Short Conceptual Overview

Yuichi Ichikawa, Yoshifumi Nishimura, Hitoshi Kurumizaka and Mitsuhiro Shimizu*

Nucleosome organization and chromatin dynamics in telomeres

Abstract: Telomeres are DNA-protein complexes located at the ends of linear eukaryotic chromosomes, and are essential for chromosome stability and maintenance. In most organisms, telomeres consist of tandemly repeated sequences of guanine-clusters. In higher eukaryotes, most of the telomeric repeat regions are tightly packaged into nucleosomes, even though telomeric repeats act as nucleosome-disfavoring sequences. Although telomeres were considered to be condensed heterochromatin structures, recent studies revealed that the chromatin structures in telomeres are actually dynamic. The dynamic properties of telomeric chromatin are considered to be important for the structural changes between the euchromatic and heterochromatic states during the cell cycle and in cellular differentiation. We propose that the nucleosome-disfavoring property of telomeric repeats is a crucial determinant for the lability of telomeric nucleosomes, and provides a platform for chromatin dynamics in telomeres. Furthermore, we discuss the influences of telomeric components on the nucleosome organization and chromatin dynamics in telomeres.

Keywords: chromatin; nucleosome; telomere; telomere-binding proteins; telomeric repeat sequences.

Introduction

In eukaryotic chromosomes, genomic DNA is assembled into arrays of nucleosomes, the structural and functional units of chromatin. The nucleosome comprises the histone octamer and 146 bp of DNA (1, 2). Eukaryotic chromosomes are linear DNA molecules, and their ends, called telomeres, are composed of special nucleoprotein complexes. Telomeres are essential for genome stability and maintenance, and their structures change dynamically throughout the cell cycle and in cell differentiation (3–7). In telomeric repeats with the normal length, the telomere is in the closed form called the protected state, which is maintained by the shelterin complex with heterochromatic marks. As cell division progresses, telomere shortening leads to an open state called the deprotected state. Further shortening of telomeres causes the dysfunctional state, in which the telomeres are not protected from the DNA damage response machinery, resulting in disease-causing telomere fusions (3–7).

In most organisms, telomeric DNA consists of tandem 5–8 bp repeats with guanine-clusters (G-clusters) (Table 1), and the G-rich strand is extended to form a 3′-single-stranded overhang (G-tail) at the chromosome ends (8). For example, the telomeric repeats in human and budding yeast are (TTAGGG)_n and (TG)_{1-3}n, respectively. However, the telomere organizations are significantly different between lower and higher eukaryotes (4, 5, 7). First, the lengths of the telomeric repeats are relatively longer and variable in mammals, about 10–15 and 20–60 kbp in humans and mice, respectively. In contrast, the repeats are shorter, about 300 bp, in both fission yeast and budding yeast. In addition, in lower eukaryotes, the telomeric repeats have a nonnucleosomal chromatin structure, whereas in higher eukaryotes, they are organized within arrays of nucleosomes with short linker DNAs (Table 1). Furthermore, the components of the telomeric architecture are diverse among organisms. This review focuses on the nucleosome organization and the chromatin dynamics in telomeres, in terms of the structural properties of the telomeric repeat DNAs,
Nucleosome formation, and architectural components of the telomere.

**Nucleosome formation with telomeric repeat sequences in vitro**

Telomeres have the common sequence feature of a cluster of G-residues in the tandem repeats, in most organisms (Table 1). Nucleosome assembly on the telomeric sequences from various organisms has been extensively analyzed, by competitive reconstitution assays in vitro (4, 9–11). These studies revealed that all telomeric sequences tested show comparative positive free energies in nucleosome formation, relative to average nucleosomal DNA. These results imply that the affinities of telomeric sequences for histone octamers are the lowest among the sequences tested. Thus, telomeric DNAs have a feature that disfavors nucleosome formation in vitro.

The structures of nucleosomes reconstituted from various telomeric sequences have been examined by micrococcal nuclease (MNase) digestion. *Arabidopsis* telomeric DNA reconstituted in nucleosomes was protected from MNase digestion, resulting in an approximately 147-bp fragment, which is the canonical nucleosome length (9). It is worthwhile to note that footprinting studies, using DNase I, hydroxyl radicals, exonuclease III or λ exonuclease, revealed that telomeric nucleosomes occupy multiple positions without a rotational setting (9–11). Thus, there is no preferred location for nucleosomes on telomeric DNA sequences.

Pisano et al. (12) compared the mobility of nucleosomes reconstituted from the human telomeric sequence (TTAGGG)₂₇ (hTEL27) with those reconstituted from the standard average sequence, called TAND1. After the reconstitution, the nucleosomes were ligated to a longer sequence containing a nucleosome positioning sequence, the TG pentamer, and their mobility was examined. The histone octamer moved from the hTEL27 sequence to the TG pentamer, whereas the nucleosomes formed on the control TAND1 sequence remained, showing that telomeric nucleosomes have high intrinsic mobility under physiological conditions (12). Atomic force microscope (AFM) imaging analyses revealed that the nucleosome spacing on the relatively longer human telomeric repeats (about 800–1500 bp) varied randomly, whereas the spacing on the 5S ribosomal DNA and the 601 DNA, which are both known as nucleosome positioning sequences, was mostly uniform (13, 14). These studies indicated that the telomeric nucleosomes are labile and intrinsically mobile, consistent with the results of the competitive reconstitution and footprinting assays in vitro.

**Nucleosome organization in telomeric repeat sequences in vivo**

The nucleosome organization at telomeres on chromosomes has been extensively analyzed by MNase and DNase I digestions of nuclei isolated from various organisms (15–24). In a typical experiment, the DNA purified from the MNase- or DNase I-digests is resolved by agarose gel electrophoresis and analyzed by subsequent Southern blotting. In *Saccharomyces cerevisiae*, the subtelomeric repeats were assembled into nucleosomes, whereas the telomeric tracts of TG₁₋₃ were protected in DNA fragments with larger and more heterogeneous sizes than that of a nucleosome (23). These studies suggested that the yeast telomere is a nonnucleosomal chromatin structure.
Instead, the specific DNA-protein complex called the telosome is formed in the telomeric repeats (7). The yeast Rap1 protein binds to the telomeric repeat sequences to form the telosome, along with Rif1 and Rif2. A recent report revealed the X-ray structure of Rif1 and Rif2 bound to the Rap1 C-terminal domain, and proposed a model for the Rap1-Rif1-Rif2 assembly at telomeres (25). This structure led to the Velcro model for telomere organization, which may explain telomere homeostasis.

In contrast to the yeast telomeres, the telomeric repetitive elements in higher eukaryotes, such as mammals and some plants, are organized into packed arrays of nucleosomes, i.e., closely spaced nucleosomes (15–22). This is inconsistent with the fact that both the human and yeast telomeric sequences have lower affinities for histone octamers. Regarding this issue, it should be noted that Tommerup et al. (17) suggested that nucleosome-free regions exist in the unusual chromatin in the human telomere. They examined the telomeric chromatin in two subclones of a HeLa cell line, in which one subclone (HeLa-L) carries long telomeric repeats of ~22.5 kb, and the other subclone (HeLa-S) has much shorter telomeres with only 2–7 kb telomeric repeats. An array of tightly packed nucleosomes is formed in the longer telomeric repeats in HeLa-L cells, whereas an altered form of chromatin, characterized by diffuse MNase patterns, is formed in the shorter telomeric repeat region in HeLa-S cells. Thus, it was speculated that human telomeres have a bipartite chromatin structure, with an array of packed nucleosomes in the proximal region of the telomere and unusual chromatin near the telomere terminus. The diffuse MNase pattern could be derived from the nucleosome-free regions, but the numbers and sizes of the nucleosome-free regions, as well as their locations, cannot be predicted (17).

The long tandem repeats in human telomeres make it difficult to obtain detailed structural information. It is also challenging to discuss the exclusive effect of the DNA sequence because such cell biological studies cannot exclude the effects of endogenous telomere-binding factors. Therefore, we utilized the yeast minichromosome system, which consists of an array of positioned nucleosomes (26, 27) to examine the effects of human and yeast telomeric repeat sequences on nucleosome positioning in vivo (28). It allows a detailed analysis of properly positioned nucleosomes, without human telomere-associated factors, in yeast cells. Thus, the effects of the DNA elements of telomeric repeats, such as the number of repeats and the base composition, could be solely examined in vivo, as described in the next section.

Nucleosome-disfavoring properties of telomeric repeats in vivo, as revealed by the yeast minichromosome system

Previously, we developed an in vivo assay system to examine the effects of DNA sequences on nucleosome formation, using the defined yeast minichromosome system (26, 27). The TALS circular DNA is a derivative of the TRPIARS1 plasmid (29), which contains the α2 operator, and it is assembled into minichromosomes harboring an array of positioned nucleosomes in MAT α cells (30, 31). The minichromosome system comprising positioned nucleosomes allowed us to evaluate the detailed effects of DNA structures on the formation of nucleosomes in vivo.

We inserted various lengths of human and yeast telomeric repeat sequences (hTELn and yTELn, respectively, see Figure 1A) into the center of nucleosome IV in the TALS minichromosome (28). MNase digestion-based analyses of the yeast minichromosomes revealed that both the human and yeast telomeric repeat sequences disfavor nucleosome formation in vivo, in a length-dependent manner. Specifically, the shorter hTEL2 and hTEL4 inserts were incorporated near the pseudo-dyad axis of nucleosome IV, whereas the longer inserts of hTELn (n=6, 12, 29) and yTELn (n=4, 5, 9) (see Figure 1A) disrupted the positioning of nucleosome IV. To clarify the sequence requirements for this feature, we further analyzed the insertion of sequence isomers of the telomeric repeat sequences, called SI-An and SI-Bn [(TGTAGG)n and (TGTGAG)n] (n=6, 12, respectively). These SI-An and SI-Bn sequences have the same base composition as the human telomeric repeat, but the G-cluster portion, which is a characteristic of the telomeric repeat, is disrupted in the repeats. In contrast to the longer hTELn (n≥6) insert, all of the sequence isomers (SI-An and SI-Bn, n=6, 12) were incorporated in the positioned nucleosome IV (the results are summarized in Figure 1B). Thus, the periodic appearance of the G-clusters is responsible for the nucleosome-disfavoring properties of the telomeric repeats (28).

Two mechanisms are mainly considered for the nucleosome-disfavoring properties of the longer telomeric repeats. In one mechanism, the human telomeric repeats are occupied by yeast proteins, such as Tbf1 (TATAGG-binding protein factor 1) and Rap1 (a yeast telomeric-binding protein) (32–34), to prevent histone octamer binding. In the other mechanism, the anti-bending property of the telomeric repeat DNA confers its resistance to nucleosome formation. As we discussed previously (28), it is most likely that an inherent structural feature of the DNA is the major
determinant for the nucleosome-disfavoring properties. We will discuss the components influencing the nucleosome organization and chromatin dynamics in telomeric repeat sequences, and clarify their properties in the next section.

Components that influence nucleosome organization in telomeric chromatin

Structural properties of telomeric DNA repeat sequences

Our studies support the idea that the inherent structural features of the DNA in telomeric repeat sequences are critical for nucleosome formation (28). The 24-bp insert (hTEL4) was incorporated into nucleosome IV, in contrast to the 36-bp insert (hTEL6). Figure 1C shows the putative DNA path of the 36-bp insert, which was superimposed on the crystal structure of the yeast nucleosome (35), and would span from superhelix axis location (SHL) +1 to -2.5. As the telomeric repeat length increases, the flexibility of the longer repeat DNA could become more limited, especially around at SHL -1.0 to -2.5, thus inhibiting the formation of nucleosome IV in the yeast minichromosome (28). As described below, several lines of evidence support the idea that the structures of DNA sequences containing consecutive G-residues differ from that of the canonical B-DNA.

The crystal structure of the DNA oligomer d(G-G-G-G-C-C-C-C) revealed that its overall structure is similar to an A-form duplex, with a shallow minor groove and a deep
major groove (36). The aqueous structures of this oligomer and its sequence isomer d(C-C-C-C-G-G-G) were analyzed by CD and NMR spectroscopies, together with unrestrained molecular dynamics (37). The structure of d(C-C-C-C-G-G-G) is quite unusual, as it forms an A-type double-helix with B-type sugar puckering. In the aqueous structure of the DNA duplex of d(G-G-G-G-C-C-C), the G-G stacking is A-like (38). In addition, poly(dG)-poly(dC) in solution can adopt the A-form, depending on the environment (39). Thus, the A-like structural properties of the G-cluster segments (36–40), which may be more rigid than the canonical B-form, would be responsible for inhibiting the incorporation of longer telomeric repeats within the nucleosome.

Alternatively, telomeric repeat sequences can adopt four-stranded DNA structures, called G-quadruplexes or G4 DNA (41). The G-quartet is the basic structural unit of G4 DNA, in which four G-bases are connected by Hoogsteen-base pairs and assembled in a square planar manner. In telomeres, the 3′ single-stranded overhang of the G-rich strand has the potential to form the G4 structure, which may be important for telomeric DNA protection (41, 42). It was also proposed that G4 structures can form on the G-rich template for lagging strand synthesis during the replication of the double-stranded telomeric repeat regions, causing replication fork arrest at telomeric repeats (41, 42). In fact, G4 structures have been quantitatively visualized in human cells, using an engineered, structure-specific antibody (43). However, it was recently shown that G4 formation is not required for DNA polymerase δ stalling on telomeric lagging strands (44). Taken together with consideration of our structural studies (28), it is uncertain whether the double-stranded telomeric repeats in the yeast minichromosomes form G4 DNA during DNA replication, which should prevent nucleosome formation.

**Telomere repeat direct binders TRF1 and TRF2**

As mentioned earlier, the nucleosome-disfavoring properties of the human telomeric repeats, as revealed by the in vitro reconstitution studies and the in vivo assay system using the yeast minichromosome, argue against the fact that packed nucleosomes with short spacing are formed in the tandem repeats of human telomeres in vivo. This inconsistency suggests that telomeric components may play an important role in both the formation and dynamics of nucleosomal arrays in the telomeric repeats in human telomeres.

Mammalian telomeres are assembled into a specialized chromatin structure containing the shelterin complex, composed of six proteins: TRF1, TRF2, POT1, RAP1, TIN2, and TPP1 (3–5). The shelterin complex is essential for telomere length maintenance and chromosome protection. TRF1 and TRF2 have a Myb-like domain and specifically bind to the double-stranded DNA sequence of GGGTAGGG in two repeats (45, 46), and occupy approximately 30% and 15% of all repeat regions, respectively (47).

To examine the interaction between TRF1 and nucleosome, Cacchione and co-workers (48) performed a gel mobility shift assay and DNase I footprinting in vitro. TRF1 specifically recognizes its binding sites located within nucleosomes, and forms a ternary complex. TRF1 binding to a nucleosome depends on the orientation of the binding sites on the nucleosome surface, and causes alterations in the nucleosome structure without histone octamer dissociation. Furthermore, TRF1 specifically induces the mobility of telomeric nucleosomes, which are intrinsically mobile, and also causes the compaction of telomeric DNA through TRF1-nucleosome interactions (49). Thus, TRF1 has the ability to regulate the telomeric chromatin structure, including the nucleosome organization, through its various interactions with nucleosomes, such as binding to a site within nucleosomes, inducing nucleosome mobility, and compacting telomeric DNA regions.

The effects of TRF2 on nucleosome organization have also been extensively studied in vivo (50–52) and in vitro (52–54). A recent chromatin immunoprecipitation and MNase mapping assay demonstrated that TRF2 overexpression reduces the nucleosome density and increases the internucleosomal distance at telomeres in vivo (52). These results confirmed a previous report that TRF2 overexpression increases nucleosome spacing at telomeres in transgenic mice (50). When nucleosome arrays were assembled on telomeric repeats in the presence of TRF2 in vitro, the distance between nucleosomes increased and became less regular (52). In vitro analytical agarose gel electrophoresis and AFM studies revealed that TRF2, as well as the TRF2 DNA binding domain, promotes the folding of nucleosomal arrays into more compact structures (53, 54), thus generating results that seem to be inconsistent with the in vivo studies mentioned earlier. Although the relationship between nucleosome spacing and chromatin compaction remains uncertain, it is apparent that TRF2 affects the telomeric chromatin structure.

However, Wu and de Lange (51) reported no evidence of altered nucleosomal spacing in telomeric chromatin or nucleosome eviction near the telomere terminus by the deletion of TRF2, although the telomere structures became deprotected. Furthermore, even when the entire shelterin complex was removed from mouse telomeres, by the double-knockout of TRF1 and TRF2, no significant
differences were observed in the nucleosome ladders generated by MNase digestion (55). Specifically, the telomeric repeat DNA remained packaged within nucleosomes in the absence of the shelterin complex. Thus, it is speculated that TRF1 and TRF2 can alter nucleosome organization in telomeres, but once the nucleosomes are formed, the removal of TRF1 and TRF2 does not affect the nucleosome organization. It is also possible that telomeric components other than the shelterin complex maintain nucleosomal arrays in telomeric repeat regions.

Although TRF1 and TRF2 contain similar functional domains, a remarkable difference exists in their N-terminal regions, which are acidic and basic in TRF1 and TRF2, respectively. Unlike TRF1, TRF2 introduces positive supercoiling, and hence negative torsion is created on the adjacent DNA (56). Interestingly, the deletion of the acidic domain in TRF1 causes it to introduce positive supercoiling into DNA (57). The TRF2-specific property is thought to lead to the stimulation of the invasion of single-stranded telomeric repeats to promote telomeric DNA looping (called a t-loop), and nucleosomes have the ability to facilitate this TRF2 activity, which may be involved in stabilizing the looped telomere structures (56, 57). Thus, we speculate that each characteristic N-terminal domain may exert unique effects on the nucleosome organization in telomeres.

**Histone variants, modifications, and chaperones**

Histone post-translational modifications, histone variants, and histone chaperones play important roles in the regulation of transcription, as well as DNA replication, repair and recombination, and specific modifications and variants are often referred to as epigenetic marks (2). Telomeres had previously been regarded as heterochromatin associated with condensed chromatin, in transcriptionally silent chromosomal regions. The epigenetic marks in mouse and human telomeres have been extensively studied. The telomeric repeat sequences are identical between mouse and human, but their lengths in mice are longer (about 20–60 kbp) than those in humans (about 10–15 kbp). In mouse telomeres, heterochromatic marks such as H3K9me3 and H4K20me3 are enriched, and histones H3 and H4 are hypoacetylated (58, 59). In contrast, the levels of heterochromatic marks, H3K9me3, H4K20me3 and H3K27me3, are relatively low in telomeres in human cells (60). Similarly, H3K9me3 and H3K36me are the least significant modifications detected at telomeres in CD4+ T-cells (61). Notably, a recent study revealed that the H3K9me3 density increased with the elongation of telomeres in human cells (62). Thus, the difference in the level of H3K9me3 in telomeres between mouse and human can be attributed to the difference in their telomere lengths, consistent with the report that the H3K9me3 density varies with telomere length in telomerase-deficient mice (63). The removal of the acetyl group of H3K9 residues by SIRT6, the NAD+-dependent histone deacetylase, is also important for telomere maintenance. The SIRT6 knockdown causes premature cellular senescence and telomere dysfunction (64). These results indicate that the establishment of the heterochromatin state in human telomeres is required for chromosome stability. It is likely that heterochromatin plays a role in the tightly packed nucleosomes in human telomeres, detected as nucleosome ladders by MNase digestions, as discussed earlier.

Meanwhile, euchromatic marks are detectable in human telomeres. For example, the most significant modifications at telomeres are H2BK5me1 and H3K4me3, which are often associated with actively transcribed genes in CD4+ T-cells (61). The transcriptional regulator mixed lineage leukemia (MLL), a mammalian *trithorax*-group gene product, associates with telomeres and contributes to their H3K4 methylation and transcription in a telomere length-dependent manner, and the level of H3K4 methylation decreases with the shortening of the telomere length (65). Thus, the current view is that both heterochromatic and euchromatic marks are detectable in human telomeres, dependent on the telomere lengths.

Furthermore, in murine embryonic stem cells (ES cells) and neuronal precursor cells, the histone H3 variant H3.3 is localized in actively transcribed genes, as well as in telomeres (66–68). Hira and DAXX are H3.3 chaperones; the former is required for H3.3 enrichment at active genes, whereas in the complex with ATRX, the latter facilitates the deposition of H3.3 at telomeres (67, 68). Ser31-phosphorylated H3.3 (H3.3S31P) is enriched at telomeres during mitosis in pluripotent mouse ES cells, but not at telomeres in other nonpluripotent mouse and human cells. Cellular differentiation leads to a decrease in the H3.3S31P signals and to an increase in the H3K9me3 and H4K20me3 levels, and the telomere chromatin in ES cells becomes more MNase-resistant following differentiation (66). Thus, cellular differentiation-mediated histone modifications at the telomeres affect the underlying telomeric nucleosome organization, suggesting that the chromatin dynamics in telomeres are relevant to cellular differentiation. Monitoring of the genome-wide dynamics of replication-independent H3.3 indicated that the turnover of the H3.3 enriched in telomeres is very slow, which may be necessary for the maintenance of telomeric stability (69).
Thus, the deposition of H3.3 at telomeres, its modifications, and turnover may be determinants for the dynamic nature of telomeric chromatin.

As discussed earlier, telomeric chromatin contains epigenetic marks associated with both heterochromatin and euchromatin. In fact, telomeric repeat regions are transcribed into the noncoding RNA called TERRA (telomeric repeat containing RNA) (70). TERRA has been implicated in many processes, such as heterochromatin formation, telomere end protection, control of telomere length, replicative senescence, and so on (71). In relation to these functions, TERRA interacts with several telomeric components, including TRF1, TRF2, and ORC (72). Thus, TERRA is an important regulator of telomere structure and function (71–73).

Concluding remarks and perspectives

In spite of the nucleosome-disfavoring properties of telomeric repeat DNA, human telomeres are packaged into nucleosomal arrays. The formation of nucleosomal arrays in chromosomes should be more stable, as compared to a long nucleosome-free region. The covering of telomeric repeats by nucleosomes would suppress inappropriate recombination, thereby preventing telomere fusion and loss of telomere length control. In addition, nucleosome formation retains the telomeric repeat DNA in the B-form structure, thereby restraining it from the inappropriate formation of the G4 structure. Thus, packaging nucleosomes are necessary for the proper regulation of telomere function.

Although telomeres were considered to be condensed heterochromatin structures, recent studies have revealed that the chromatin structures are highly dynamic in telomeres. For instance, the telomeric repeat regions are transcribed into the noncoding RNA called TERRA, and they contain both the heterochromatic marks and euchromatic marks of histone modifications and variants. The dynamic properties of telomeric chromatin are considered to be important for the structural changes between the euchromatic and heterochromatic states, as well as for the changes between the protected and deprotected states during the cell cycle. As the transcribed TERRA interacts with TRF1, TRF2, and ORC (origin recognition complex), it is postulated that TERRA, together with telomeric proteins, participates in the inter-conversion between the euchromatic and heterochromatic states. Heterochromatin formation with telomeric components would contribute toward stabilizing nucleosomal arrays at human telomeres. We propose that the nucleosome-disfavoring property of the telomeric repeats provides a platform for such dynamics of telomeric chromatin, as shown in Figure 1D. The molecular mechanisms underlying the dynamics of telomeric chromatin, including the alterations of the nucleosome organization, remain elusive. Furthermore, it is possible that the nucleosome organization and dynamics may be regulated by the telomere organization in the nucleus, such as telomere clustering at the nuclear periphery. Future studies on the mechanisms should provide important clues toward understanding telomere biology.

Acknowledgments: This work was partly supported by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) (Grants-in-Aid, PDIS) (YI, YN, HK, and MS) and from the Japan Society for the Promotion of Science (JSPS) (Grants-in-Aid) (to HK and MS), and the Platform for Drug Discovery, Informatics, and Structural Life Science (YN and HK) from the MEXT, Japan. HK was also supported by the Waseda Research Institute for Science and Engineering (RISE).

List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFM</td>
<td>atomic force microscope</td>
</tr>
<tr>
<td>ES cell</td>
<td>embryonic stem cell</td>
</tr>
<tr>
<td>G-cluster</td>
<td>guanine-cluster</td>
</tr>
<tr>
<td>MNase</td>
<td>micrococcal nuclease</td>
</tr>
<tr>
<td>SHL</td>
<td>superhelix axis location</td>
</tr>
</tbody>
</table>

References