) arms. The nucleosome core consists of 145–147 bp of nucleosomal DNA (N-DNA) wrapped around a histone octamer composed of two copies of each of the core histone proteins H2A, H2B, H3, and H4 (Klug et al., 1980; Luger et al., 1997). In addition to the core histones, a linker histone (LH) protein, H1 or H5, can bind to the nucleosome between the two L-DNA arms to form a chromatosome (Pruss et al., 1996; Zhou et al., 2013, 2015; Flanagan et al., 2016). Chromatosomes were first revealed by digestion of chromatin by a non-specific nuclease to consist of the nucleosome core, about 20 bp of L-DNA and one LH (Simpson, 1978). The chromatosome can therefore be considered as a fundamental unit of the chromatin structure (Widom, 1998), and the determination of the three-dimensional structure of this subnucleosomal particle has been a longstanding goal.

LHs are composed of about 200 amino acid residues, and contain three distinct domains, a short (~40 residues) unstructured N-terminal tail, a relatively conserved globular domain (GD) (~80 residues), and a basic disordered C-terminal tail (~100 residues) (Roque et al., 2016). Previous studies have shown that, even though the N- and C-terminal tails can affect the affinity and geometry of LH-nucleosome binding (Hutchinson et al., 2015), they do not appear to affect the position of the LH GD (Syed et al., 2010; Zhou et al., 2016). Furthermore, both the GD and the full-length LH protect the same L-DNA from micrococcal nuclease digestion (Puigdomènech et al., 1983). Thus, the position of the LH on the chromatosome is mainly governed by the LH GD. For this reason, and because of the difficulties of studying disordered regions of proteins, most *in vitro* studies aimed at revealing the structure of the chromatosome have been focused on LH GD-nucleosome complexes.

Recent research has shown that LH proteins have a range of functions, including roles in DNA replication, epigenetic regulation, genome stability, and DNA repair (for a recent review see Fyodorov et al., 2017). Higher eukaryotes have a family of LH proteins consisting of a number of variants, also referred to as subtypes, which have a relatively conserved GD and more variable N and C termini (Kowalski and Pałyga, 2016). It has been shown that LH variants can have different functions, tissue expression levels, and DNA binding affinities (Millán-Ariño et al., 2016; Parseghian, 2015; Parseghian and Hamkalo, 2001). In mammals, there are seven standard H1 subtypes with varying sequence conservation, chromatin binding affinity, and genomic distribution (Kowalski and Pałyga, 2016). H1 LH proteins have been shown to be essential for mouse development (Pan and Fan, 2016). For example, even though a single H1 isoform knockout did not result in any significant phenotypic change, deletion of three isoforms was shown to be embryonically lethal (Drabent et al., 2000; Fan et al., 2001, 2003). On the other hand, studies in unicellular eukaryotes, such as *Aspergillus nidulans* and *Tetrahymena thermophila*, have indicated that knockout of the sole H1 isoform is not lethal but can cause some genes to be up- or downregulated (Ramón et al., 2000;

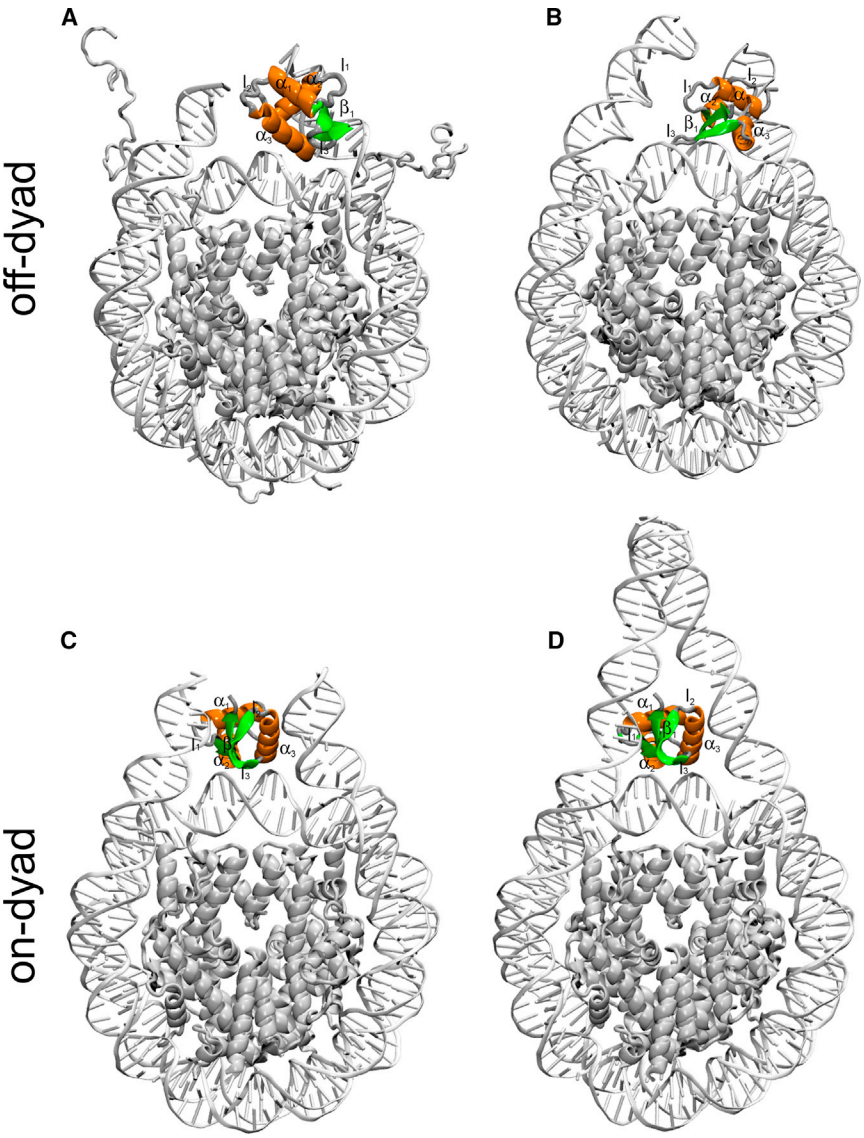

Figure 1. Four Recently Determined Three-Dimensional Structures of LH-Nucleosome Complexes

(A) Off-dyad binding of *D. melanogaster* H1 GD to a nucleosome, as reported by Zhou et al. (2013) (structure kindly provided by Yawen Bai).

(B) Off-dyad binding of *H. sapiens* H1.4 GD to a nucleosome as reported by Song et al. (2014) (structure kindly provided by Ping Zhu).

(C) On-dyad binding of *G. gallus* H5 GD to a nucleosome as reported by Zhou et al. (2015), PDB: 4QLC.

(D) On-dyad binding of *X. laevis* H1 GD to a nucleosome as reported by Bednar et al. (2017), PDB: 5NL0.

LH proteins are shown in cartoon representation and colored and labeled according to secondary structure elements: α helices in orange, β sheets in green and unstructured regions (including loops, *l*) in gray. DNA is shown in light gray and core histones are shown in dark gray. See Table 1 and the text for further details.

Shen and Gorovsky, 1996). Furthermore, it was previously reported that LHs behave as regulators of specific genes by affecting nucleosome spacing (Fan et al., 2003).

Even with many biological and physiological roles being associated with the LH, its function remains enigmatic. However, recent breakthroughs in the determination of LH-nucleosome structures, coupled with a growing number of epigenetic studies, open up the possibility of achieving a thorough understanding of the mechanism of formation of LH-nucleosome complexes (for a recent review see Cutter and Hayes, 2017). Here, we compare the five recently determined three-dimensional structures of LH-nucleosome complexes (Bednar et al., 2017; Song et al., 2014; Zhou et al., 2013, 2015, 2016) (see Figure 1), and consider what can be learnt from these experimentally determined structures, as well as from modeling and simulation studies. Our analysis suggests that the different structures of LH-nucleosome complexes revealed in these studies can be reconciled by a paradigm shift away from the concept of "the structure of the chromatosome" toward "the structural ensemble distributions of individual chromatosomes," in which alternative configurations of LH-nucleosome complex structures can exist. In particular, these configurations differ in the position and orientation of the LH GD with respect to the nucleosome. Our analysis also highlights the importance of comprehensive documentation of protein and DNA sequences and post-translational modifications (PTMs) in future studies of LH-nucleosome complexes.

What Are the LH-Nucleosome Systems for Which Structures Have Been Determined?

Reconstitution of nucleosomes requires certain conditions that are far from physiological conditions, such as 2 M salt concentration, as well as suitable DNA and protein sequences (Luger et al., 1999). Obtaining chromatosomes in a form suitable for structure determination has been difficult. As can be seen in Table 1, the systems studied have a combination of DNA, core histone, and LH sequences of different origins and DNA and protein constructs of different lengths. Moreover, the LH-nucleosome complexes were reconstituted and their structures determined under a range of environmental conditions, with different LH-nucleosome ratios, with different buffers and at different pH values and temperatures. Bednar et al. (2017) even used a chaperone, the NAP-1 histone chaperone, for reconstitution of LH-nucleosome complexes. Notably, the nucleosomes were reconstituted using salt dialysis against a gradually decreasing high salt buffer, the LH-nucleosome complexes were reconstituted by incubation at various ionic strength conditions, and the structural measurements were made at salt concentrations ranging from about 10 mM up to close to physiological ionic strength (Table 1). On the other hand, Schlick and colleagues showed that salt and

Please cite this article in press as: Öztürk et al., Toward an Ensemble View of Chromatosome Structure: A Paradigm Shift from One to Many, Structure (2018), https://doi.org/10.1016/j.str.2018.05.009

Table 1. Experimentally Determined Structures of LH-Nucleosome Complexes

References	Zhou et al. (2013)	Song et al. (2014)	Zhou et al. (2015)	Zhou et al. (2016)	Bednar et al. (2017)
Experimental Details					
Structure determination methods	NMR, PRE ITC, HADDOCK	cryo-EM ultracentrifugation	NMR, ITC, X-ray FRET, ultracentrifugation	NMR, ITC HADDOCK ultracentrifugation	cryo-EM, X-ray OH footprint, CL
147-bp N-DNA	synthetic DNA Widom 601	synthetic DNA Widom 601	synthetic DNA Widom 601	synthetic DNA Widom 601	synthetic DNA Widom 601 Widom 601 L[a]
No. of L-DNA bp[b]	10 + 10 30 + 30	15 + 15 20 + 20	10 + 10	10 + 10 30 + 30	25 + 25
Core histones	*D. melanogaster*	*X. laevis*	*D. melanogaster*	*D. melanogaster*	*H. sapiens*
Linker histone (LH)[c]	*D. melanogaster* H1 (WT, 37-132, 45–119, 37–211, 37–256)	*H. sapiens* H1.4	*G. gallus* H5 (22–98, 24–98, 22–102 and 22–142) *D. melanogaster* (WT and 44–118), *X. laevis* H1	WT and mutant *G. gallus* H5 (24–98) *D. melanogaster* H1 (WT and 45–119) *X. laevis* H1.0 *H. sapiens* H1.0	*X. laevis* H1.0b *H. sapiens* H1.5 (1–177 and 40–112)
Environmental conditions for LH-nucleosome structural measurements[d]	low IS, pH 6.0–7.4	low IS, pH 8.0	NMR: low IS, X-ray, ITC FRET: medium IS, pH 3.75–8.0	NMR: low IS ITC: medium IS, pH 7.4–8.0	Cryo-EM: low IS X-ray: medium IS, pH 6.4
Resolution (Å)	–	11 and 25	3.5	–	5.5 (X-ray)
Structure Details					
Basis for nucleosome structure	nucleosome from PDB: 1ZBB and 1KX5	cryo-EM map fitted with nucleosome PDB: 1AOI and 1ZBB	electron density fitted with nucleosome PDB: 4INM and 3MVD	DNA from PDB: 4QLC	electron density fitted with nucleosome PDB: 3UT9
N-DNA	*H. sapiens* X chromosome α satellite DNA palindromic 147 bp	*H. sapiens* X chromosome α satellite DNA palindromic 146 bp	synthetic DNA Widom 601 147 bp	synthetic DNA Widom 601 147 bp	synthetic DNA Widom 601 L[a] 145 bp
No. of L-DNA bp[b]	10 + 10	15 + 15 20 + 20	10 + 10	0 + 0	26 + 26
Core histones	*X. laevis*	*X. laevis*	*D. melanogaster*	none	*X. laevis*
CH tails	yes	no	no	no	no
Modeled LH sequence	*D. melanogaster* H1	*G. gallus* H5	*G. gallus* H5	*G. gallus* H5, *D. melanogaster* H1	*X. laevis* H1.0b
Modeled LH structure	from closed *G. gallus* LH PDB: 1HST, chain B	from open *G. gallus* LH PDB: 1HST, chain A	from closed *G. gallus* LH PDB: 1HST, chain B	H5, from closed *G. gallus* LH PDB: 1HST, chain B H1, from closed *G. gallus* LH PDB: 1HST, chain B	from closed *G. gallus* LH PDB: 1HST, chain B
LH position	off-dyad	off-dyad	on-dyad	on-dyad off-dyad	on-dyad
PDB ID of model			4QLC		5NL0

The methods used and the sequences studied are given, followed by the details of the structural models derived from the experimental results. CH, core histone; N-DNA, nucleosomal DNA; L-DNA, linker DNA; NMR, nuclear magnetic resonance; PRE, paramagnetic relaxation enhancement; ITC, isothermal titration calorimetry; HADDOCK, High Ambiguity-Driven protein-protein DOCKing; cryo-EM, cryoelectron microscopy; X-ray, X-ray crystallography; FRET, Förster resonance energy transfer; OH footprint, hydroxyl radical footprinting; CL, chemical crosslinking; WT, wild-type.

[a]The Widom 601L N-DNA sequence is the palindrome of the left half of the Widom 601 N-DNA sequence.

[b]The number of bp for each L-DNA arm is given, e.g., 10 + 10 denotes L-DNA1 with 10 bp and L-DNA2 with 10 bp.

[c]Residue ranges are given in parentheses.

[d]The ionic strength is classified as low: ca. 10–20 mM, and medium: ca. 100–150 mM.

LH concentration, L-DNA length, the presence of oligo-nucleosome systems, and synergistic folding of the LH C-terminal domain affect chromatin condensation and LH contacts with L-DNAs (Luque et al., 2014, 2016; Perišić and Schlick, 2017). Thus, the heterogeneity of the studied systems should be borne in mind in considering the relevance of results with these *in vitro* systems for understanding chromatosome and chromatin structure in cell nuclei.

The nucleosome systems vary in the lengths of the L-DNA arms, which each range from 10 to 30 bp. The sequences of the N-DNA and L-DNA can influence the binding location and the orientation of the LH, which may interact directly with one or both of the L-DNA arms (Öztürk et al., 2016). The first chromatosome structure solved (Zhou et al., 2013) had a Widom 601 DNA sequence and core histone proteins from *Drosophila melanogaster*. A common component of the recent structural nucleosome studies is the synthetic 147-bp Widom 601 N-DNA sequence that wraps around the core histones and has a strong core histone octamer binding affinity (Lowary and Widom, 1998). The choice of Widom 601 sequence, albeit unnatural, allowed researchers to obtain more stable nucleosomes (Tóth et al., 2013), thereby facilitating structure determination. To the best of our knowledge, the first published report of the sequence of Widom 601 DNA was given in the study of Schalch et al. (2005). A palindromic variant, Widom 601 L, with higher core histone octamer affinity (L indicates that it was generated from the left half of the Widom 601 sequence), was also used by (Bednar et al., 2017; Chua et al., 2012).

In the published studies of the structures of LH-nucleosome complexes, the core histones vary in origin, as shown in Table 1. The core histones have flexible tails, which are present in the sequences used in the experiments but often missing in the final structures determined. The extent to which the flexible tails affect LH binding is unknown. Öztürk et al. (2016) found that off-dyad binding would mean little interaction of the *Gallus gallus* gH5 with core histone tails. On the other hand, Zhou et al. (2013) reported that *D. melanogaster* H1 methyl groups are affected by paramagnetic relaxation enhancement (PRE) labeling of T119 in the H2A tail, and that the disordered C-terminal tail of H2A folds upon LH binding. These results suggest that further research is required to understand the effects of the core histone tails on LH binding to the nucleosome. Experimentally, the LHs have been studied as full-length proteins and as GD constructs of varying lengths and, in some cases, with mutations to improve stability or switch key isoform residues. The N- and C-terminal domains are highly flexible and, therefore, their removal can be expected to facilitate crystallization.

How Have the Structural Models of LH-Nucleosome Complexes Been Derived?

The experimental methods used vary in the level of detail and the amount of information that they provide, as well as the associated uncertainties (for a recent review, see Mackay et al., 2017). For the first structure of an LH-nucleosome complex determined, Zhou et al. (2013) mutated four residues of the *D. melanogaster* H1 GD and obtained a more stable LH domain, similar to the *G. gallus* H5 GD. By using a gel shift assay and isothermal titration calorimetry (ITC), they showed that various mutant *D. melanogaster* H1 constructs (residues 37–132, 45–119, 37–211, and 37–256) have the same nucleosome binding affinities. The authors derived experimental constraints with nuclear magnetic resonance (NMR) shifts and PRE for wild-type and mutant *D. melanogaster* H1 binding to a nucleosome. The structure of the *D. melanogaster* H1 GD was modeled by homology, based on the closed conformation of the *G. gallus* H5 GD in the crystal structure (PDB: 1HST, chain B) and a structural model of an LH GD-nucleosome complex was obtained by docking the LH GD and nucleosome with the HADDOCK program (Dominguez et al., 2003) using a small number of restraints derived from the combined experimental results. It should be noted that, even though Zhou et al. (2013) did their experiments with a Widom 601 N-DNA sequence, in their docking calculations they used the nucleosome structure with PDB: 1ZBB, whose DNA sequence is not Widom 601 but a palindromic sequence extracted from PDB: 1KX5 (Schalch et al., 2005). Later, Zhou et al. (2016) used a similar approach to study the binding of LHs from two different organisms.

A detailed model was only obtained when the first structure of an LH GD-nucleosome complex was determined by X-ray crystallography at 3.5 Å resolution (PDB: 4QLC) (Zhou et al., 2015). This model was supported by NMR data in the same publication. The first cryoelectron microscopy (cryo-EM) study (Song et al., 2014) gave structures of chromatin fibers with 12 × 177- and 12 × 187-bp nucleosomes in the presence of full-length and wild-type *Homo sapiens* H1.4 with ~25 and 11 Å resolution, respectively. Both structures were in agreement with a zigzag two-start helix model for the 30 nm chromatin fiber. By averaging the densities of the central four nucleosomes in the 11 Å resolution map, Song et al. (2014) deduced an off-dyad binding mode for H1. Although they proposed a specific orientation of H1 in the chromatosome, the low resolution of the electron density map means that other orientations, are also consistent with the data. In our fitting of the LH GD into the chromatin fiber EM density, we found that it is also possible to obtain various off-dyad LH configurations (M.A.Ö., unpublished data). This may reflect the variable and dynamic nature of chromatin due to molecular flexibility and variable DNA length, histone variants, and PTMs of the core histones and DNA.

Recently, Bednar et al. (2017) reported the first X-ray crystal structure for a complex containing a nucleosome with a full-length LH at 5.5 Å resolution (PDB: 5NL0). In addition, they applied cryo-EM, site-directed crosslinking, and hydroxyl radical footprinting methods in the same study. For experiments, they used standard Widom 601 and palindromic Widom 601L L-DNA sequences, together with *H. sapiens* core histone, and either *H. sapiens* H1.5 or *Xenopus laevis* H1.0b LH proteins. For deriving structural models, they used *X. laevis* core histone and LH protein sequences.

What Is the Position of the LH with Respect to the Nucleosome?

Biochemical experiments performed by micrococcal nuclease digestion (Allan et al., 1980; An et al., 1998), chemical crosslinking (Pruss et al., 1996; Zhou et al., 1998), fluorescence recovery after photobleaching (FRAP) (Brown et al., 2006), and hydroxyl radical footprinting (Syed et al., 2010) have previously indicated either on- or off-dyad binding of LH proteins to nucleosomes. In addition, molecular modeling and simulation studies resulted in various on- and off-dyad LH binding modes (Bharath et al., 2003; Brown et al.,

2006; Cui and Zhurkin, 2009; Fan and Roberts, 2006; Öztürk et al., 2016; Pachov et al., 2011). The recent structure determinations by NMR, X-ray crystallography, and cryo-EM reported by Zhou et al. (2013, 2015, 2016), Song et al. (2014), and Bednar et al. (2017), show both on- and off-dyad binding modes for the LH (see Figure 1 and Table 1).

In the off-dyad configuration, the LH GD interacts with only one of the L-DNAs and binds to the N-DNA adjacent to the dyad axis. Zhou et al. (2013) showed that *D. melanogaster* H1, both as a full-length wild-type construct (residues 1–256) and in a truncated form (residues 37–211), binds off-dyad to a nucleosome with a 147-bp Widom 601 DNA sequence, two 10-bp L-DNAs and *D. melanogaster* core histones. Song et al. (2014) showed that full-length wild-type *H. sapiens* H1.4 (residues 1–219) binds off-dyad to a chromatin composed of nucleosomes of 147 bp Widom 601 DNA sequence and two 15- or 20-bp L-DNAs wound around *X. laevis* core histones. However, it should be noted that the authors crosslinked LHs to the nucleosomes which may cause artifacts. Zhou et al. (2016) reported that both full-length and truncated (residues 45–119) *D. melanogaster* H1 bind in an off-dyad position to a nucleosome of a 147-bp Widom 601 DNA sequence and two 10- or 30-bp L-DNAs with *D. melanogaster* core histones.

In the on-dyad configuration, the LH interacts with both L-DNAs and the N-DNA on the dyad axis. Zhou et al. (2015) found that *G. gallus* H5 (residues 22–98, 22–102, and 22–142) binds on-dyad to a nucleosome of a 147-bp Widom 601 DNA sequence and two 10-bp L-DNAs with *D. melanogaster* core histones. Most recently, Bednar et al. (2017) reported that full-length *X. laevis* H1, *H. sapiens* H1.5 (residues 1–177, lacking the 50 C-terminal residues), and *H. sapiens* H1.5 GD (residues 40–112) bind with the GD on-dyad to nucleosomes with a 147-bp Widom 601 DNA sequence and two 25-bp L-DNAs with *H. sapiens* core histones.

What Is the Orientation of the LH-Nucleosome Complexes?

The experimentally derived structures not only show two positions of the LH on the nucleosome—on- and off-dyad—but also show different orientations, with the l_3 loop of the LH interacting with either the minor groove on the dyad or the L-DNA and N-DNA major grooves (see Figure 1). In several cases, the experimental data can be fit with more than one orientation of the LH GD, i.e., its orientation cannot be unambiguously defined from the experimental results. Computational docking can help to identify the preferred orientation, for example, as applied by Zhou et al. (2015, 2016). However, in such efforts, the sequences of the modeled structures should ideally be exactly the same as the sequences used experimentally, which has not always been the case (Table 1). Computer simulations can also provide insights into the mechanism of association. Brownian and molecular dynamics simulations of the association of the H5 GD to a nucleosome showed that off-dyad LH-nucleosome binding involves conformational selection and induced fit mechanisms. The position and orientation of the H5 GD was dependent on the opening of the L-DNA arms, and the H5 GD was observed to undergo induced fit toward an open conformation to optimize LH l_3 loop-DNA interactions and the off-dyad orientation on the nucleosome (Öztürk et al., 2016).

Consideration of the positions occupied by the flexible LH N- and C-termini also serves to limit the possible orientations an LH can adopt on a nucleosome. For example, the LH configuration proposed by Zhou et al. (2013) has a very close contact between the C-terminus of the H1 and L-DNA, which may not represent the full-length LH system *in vivo*. Recently, the cryo-EM structure of Bednar et al. (2017) revealed that the C-terminal domain of the LH localizes on one of the L-DNAs and introduces an asymmetry in the structure of the nucleosome, which is consistent with their crystal structure, and site-directed crosslinking data showing closer association of the LH GD to one L-DNA than the other, despite binding of the LH GD in an on-dyad position.

Zhou et al. (2013, 2015) indicated that single-residue mutations in the LH GD can significantly affect the LH-nucleosome binding affinity. Further experimental and computational analysis is necessary with mutant LHs to understand whether point mutations lead to a positional or an orientational shift of the LH GD with respect to the nucleosome. This aspect is important as, generally, experiments on LH-nucleosome complexes are carried out with mutant forms of LH and care is therefore required in interpretation of the data with respect to wild-type or post-translationally modified LHs.

What Else Have We Learnt About the Structures of LH-Nucleosome Complexes?

Zhou et al. (2013) provided the first systematic approach to combining various experimental methods to determine the structures of LH-nucleosome complexes. Apart from using various lengths and mutants of *D. melanogaster* LH, they showed that the construction of nucleosomes with H2A.Z core histones resulted in an undetectable level of LH binding as measured by ITC. This suggests that, depending on the composition of the core histones of the nucleosome, there could be various LH-nucleosome binding affinities and different chromatosome ensembles. Song et al. (2014) reported that the tetranucleosomal units of the 12mer structure have an interaction of the N terminus of H4 and the acidic patch of the H2A-H2B dimer, which was suggested to be the reason for the twist between the tetranucleosomal units. Such a twist would allow a wide range of L-DNA angles to be present in higher-order nucleosome structures. Recently, Öztürk et al. (2016) showed that chromatosome configuration is dependent on the L-DNA opening angle and sequence composition. Multi-nucleosome unit twists in the higher-order chromatin structure could allow various chromatosome ensembles, as also shown by mesoscale simulations of chromatin (Perišić and Schlick, 2017).

Zhou et al. (2015) reported the first crystal structure of nucleosome bound to *G. gallus* H5 GD. The clear observation of the side chains of four arginine residues of the H5 GD interacting with DNA implied that the H5 GD makes stable interactions with the dyad N-DNA and both L-DNA arms. This suggests that specific residues are responsible for the affinity to the nucleosome and the stability of the structure of the complex. Similarly, the authors showed that the H5 GD undergoes conformational rearrangement upon nucleosome binding, and they reported that the H5 GD l_2 loop is more stable than the l_3 loop in its free form. Such conformational rearrangement of the LH upon nucleosome binding was also reported by Öztürk et al. (2016) based on extensive molecular dynamics simulations. Computational

techniques can give further insights into the molecular features of the chromatosome due to the ability to construct various LH-nucleosome systems *in silico* without experimental constraints, such as those for achieving protein expression.

Bednar et al. (2017) found, by cryo-EM of unbound and bound nucleosomes, that binding of the LH leads to a more closed nucleosome structure with closer and less-flexible L-DNA arms. Consistently, Zhou et al. (2015) found by crystallography that, relative to the LH-bound state, the L-DNA arms of the free nucleosome are ~10 Å further apart, which would affect the higher-order chromatin structure and dynamics. Furthermore, in the same publication, the authors applied sedimentation assays on 12 × 177-bp nucleosomes with *D. melanogaster* full-length H1 and H1 GD, *X. laevis* H1, and *G. gallus* H5 GD. They found that, in nucleosome arrays, H5 GD has a ~6S higher sedimentation coefficient compared with the *D. melanogaster* H1 GD, and this could be an indication of different nucleosome complexation mechanisms for the respective LHs. The authors also mention that NMR analysis showed that the human LH subtype, H1x, has an α_1 helix that is two helical turns longer than that of the H5 GD. This difference in length could lead to a distinct nucleosome binding mode.

Lastly, Bednar et al. (2017) showed that LH tails introduce an asymmetry into higher-order chromatin structure as the C-terminal tail of the LH only interacts with one of the L-DNAs in the nucleosome. This feature could affect the accessibility of the nucleosomes for LH binding in the chromatin structure, and could facilitate or block LH binding to certain conformations. In addition to cryo-EM and crystal structure determination, the authors conducted hydroxyl radical footprinting analysis and site-specific crosslinking experiments in solution. They showed that both *X. laevis* full-length and GD H1 have similar symmetric footprints on the core DNA, indicating that the LH tails have a limited effect on the LH binding site. To distinguish on- and off-dyad binding, the authors also conducted hydroxyl radical footprinting experiments on nucleosomes that lack either one or the other L-DNA, and observed similar DNA protection patterns, as observed for the full nucleosome particle, supporting an on-dyad binding mode for GD H1. Interestingly, site-specific crosslinking experiments in solution with full-length H1 were consistent with the two on-dyad orientations of GD H1 observed in their crystal structure. Finally, the authors pointed out that the PTMs of the LH could change the electrostatic potential of the protein and result in regulation of chromatin structure.

These recent studies show that different experimental approaches can be combined to understand the structure and dynamics of the chromatosome. Taken together, and contrary to the common concept of a single unique chromatosome structure, experiments and computations suggest that LHs can recognize nucleosomes in alternative configurations depending on LH protein and nucleosome DNA sequences as well as PTMs. As a result, chromatosomes can be considered to be composed of varying distributions of structural ensembles.

Suggestions for Future Structural Studies of LH-Nucleosome Complexes

To facilitate investigations to achieve an in-depth understanding of chromatosome structure and dynamics, we would like to emphasize the importance of comprehensive reporting of the details of the sequences and environmental conditions of the systems studied.

The DNA sequences used in nucleosome construction, the labeling of the L-DNAs, the source organism of the core histones, and the presence or absence of core histone tails, are some of the nucleosome descriptors that need to be clearly specified. The LH constructs used in experiments, the length of the LH proteins, and the presence of mutations need to be given. This is important for reproducibility and comparison of the studied systems both experimentally and computationally.

In some of the recent studies, the experimental system investigated and the structural model derived do not match each other in sequence (see Table 1). This can lead to inconsistencies and issues with data interpretation. Such problems can be ameliorated if the coordinates of models are deposited and detailed protocols for the computational approaches used to derive the models provided.

As well as providing further LH-nucleosome structures, future studies should address binding mechanisms. For example, they could test the possibility of stepwise assembly of chromatosomes in which the off-dyad configuration of the LH GD may represent an intermediate binding mode with shorter life span compared with the on-dyad configuration of the LH GD. For this purpose, complementary structural, thermodynamic, and kinetic experiments will be necessary.

For a long time, the question was "What is *the* structure of *the* chromatosome?" and a range of modeling approaches were used to try to answer it. The recent experimental structure determinations reviewed here suggest that the question should be reformulated as "What are *the* structural ensembles of chromatosomes and how are they dependent on sequences, post-translational modifications and environmental conditions?" Considering recent experimental and computational advances, we can expect many interesting answers that will contribute toward demistifying the enigmatic nature of chromatosomes.

ACKNOWLEDGMENTS

The authors thank Dr. Katalin Toth, Dr. Neil Bruce, Ina Pöhner, and Madhura De for their helpful comments on the manuscript and Dr. Yawen Bai and Prof. Ping Zhu for providing their structures. M.A.Ö. and R.C.W. thank the Klaus Tschira Foundation and HITS gGmbH for funding, and acknowledge support by the state of Baden-Württemberg through bwHPC for computing resources. V.C. thanks the Max Planck Society, and in particular Hans Schöler for support.

AUTHOR CONTRIBUTIONS

M.A.Ö. and R.C.W. conceived, designed, and wrote the manuscript with contributions from V.C.

REFERENCES

Allan, J., Hartman, P.G., Crane-Robinson, C., and Aviles, F.X. (1980). The structure of histone H1 and its location in chromatin. Nature *288*, 675–679.

An, W., Leuba, S.H., van Holde, K., and Zlatanova, J. (1998). Linker histone protects linker DNA on only one side of the core particle and in a sequence-dependent manner. Proc. Natl. Acad. Sci. USA *95*, 3396–3401.

Bednar, J., Garcia-Saez, I., Boopathi, R., Cutter, A.R., Papai, G., Reymer, A., Syed, S.H., Lone, I.N., Tonchev, O., Crucifix, C., et al. (2017). Structure and dynamics of a 197 bp nucleosome in complex with linker histone H1. Mol. Cell *66*, 384–397.e8.

Bharath, M.M.S., Chandra, N.R., and Rao, M.R.S. (2003). Molecular modeling of the chromatosome particle. Nucleic Acids Res. *31*, 4264–4274.

Brown, D.T., Izard, T., and Misteli, T. (2006). Mapping the interaction surface of linker histone H10 with the nucleosome of native chromatin in vivo. Nat. Struct. Mol. Biol. *13*, 250–255.

Chua, E.Y.D., Vasudevan, D., Davey, G.E., Wu, B., and Davey, C.A. (2012). The mechanics behind DNA sequence-dependent properties of the nucleosome. Nucleic Acids Res. *40*, 6338–6352.

Cui, F., and Zhurkin, V.B. (2009). Distinctive sequence patterns in metazoan and yeast nucleosomes: implications for linker histone binding to AT-rich and methylated DNA. Nucleic Acids Res. *37*, 2818–2829.

Cutter, A.R., and Hayes, J.J. (2017). Linker histones: novel insights into structure-specific recognition of the nucleosome. Biochem. Cell Biol. *95*, 171–178.

Dominguez, C., Boelens, R., and Bonvin, A.M.J.J. (2003). HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J. Am. Chem. Soc. *125*, 1731–1737.

Drabent, B., Saftig, P., Bode, C., and Doenecke, D. (2000). Spermatogenesis proceeds normally in mice without linker histone H1t. Histochem. Cell Biol. *113*, 433–442.

Fan, L., and Roberts, V.A. (2006). Complex of linker histone H5 with the nucleosome and its implications for chromatin packing. Proc. Natl. Acad. Sci. USA *103*, 8384–8389.

Fan, Y., Sirotkin, A., Russell, R.G., Ayala, J., and Skoultchi, A.I. (2001). Individual somatic H1 subtypes are dispensable for mouse development even in mice lacking the H1(0) replacement subtype. Mol. Cell. Biol. *21*, 7933–7943.

Fan, Y., Nikitina, T., Morin-Kensicki, E.M., Zhao, J., Magnuson, T.R., Woodcock, C.L., and Skoultchi, A.I. (2003). H1 linker histones are essential for mouse development and affect nucleosome spacing in vivo. Mol. Cell. Biol. *23*, 4559–4572.

Flanagan, T.W., Files, J.K., Casano, K.R., George, E.M., and Brown, D.T. (2016). Photobleaching studies reveal that a single amino acid polymorphism is responsible for the differential binding affinities of linker histone subtypes H1.1 and H1.5. Biol. Open *5*, 016733.

Fyodorov, D.V., Zhou, B.-R., Skoultchi, A.I., and Bai, Y. (2017). Emerging roles of linker histones in regulating chromatin structure and function. Nat. Rev. Mol. Cell Biol. *19*, 192–206.

Grigoryev, S.A., and Woodcock, C.L. (2012). Chromatin organization — the 30nm fiber. Exp. Cell Res. *318*, 1448–1455.

van Holde, K., and Zlatanova, J. (2007). Chromatin fiber structure: where is the problem now? Semin. Cell Dev. Biol. *18*, 651–658.

Hutchinson, J.B., Cheema, M.S., Wang, J., Missiaen, K., Finn, R., Gonzalez Romero, R., Th'ng, J.P.H., Hendzel, M., and Ausió, J. (2015). Interaction of chromatin with a histone H1 containing swapped N- and C-terminal domains. Biosci. Rep. *35*, e00209.

Klug, A., Rhodes, D., Smith, J., Finch, J.T., and Thomas, J.O. (1980). A low resolution structure for the histone core of the nucleosome. Nature *287*, 509–516.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868–871.

Kowalski, A., and Pałyga, J. (2016). Modulation of chromatin function through linker histone H1 variants. Biol. Cell *108*, 339–356.

Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J. Mol. Biol. *276*, 19–42.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251–260.

Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Expression and purification of recombinant histones and nucleosome reconstitution. Methods Mol. Biol. *119*, 1–16.

Luque, A., Collepardo-Guevara, R., Grigoryev, S., and Schlick, T. (2014). Dynamic condensation of linker histone C-terminal domain regulates chromatin structure. Nucleic Acids Res. *42*, 7553–7560.

Luque, A., Ozer, G., and Schlick, T. (2016). Correlation among DNA linker length, linker histone concentration, and histone tails in chromatin. Biophys. J. *110*, 2309–2319.

Mackay, J.P., Landsberg, M.J., Whitten, A.E., and Bond, C.S. (2017). Whaddaya know: a guide to uncertainty and subjectivity in structural biology. Trends Biochem. Sci. *42*, 155–167.

Millán-Ariño, L., Izquierdo-Bouldstridge, A., and Jordan, A. (2016). Specificities and genomic distribution of somatic mammalian histone H1 subtypes. Biochim. Biophys. Acta *1859*, 510–519.

Öztürk, M.A., Pachov, G.V., Wade, R.C., and Cojocaru, V. (2016). Conformational selection and dynamic adaptation upon linker histone binding to the nucleosome. Nucleic Acids Res. *44*, 6599–6613.

Pachov, G.V., Gabdoulline, R.R., and Wade, R.C. (2011). On the structure and dynamics of the complex of the nucleosome and the linker histone. Nucleic Acids Res. *39*, 5255–5263.

Pan, C., and Fan, Y. (2016). Role of H1 linker histones in mammalian development and stem cell differentiation. Biochim. Biophys. Acta *1859*, 496–509.

Parseghian, M.H. (2015). What is the role of histone H1 heterogeneity? A functional model emerges from a 50 year mystery. Biophys *2*, 724–772.

Parseghian, M.H., and Hamkalo, B.A. (2001). A compendium of the histone H1 family of somatic subtypes: an elusive cast of characters and their characteristics. Biochem. Cell Biol. *79*, 289–304.

Perišić, O., and Schlick, T. (2017). Dependence of the linker histone and chromatin condensation on the nucleosome environment. J. Phys. Chem. B *121*, 7823–7832.

Pruss, D., Bartholomew, B., Persinger, J., Hayes, J., Arents, G., Moudrianakis, E.N., and Wolffe, A.P. (1996). An asymmetric model for the nucleosome: a binding site for linker histones inside the DNA gyres. Science *274*, 614–617.

Puigdomènech, P., José, M., Ruiz-Carrillo, A., and Crane-Robinson, C. (1983). Isolation of a 167 basepair chromatosome containing a partially digested histone H5. FEBS Lett. *154*, 151–155.

Ramón, A., Muro-Pastor, M.I., Scazzocchio, C., and Gonzalez, R. (2000). Deletion of the unique gene encoding a typical histone H1 has no apparent phenotype in *Aspergillus nidulans*. Mol. Microbiol. *35*, 223–233.

Roque, A., Ponte, I., and Suau, P. (2016). Interplay between histone H1 structure and function. Biochim. Biophys. Acta *1859*, 444–454.

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. Nature *436*, 138–141.

Shen, X., and Gorovsky, M.A. (1996). Linker histone H1 regulates specific gene expression but not global transcription in vivo. Cell *86*, 475–483.

Simpson, R.T. (1978). Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. Biochemistry *17*, 5524–5531.

Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.-M., Zhu, P., and Li, G. (2014). Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. Science *344*, 376–380.

Syed, S.H., Goutte-Gattat, D., Becker, N., Meyer, S., Shukla, M.S., Hayes, J.J., Everaers, R., Angelov, D., Bednar, J., and Dimitrov, S. (2010). Single-base resolution mapping of H1-nucleosome interactions and 3D organization of the nucleosome. Proc. Natl. Acad. Sci. USA *107*, 9620–9625.

Taube, J.H., and Barton, M.C. (2006). Chromatin and regulation of gene expression. In Gene Expression and Regulation, J. Ma, ed. (Springer), pp. 95–109.

Tóth, K., Böhm, V., Sellmann, C., Danner, M., Hanne, J., Berg, M., Barz, I., Gansen, A., and Langowski, J. (2013). Histone- and DNA sequence-dependent stability of nucleosomes studied by single-pair FRET. Cytometry A *83*, 839–846.

Widom, J. (1998). Chromatin structure: linking structure to function with histone H1. Curr. Biol. *8*, R788–R791.

Zhou, B.-R., Feng, H., Kato, H., Dai, L., Yang, Y., Zhou, Y., and Bai, Y. (2013). Structural insights into the histone H1-nucleosome complex. Proc. Natl. Acad. Sci. USA *110*, 19390–19395.

Zhou, B.-R., Jiang, J., Feng, H., Ghirlando, R., Xiao, T.S., and Bai, Y. (2015). Structural mechanisms of nucleosome recognition by linker histones. Mol. Cell *59*, 628–638.

Zhou, B.-R., Feng, H., Ghirlando, R., Li, S., Schwieters, C.D., and Bai, Y. (2016). A small number of residues can determine if linker histones are bound on or off dyad in the chromatosome. J. Mol. Biol. *428*, 3948–3959.

Zhou, Y.-B., Gerchman, S.E., Ramakrishnan, V., Travers, A., and Muyldermans, S. (1998). Position and orientation of the globular domain of linker histone H5 on the nucleosome. Nature *395*, 402–405.