Comparison between the CENP-A and histone H3 structures in nucleosomes

Hiroaki Tachiwana, Wataru Kagawa and Hitoshi Kurumizaka*
Laboratory of Structural Biology; Graduate School of Advanced Science and Engineering; Waseda University;Tokyo, Japan

Keywords: CENP-A, nucleosome, centromere, chromatin, histone, histone H3, CENP-B, CENP-C
Submitted: 09/15/11
Revised: 10/10/11
Accepted: 10/11/11
http://dx.doi.org/10.4161/nucl.3.1.18372
*Correspondence to: Hitoshi Kurumizaka; Email: kurumizaka@waseda.jp

Extra View to: Tachiwana H, Kagawa W, Shiga T, Osakabe A, Miya Y, Saito K, et al. Crystal structure of the human centromeric nucleosome containing CENP-A. Nature 2011; 476:232–5; PMID:21743476; http://dx.doi.org/10.1038/nature10258

Centromeres are epigenetically marked by the assembly of nucleosomes containing the centromere-specific histone H3 variant, CENP-A (CENP-A nucleosome) and their inheritance is probably dictated by the architecture of the centromeric nucleosome. We previously determined the crystal structure of the human CENP-A nucleosome. CENP-A forms a histone octamer containing two each of histones H2A, H2B, H4 and CENP-A and the DNA is left-handedly wrapped around the histone octamer, as in canonical nucleosomes containing histone H3. In the CENP-A nucleosome structure, 13 base pairs of the DNA ends are detached from the histone surface and two CENP-A regions, the αN helix and loop 1, adopt different structures from those in the H3 nucleosome. In this Extra View article, we provide a detailed structural comparison between CENP-A and H3 in nucleosomes and describe their distinctions and similarities.

Introduction

In eukaryotes, chromosome segregation is accurately ensured during mitosis and meiosis. Kinetochores are unique chromosomal sites for microtubule attachment and are assembled on a specific chromosomal region called the centromere. The centromere is epigenetically marked by the centromere-specific histone H3 variant, CENP-A. Therefore, CENP-A is considered to form a unique chromatin architecture in centromeres. The fundamental structural unit of chromatin is the nucleosome, which contains two each of histones H2A, H2B, H3 and H4. Since CENP-A is an H3 variant, the conventional H3 is replaced by CENP-A in the centromeric nucleosome. Three models have been proposed for the centromeric nucleosome structure. (1) The octasome model, in which two each of histones H2A, H2B, H4 and CENP-A form a histone octamer, as in the canonical H3 nucleosome and the DNA is left-handedly wrapped around the histone octamer. (2) The compact octasome model, in which the DNA is left-handedly wrapped around the histone octamer, which forms a more compact structure than that of the canonical histone octamer. (3) The hemisome model, in which a heterotypic histone tetramer, containing one each of histones H2A, H2B, H4 and CENP-A, right-handedly wraps the DNA.

We prepared the CENP-A nucleosome, using bacterially-expressed human histones H2A, H2B, H4 and CENP-A, and determined its crystal structure. The crystal structure revealed that the histone octamer is formed with two each of histones H2A, H2B, H4 and CENP-A and the DNA is wrapped around it in a left-handed supercoil. The shape and size of the histone octamer of the CENP-A nucleosome are similar to those of the canonical H3 nucleosome. Therefore, the crystal structure of the CENP-A nucleosome is consistent with the octasome model.

In this Extra View article, we provide a detailed structural comparison between CENP-A and H3 in the nucleosomes.
Comparison of the CENP-A Structure with the H3 Structure in Nucleosomes

Eight H3 variants, H3.1, H3.2, H3.3, H3T (H3.4), H3.5, H3.X, H3.Y and CENP-A, have been identified in humans.[21] The H3.1, H3.2, H3.3, H3T and H3.5 variants have only one to several amino acid differences among them. Consistent with their high sequence conservation, minimal differences in the local structures were observed among H3.1, H3.2, H3.3 and H3T in nucleosomes.[22-24] On the other hand, CENP-A shares about 50% amino acid identity with the canonical H3.1, H3.2 or H3.3. Therefore, CENP-A is the most distant H3 variant. In the CENP-A nucleosome structure, amino acid residues 46–134 and 48–135 of each of the two CENP-A molecules are visible and the N-terminal residues (about 45 residues) and the C-terminal five or six residues are completely disordered (Fig. 1). Since the two CENP-A structures in the mono-nucleosome are almost identical, we averaged the two CENP-A structures for the comparison with the H3.1 structure in the nucleosomes.[22-24] Therefore, CENP-A is the most distant H3 variant. In the CENP-A nucleosome structure, amino acid residues 46–134 and 48–135 of each of the two CENP-A molecules are visible and the N-terminal residues (about 45 residues) and the C-terminal five or six residues are completely disordered (Fig. 1). Since the two CENP-A structures in the mono-nucleosome are almost identical, we averaged the two CENP-A structures for the comparison with the H3.1 structure in the nucleosomes. The CENP-A and H3.1 structures were superimposed and the root mean square deviation (RMSD) for each residue pair was calculated. The RMSD values of the main chain Cα atoms are plotted (Fig. 2). We found four CENP-A regions that exhibited large deviations (>1Å), which excludes amino acid residues 48 and 134 that are located at the edges of the structured region (Fig. 2). These CENP-A regions are amino acid residues 49–51, 79–85, 108–109 and 127–129.

The CENP-A 49–51 region is located just before the αN helix and does not form a stable α-helix, as discussed previously, although the corresponding region of the canonical H3 forms an α-helix (Fig. 2).[20] Therefore, the αN helix of CENP-A is shorter than that of the canonical H3. The CENP-A 79–85 region forms a loop structure (L1) connecting the α1 and α2 helices (Figs. 1 and 2). The CENP-A L1 loop is longer by two amino acid residues than the H3 L1 loop. Previously, we reported that both the length and the amino acid residue composition of the tip of the CENP-A L1 loop are important to maintain the CENP-A nucleosome at centromeres in cells.[20] Therefore, the CENP-A-specific L1 loop structure may function to interact with (α) factor(s) that associate(s) with CENP-A to stabilize the CENP-A nucleosome at centromeres.

The CENP-A–CENP-A′ Interface Structure in the CENP-A Nucleosome

Relatively large deviations were found in the CENP-A 108–109 and 127–129 regions (Fig. 2). Interestingly, these regions directly constitute the CENP-A–CENP-A′ interface in the CENP-A nucleosome (the dyad symmetry-related histones and its residues are denoted by ′) (Fig. 3A). The structure of the DNA-free CENP-A/H4 tetramer was reported by another group,[24] and showed that the CENP-A–CENP-A′ interface in the tetramer is significantly rotated.
as compared with the H3–H3' interface. However, unlike the free CENP-A/H4 tetramer, the CENP-A–CENP-A' interface was very similar to the H3–H3' interface in the CENP-A nucleosome (Fig. 3). The crystal structure of the HJURP–CENP-A–H4 complex revealed that the CENP-A–CENP-A' interface found in the CENP-A nucleosome is disrupted by the N-terminal α-helix of HJURP.25 HJURP has been identified as the CENP-A-specific chaperone;26,27 therefore, CENP-A may exist as the CENP-A–H4 dimer in the pre-assembled complex with HJURP, before its incorporation into chromatin.

In nucleosomes, van der Waals contacts are formed at both the CENP-A–CENP-A' interface and the H3–H3' interface (Fig. 3A and B). In the CENP-A nucleosome, the CENP-A Leu112, Leu128 and Ile132 residues form van der Waals contacts with the CENP-A' Leu111', Leu112' and His115' residues (Fig. 3A). On the other hand, in the H3 nucleosome, the H3 Cys110, Leu126 and Ile130 residues formed van der Waals contacts with the H3' Leu109' and His113' residues (Fig. 3B). In addition, in the H3.1 nucleosome, only one hydrogen bond, between His113 and Asp123' (and between Asp123 and His113'), is formed at the H3–H3' homodimeric interface (Fig. 3D). This hydrogen bond is conserved in the CENP-A–CENP-A' homodimeric interface between His115 and Asp125' (and between Asp125 and His115') in the CENP-A nucleosome (Fig. 3C). In contrast to the H3.1 nucleosome, an additional hydrogen bond between Arg131 and Asp108' (and between Asp108 and Arg131') was found in the CENP-A–CENP-A' interface (Fig. 3C). Arg131 also forms a hydrogen bond with Glu107'; however, this was not observed in the symmetry-related CENP-A' molecule (Glu107-Arg131'). The CENP-A Asp108 residue is located in the α2 helix and the hydrogen bond is also disrupted by the N-terminal α-helix of HJURP in the HJURP–CENP-A–H4 complex.25 These CENP-A-specific interactions were not seen in the H3.1 nucleosome, although the corresponding H3.1 residues, Arg129 and Asp106, are conserved (Fig. 3D). The formation of these CENP-A-specific hydrogen bonds induces structural deviations in the CENP-A 108–109 and 127–129 regions (Fig. 2).
The CENP-A C-Terminal Region

The constitutive centromere-associated network (CCAN) proteins were identified as factors required for the formation of the mitotic kinetochore. Among them, CENP-C and CENP-N reportedly bind to both CENP-A and DNA. The C-terminal region of CENP-A may be the binding site for CENP-C. In the CENP-A nucleosome structure, however, the CENP-A C-terminal 135–140 region is disordered (Fig. 4A). The C-terminal disordered region of CENP-A is exposed on the disk surface of the CENP-A nucleosome and may be accessible to CENP-C (Fig. 4B). In addition, the C-terminal region is located near the dyad axis of the nucleosomal DNA (Fig. 4B). CENP-C is a DNA-binding protein; therefore, it may simultaneously bind to both the CENP-A C-terminal tail and the nucleosomal DNA around the nucleosomal dyad.

CENP-B Orientation with Respect to the CENP-A Nucleosome

As we previously reported, both ends of the nucleosomal DNA are detached in the CENP-A nucleosome, unlike the canonical H3 nucleosomes. CENP-B is known to be required for de novo centromere formation on DNA lacking a functional centromere, and is found in the CENP-A associated kinetochore complex. CENP-B binds to the 17 base-pair CENP-B recognition sequence, called the CENP-B box, with its N-terminal region and dimerizes with its C-terminal region. For the crystallization of the CENP-A nucleosome, we used a human centromeric α-satellite DNA fragment containing the 17 base-pair CENP-B box sequence. Interestingly, in the CENP-A nucleosome structure, about 3/4 of the CENP-B box sequence is detached from the histone octamer surface, at the entrance and exit of the nucleosome. Our previous biochemical study revealed that the CENP-B box sequence is preferentially located near the entrance and exit of the nucleosome. Therefore, this position of the CENP-B box sequence in the CENP-A nucleosome may be physiologically relevant.

We previously determined the crystal structure of the DNA-binding domain of human CENP-B (CENP-B DBD) complexed with the CENP-B box DNA and found that the CENP-B box DNA is sharply kinked at two sites in the complex (Fig. 5A). Intriguingly, the CENP-B DBD-DNA complex can be accommodated in the CENP-A nucleosome without steric clash, when it is superimposed on the CENP-A nucleosome structure (Fig. 5B). This is consistent with the idea...
that the CENP-A nucleosome provides the fundamental architecture that induces localization of the assembly of the centromere-specific proteins and the flexible nature of the CENP-A nucleosomal DNA may be important for the assembly.

**Concluding Remarks**

In this Extra View article, we described the structural differences between CENP-A and canonical histone H3 in the nucleosome and discussed their functional significance. The global structure of the CENP-A nucleosome resembles that of the canonical nucleosomes, however, we found CENP-A nucleosome-specific structural features. (1) The nN helix of CENP-A is shorter than that of H3. (2) The L1 loop of CENP-A is longer than that of H3. (3) In the CENP-A nucleosome, specific hydrogen bonds are formed between the CENP-A protomers and induce structural deviations around the CENP-A-CENP-A interface. (4) The DNA segments are detached from the histone surface at the entrance and exit of the CENP-A nucleosomes. These structural features of the CENP-A nucleosome may be important for the specific chromatin architecture of centromeres.

**References**

1. Chomienne IM, Chen A. Molecular architecture of the kinesin-like motor molecule. Nat Rev Mol Cell Biol 2004; 5:3-15. PMID:14707779; http://dx.doi.org/10.1038/nrm1116
2. Santaguida S, Musacchio A. The life and miracles of ancient wrap. Nat Rev Mol Cell Biol 2008; 9:33-46. PMID:18097444; http://dx.doi.org/10.1038/nrm2310
3. Talbert PB, Henikoff S. Histone variants ancient wrap array of nucleosomal DNA contact sites. J Cell Biol 2007; 179:931-43. PMID:17520546; http://dx.doi.org/10.1083/jcb.200702022
4. Black BE, Cleveland DW. Epigenetic Centromere Propagation and the Nature of CENP-A Nucleosomes. doi:10.1038/nrc2861
5. Palmer DK, O’Day K, Wener MH, Andrews BS, Culture, Sports, Science and Technology (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. H.K. was also supported by the Waseda Research Institute for Science and Engineering.

**Acknowledgments**

We thank the beamline scientists for their assistance with data collection at the BL41XU beamline of SPring-8. We also thank T. Shiga, R. Matsumoto, Y. Miya, A. Osakabe and K. Saito (Waseda University) for technical assistance. This work was supported in part by Grants-in-Aid from the Japanese Society for the Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. H.K. was also supported by the Waseda Research Institute for Science and Engineering.

We thank the beamline scientists for their assistance with data collection at the BL41XU beamline of SPring-8. We also thank T. Shiga, R. Matsumoto, Y. Miya, A. Osakabe and K. Saito (Waseda University) for technical assistance. This work was supported in part by Grants-in-Aid for Scientific Research (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. H.K. was also supported by the Waseda Research Institute for Science and Engineering.
32. Saitoh H, Tomkitel J, Cooke CA, Rafter H, Jol, Marzi A, Rothfield NF, et al. CENP-C, an auto-antigen in scleroderma, is a component of the human inner kinetochore plate. Cell 1992; 70:115-25; PMID:1339310. http://dx.doi.org/10.1016/0092-8674(92)90538-N.

33. Sugimoto K, Yama H, Mizu Y, Hamata H, Human centromere protein C (CENP-C) is a DNA-binding protein which possesses a novel DNA-binding motif. J Biol Chem 1994; 269:117-31; PMID:8197574.

34. Carroll CW, Silva MC, Ghosh K, Zweis LE, Straight AF. Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. Nat Cell Biol 2009; 11:370-82; PMID:19453726. http://dx.doi.org/10.1038/nclb.2009.89.

35. Carroll CW, Mike AJ, Straight AF. Dual recognition of CENP-A nucleosomes is required for centromere assembly. J Cell Biol 2010; 189:1143-55; PMID:20566683. http://dx.doi.org/10.1083/jcb.201001013.

36. Guse A, Carroll CW, Mike AJ, Fuller CJ, Straight AF. In vitro centromere and kinetochore assembly on defined chromatin templates. Nature 2011; 477:354-8; PMID:21874020. http://dx.doi.org/10.1038/nature10379.

37. Okada T, Ohtaki J, Nakane M, You K, Entwistle WR, Lehto V, et al. CENP-C and CENP-B control centromere formation and affect the chromatid position. Cell 2001; 104:327-34; PMID:11296136. http://dx.doi.org/10.1016/S0092-8674(01)01047-0.

38. Saitoh H, Tomkitel J, Cooke CA, Rafter H, Jol, Marzi A, Rothfield NF, et al. CENP-C, an auto-antigen in scleroderma, is a component of the human inner kinetochore plate. Cell 1992; 70:115-25; PMID:1339310. http://dx.doi.org/10.1016/0092-8674(92)90538-N.

39. Sugimoto K, Yama H, Hamata H, Human centromere protein C (CENP-C) is a DNA-binding protein which possesses a novel DNA-binding motif. J Biol Chem 1994; 269:117-31; PMID:8197574.

40. Carroll CW, Silva MC, Ghosh K, Zweis LE, Straight AF. Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. Nat Cell Biol 2009; 11:370-82; PMID:19453726. http://dx.doi.org/10.1038/nclb.2009.89.

41. Muro Y, Masumoto H, Yoda K, Nishida N, Ohashi M, Okazaki T. Centromere protein B assembles human centromeric alpha-satellite DNA at the 17-bp sequence, CENP-B box. J Cell Biol 1992; 116:385-396; PMID:1375778. http://dx.doi.org/10.1083/jcb.116.3.385.

42. Pluta AF, Saitoh N, Goldberg S, Zweis LE, Identification of a subdomain of CENP-B that is necessary and sufficient for localization to the human centromere. J Biol Chem 2004; 279:3361-66; PMID:14662465. http://dx.doi.org/10.1074/jbc.M306477200.

43. Zweis LE, Celliers KR, Medlin PS, Cooke CA, Krane DA, Politi TL, et al. Molecular cloning of the DNA-binding domain containing MEZ region responsible for centromere localization. J Biol Chem 2004; 279:3361-66; PMID:14662465. http://dx.doi.org/10.1074/jbc.M306477200.

44. Muro Y, Masumoto H, Yoda K, Nishida N, Ohashi M, Okazaki T. Centromere protein B assembles human centromeric alpha-satellite DNA at the 17-bp sequence, CENP-B box. J Cell Biol 1992; 116:385-396; PMID:1375778. http://dx.doi.org/10.1083/jcb.116.3.385.

45. Tanaka Y, Nureki O, Kurumizaka H, Fukai S, Kurokawa S, Kora M, et al. Crystal structure of the CENP-B protein-DNA complex: the DNA-binding domain contains a novel alpha helix-turn-alpha helix-turn motif. J Mol Biol 1998; 282:769-84; PMID:9549656. http://dx.doi.org/10.1006/jmbi.1998.2119.

46. Todo H, Muro Y, Masumoto H, Yoda K, Nishida N, Ohashi M, Okazaki T. A human centromere antigen (CENP-B) interacts with a short specific sequence in alpha-satellite DNA, a human centromeric satellite. J Biol Chem 1989; 264:5934-46; PMID:1469042. http://dx.doi.org/10.1083/jbc.119.6.1413.

47. Okada T, Ohtaki J, Nakane M, You K, Entwistle WR, Lehto V, et al. CENP-C, an auto-antigen in scleroderma, is a component of the human inner kinetochore plate. Cell 1992; 70:115-25; PMID:1339310. http://dx.doi.org/10.1016/0092-8674(92)90538-N.

48. Masumoto H, Maeda H, Moro T, Nishida N, Ohashi M, Okazaki T. A human centromere antigen (CENP-B) interacts with a short specific sequence in alpha-satellite DNA, a human centromeric satellite. J Biol Chem 1989; 264:5934-46; PMID:1469042. http://dx.doi.org/10.1083/jbc.119.6.1413.

49. Tanaka Y, Nureki O, Kurumizaka H, Fukai S, Kurokawa S, Kora M, et al. Crystal structure of the CENP-B protein-DNA complex: the DNA-binding domain contains a novel alpha helix-turn-alpha helix-turn motif. J Mol Biol 1998; 282:769-84; PMID:9549656. http://dx.doi.org/10.1006/jmbi.1998.2119.