The multistructural forms of box C/D ribonucleoprotein particles

GE YU,1 YU ZHAO,2 and HONG LI1,2

1Department of Chemistry and Biochemistry, Florida State University, Tallahassee, Florida 32306, USA
2Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida 32306, USA

ABSTRACT

Structural biology studies of archaeal and yeast box C/D ribonucleoprotein particles (RNPs) reveal a surprisingly wide range of forms. If form ever follows function, the different structures of box C/D small ribonucleoprotein particles (snoRNPs) may reflect their versatile functional roles beyond what has been recognized. A large majority of box C/D RNPs serve to site-specifically methylate the ribosomal RNA, typically as independent complexes. Select members of the box C/D snoRNPs also are essential components of the megadalton RNP enzyme, the small subunit processome that is responsible for processing ribosomal RNA. Other box C/D RNPs continue to be uncovered with either unexpected or unknown functions. We summarize currently known box C/D RNP structures in this review and identify the Nop56/58 and box C/D RNA subunits as the key elements underlying the observed structural diversity, and likely, the diverse functional roles of box C/D RNPs.

Keywords: snoRNP structures; ribosome biogenesis; rRNA modification; fibrillarin; Nop56; Nop58; Nop5; kink-turn

INTRODUCTION

A key step in ribosome production is ribosomal RNA (rRNA) biogenesis. Ribosomal RNA are processed and covalently modified and finally assembled with ribosomal proteins in a well-maintained order by a large collection of biogenesis factors (Fatica and Tollervey 2002; Dez and Tollervey 2004; Thomson et al. 2013; Woolford and Baserga 2013). Among these, small nucleolar ribonucleolar protein particles (snoRNPs) both modify and participate in processing of rRNA, and are likely critical to the maintenance of the order of the rRNA maturation process (Maxwell and Fournier 1995; Decatur and Fournier 2003; Yu and Meier 2014). SnoRNPs are comprised of a snoRNA and partner proteins that can vary in composition depending on their actual roles in rRNA biogenesis (Decatur and Fournier 2003; Liang and Li 2011). A large majority of snoRNPs is responsible for site-specific 2′-O-methylation (box C/D-type snoRNPs) or pseudouridylation (box H/ACA-type snoRNPs) and contains four core proteins. Several members of the snoRNPs, such as U3, U14, snR30 (U17/E1 in humans), snR10 in yeast (ACA21 in humans), and U8 in Xenopus, are required for or promote pre-rRNA cleavage (Kass et al. 1990; Beltrame and Tollervey 1995; Maxwell and Fournier 1995; Allmang et al. 1996; Decatur and Fournier 2003; Liang et al. 2010; Watkins and Bohnsack 2012; Langhendries et al. 2016). The snoRNPs involved in pre-rRNA processing, in contrast to those involved in modification, are essential for cell viability. For instance, the U3 snoRNP has been found essential for rRNA processing in human, yeast, and Xenopus cells (Borovjagin and Gerbi 1999; Phipps et al. 2011). Some of these special snoRNAs, such as U3 and snR30, are much larger in size and contain more elaborate snoRNA and potentially additional proteins beyond the four core proteins associated with the modification snoRNPs (Dragon et al. 2002; Lemay et al. 2011). Regardless of the type, snoRNPs interact with pre-rRNA or other potential targets through RNA–RNA base-pairing and function cotranscriptionally with pre-rRNA, in which modification takes place first, followed by the processing events (Fatica and Tollervey 2002), although some exceptions to this order have been observed (Sloan et al. 2015).

The molecular basis of how different types of snoRNPs are assembled and function in maintaining the order of pre-rRNA maturation is of particular interest to understanding rRNA biogenesis and the related processes. The questions being addressed include if all box C/D...
snoRNPs share the same structural core, and if not, how they differ and if the differences are related to the involved functional processes. Several excellent reviews are available that describe the current understanding of the assembly, biogenesis, and impact on human diseases for both types of snoRNPs (Matera et al. 2007; Liang and Li 2011; Watkins and Bohnsack 2012; Yu and Meier 2014; Lafontaine 2015; Massenet et al. 2017). Here, we provide a summary of the structures of the box C/D type of snoRNP in light of the structural data that have recently become available, and we discuss their potential implications on functions.

Of the snoRNPs, the box C/D type is the most diverse, owing to their compositional as well as structural diversity. The core assembly of each box C/D snoRNP is comprised of a box C/D snoRNA and three (in Archaea) or four (in Eukarya) proteins (nucleolar protein 56, or Nop56p in yeast; nucleolar protein 58, or Nop58p in yeast and Nop5 in archaea; nucleolar protein 1, or Nop1p or fibrillarin; and L7Ae in archaea, small nuclear protein 13, or Snu13p in yeast or 15.5 kD, more recently named SNU13, in human; see complete description in Fig. 1). Fibrillarin contains the well-conserved S-adenosine-L-methionine (SAM) binding motif and is thus the methyltransferase center. Eukaryotic fibrillarin also contains a glycine-arginine-rich or GAR domain whose intrinsically disordered region (IDR) is believed to be important to organization of the nucleolus (Feric et al. 2016). The smallest protein subunit, Snu13p/SNU13, contains a well-conserved RNA binding motif and is also a component of the messenger RNA splicing machinery U4 snRNP (Nottrott et al. 1999; Watkins et al. 2000). Nop56p/NOP56 and Nop58p/NOP58 are a pair of homologous proteins that are also homologous to archaeal Nop5. Nop5, Nop56p/NOP56, and Nop58p/NOP58 are multidomain proteins. Interestingly, the U4 snRNP also contains a Nop56/58p-related protein Prp31p in addition to Snu13p/SNU13, suggesting a shared RNP domain.

The knowledge of box C/D snoRNP architecture first came from structural studies of the archaeal box C/D small RNP (sRNP) (Bleichert et al. 2009; Xue et al. 2010; Liang and Li 2011; Lin et al. 2011; Lapinaite et al. 2013) and more recently from electron microscopy studies of the yeast ribosome small subunit processome that contains U3 snoRNP as its core subcomplex (Barandun et al. 2017; Chaker-Margot et al. 2017; Sun et al. 2017). These studies have revealed a remarkably wide range of structural conformations as well as oligomerization in s(no)RNPs (Fig. 2). We describe these structural results and address the controversial topic of oligomerization. In analyzing both previously and recently available evidence supporting a functional diversity, we conclude that the structural diversity of box C/D snoRNPs may reflect the functional diversity.

**FIGURE 1.** Anatomical features of box C/D small (nucleolar) (sno) ribonucleoprotein particles (RNPs), U14, and U3 snoRNAs. The secondary structure of an archetypical C/D RNA is depicted by lines and conserved box C (RUGAUGA) and D (CUGA) sequences are boxed. The core proteins are represented by filled ovals and labeled with yeast protein names. The corresponding names in human and archaea are listed under the cartoon. Parentheses indicate alternative names. Each C/D box is associated with an antisense region (D or D’ guide) located upstream of box D (D guide) or D’ (D’ guide) that base-pairs with the target RNA. For methylation C/D (s)noRNAs, the target RNA is methylated on the 2’-ribose of the nucleotide base-paired with the 5th guide nucleotide upstream of box D or D’. Both U14 and U3 snoRNAs in yeast, though associated with the same proteins, do not completely comply with this rule and participate in ribosomal RNA processing.
Evidence for functional diversity of C/D snoRNPs

The functional diversity of box C/D s(no)RNPs was already known at the time of their discovery (Maxwell and Fournier 1995). Whereas most box C/D s(no)RNPs mediate site-specific 2'-O-methylation, several, such as the U3 and U14 snoRNPs, contain regions extensively paired with pre-rRNA and facilitate assembly of the 90S processome for processing pre-rRNA; although in Xenopus oocytes, U14 was shown to be primarily required for methylation (Dunbar and Baserga 1998). More recent studies have revealed functions involving box C/D snoRNPs beyond those traditionally known. Specialized box C/D snoRNAs, such as orphan box C/D snoRNAs, snR4 and snR45, can act as antisense guides to target RNA base acetylation, likely through a uniquely assembled complex (Sharma et al. 2017). A long noncoding RNA, LoNA, serves as a dual modulator of both rRNA transcription and methylation by trapping through its snoRNA motif box C/D proteins. Interestingly, depletion of the four snoRNP proteins did not impact rRNA processing but noticeably reduced the frequency of rRNA methylation, which suggests a different assembly mechanism of processing from the methylation snoRNPs (Li et al. 2018). Evidence was also found in an archaeal sRNP that fibrillarin and Nop5 complex could have RNP-independent methylation activity through a guide-free target RNA binding model (Tomkuvienë et al. 2017). A long noncoding RNA, LoNA, serves as a dual modulator of both rRNA transcription and methylation by trapping through its snoRNA motif box C/D proteins. Interestingly, depletion of the four snoRNP proteins did not impact rRNA processing but noticeably reduced the frequency of rRNA methylation, which suggests a different assembly mechanism of processing from the methylation snoRNPs (Li et al. 2018). Evidence was also found in an archaeal sRNP that fibrillarin and Nop5 complex could have RNP-independent methylation activity through a guide-free target RNA binding model (Tomkuvienë et al. 2017). A human snoRNA, SNORD27, regulates E2F7 pre-mRNA alternative splicing by masking splice sites through base-pairing. Interestingly, the SNORD27 snoRNP lacks fibrillarin, suggesting a possibly different assembly from other representative snoRNPs (Falaleeva et al. 2016). Global analysis of RNA modification revealed that inhibition of human fibrillarin led to an uneven decrease in the methylation sites—those undermethylated decreased more while those in peptidyl transferase center (PTC) did not reduce (Erales et al. 2017). This result suggests heterogeneity either in human snoRNP assembly or in rRNA accessibility.

Several human box C/D snoRNPs are found to be critical determinants of leukemic stem cell activity, via interactions with the groucho-related amino-terminal enhancer (AES)-induced RNA helicase DDX21 (Zhou et al. 2017). In addition, approximately half of human snoRNPs have no predictable RNA targets (Falaleeva et al. 2017). These select but convincing examples of snoRNP functions suggest that the box C/D s(no)RNPs may adapt to different structural forms necessary for specific interactions and functions.

EVIDENCE FOR STRUCTURAL DIVERSITY WITHIN THE C/D RNP FAMILY

Box C/D RNP structures contain flexible hinges

Crystal structures of individual box C/D proteins come from studies of archaeal homologs, with the exception of the smallest subunit SNU13/Snu13p/L7Ae, for which structures of human, yeast, and archaeal homologs are solved (Vidovic et al. 2000; Hamma and Ferré-D’Amaré 2004; Moore et al. 2004; Oruganti et al. 2005). Due to high sequence homology, structures of eukaryotic fibrillarin and Nop56/58 proteins are believed to be similar to those of their archaeal counterparts. However, eukaryotic fibrillarin and Nop proteins contain additional IDRs outside the conserved cores that are believed to serve functions unique to eukaryotic organisms. Structural analysis of individual components of the box C/D RNP is described below.

Fibrillarin belongs to the Class I methyltransferases (MTases), which share a common structural domain comprising central seven β-sheets mixed with α-helices (Cheng and Roberts 2001; Martin and McMillan 2002). The fold of fibrillarin undergoes little change upon assembly with the Nop56/58 proteins or with RNA, suggesting that the MTase contains a stable fold required for its catalytic activity. A similar structural stability is observed in the...
small subunit SNU13/Snu13p/L7Ae protein. The SNU13/Snu13p/L7Ae subunit contains a well-conserved RNA binding fold that is found in many other RNA binding proteins, especially ribosomal proteins (Koonin et al. 1994).

A prominent site of structural flexibility is found within the box C/D RNA. Each box C/D RNA contains two characteristic motifs, box C/D and box C′/D′ (or box B/C and box C′/D in U3 snoRNA) (Fig. 1), each undergoing structural transitions upon association with proteins. The structural transition and its dependence on the chemical environment have been demonstrated by fluorescence resonance energy transfer experiments (Turner et al. 2005) and by crystal structures of the SNU13/L7Ae-bound U4 small nuclear RNA (snRNA) or box C/D RNA (Vidovic et al. 2000; Watkins et al. 2000), and some messenger RNA regulatory regions (Chao et al. 2003). These K-turns play a similar role in controlling RNP assembly by introducing a protein-dependent bent in RNA. In vitro studies showed that this structural transition in the U4 snRNA and box C/D snoRNA initiated by the SNU13/Snu13p/L7Ae protein, the U4 snRNA, box C/D or box C′/D′ motifs form a sharp bent (kink-turn or K-turn) structure stabilized primarily by A-minor interactions and base stacking involving two tandem A-G mismatches. The formation of K-turn is highly predictable and is observed in ribosome (Klein et al. 2001), spliceosome (Nottrott et al. 1999; Vidovic et al. 2000; Watkins et al. 2000), and some messenger RNA regulatory regions (Chao et al. 2003). These K-turns may rotate (Fig. 3). In addition, the two long coils of the coiled-coil domain can adopt two positions: parallel or antiparallel (Fig. 3). Two important hinges around which domains may reposition are indicated. The observed placement of domains in currently known box C/D RNP structures is compared. The abbreviated organism names represent: Sulfolobus solfataricus (Ss), Archaeoglobus fulgidus (Af), Methanococcus jannaschi (Mj), Pyrococcus furiosus (Pf), and Saccharomyces cerevisiae (Sc).

FIGURE 3. Schematics of the key assembly subunit, Nop5p/S6p/S8p, of the box C/D RNPs. Nop5p/S6p/S8p comprises three domains: the NTD, the coiled-coil domain, and the CTD. Two important hinges around which domains may reposition are indicated. The observed placement of domains in currently known box C/D RNP structures is compared. The abbreviated organism names represent: Sulfolobus solfataricus (Ss), Archaeoglobus fulgidus (Af), Methanococcus jannaschi (Mj), Pyrococcus furiosus (Pf), and Saccharomyces cerevisiae (Sc).

The two K-turns of the box C/D snoRNPs could create conformational flexibility required for specific assemblies.

The most significant structural variation observed in s(no)RNP assembly lies in the Nop5p/S6p/S8p subunit. Unlike that which takes place within the box C/D RNA, structural transitions within the Nop5p/S6p/S8p are ill-defined and span a wide spatial range that seems to depend on the associated s(no)RNA, the presence of the target RNA, and the species where the s(no)RNP are from. Nop5p/S6p/S8p contains three distinct domains, an N-terminal domain (NTD), a coiled-coil domain, and a C-terminal domain (CTD) that is also known as the NOP domain. The elements connecting the three domains form two hinges around which the NTD and CTD of Nop5p/S6p/S8p may rotate (Fig. 3). In addition, the two long coils of the coiled-coil domain can adopt two positions: parallel or crossed (see illustration in Fig. 3). Since these two coiled-coil positions do not seem to correlate with the species analyzed, the quaternary structure, the method used for structural determination, or the functional state of box C/D RNPs, they are illustrated in Figure 3 but are not discussed in depth here.

Studies of the archaean Nop5 protein revealed a large range of domain movements (Aittaleb et al. 2003; Oruganti et al. 2007; Ye et al. 2009) that are likely shared by its eukaryotic counterparts. The most flexible hinge within Nop5/Nop5p/S6p/S8p is that connecting NTD to the
coiled-coil domain (hinge N, Fig. 3). NTD binds fibrillarin, and its varying positions place fibrillarin differently in space and are related to the state of box C/D RNP catalytic function. In the crystal structures of the isolated archaeal Nop55-fibrillarin, whether bound or not with box C/D RNA, a range of fibrillarin placement has been observed (Figs. 3, 4). In the Pf box C/D RNP halfmer (two separate RNA oligos representing box C/D and C/D′ motif, respectively), NTD is well separated from CTD, resulting in a horizontal “Y” shape with CTD and NTD forming two branches and the coiled-coil in the lower stem (Fig. 4). In a similar halfmer structure of the Ss box C/D RNP, NTD rotates further away from CTD but maintains the general “Y” shape. Several RNA-free Nop55-fibrillarin structures showed an even greater range of NTD placement through rotation around the hinge N (Ye et al. 2009). Interestingly, when Ss box C/D RNP was crystallized using a symmetric box C/D RNA and the corresponding target RNA, NTD and the bound fibrillarin was observed to be placed perfectly for catalyzing methylation of the target nucleotide (Lin et al. 2011; Yang et al. 2016). This conformation of NTD is believed to represent one competent for catalysis.

Substantial changes in CTD orientation around the hinge connecting the coiled-coil domain to CTD (hinge C, Fig. 3) have not been observed except for the target-bound di-RNP from Pf (discussed below). In previously determined crystal structures of archaeal Nop5 including both active and inactive RNPs, the CTD domain superimposes well, suggesting a relative stability of the CTD domain (Fig. 4). The CTD domain is the hallmark of the Nop family of proteins that include other RNA binding proteins such as the splicing factor Prp31 in U4 snoRNP (Watkins et al. 2000; Liu et al. 2007). The NOP domain is responsible for binding the SNU13/Snu13p/L7Ae-stabilized box C/D (C/D′) motif, thereby bringing RNA and the catalytic subunit fibrillarin into one place.

**Mono- or di-RNPs**

The most confounding variation in box C/D s(no)RNP assembly is the observation of both the mono- and di-RNPs (Figs. 2, 5). The mono-RNP assembly contains exactly one copy of RNA but two (or the heterodimer of Nop56 and Nop58) copies of each protein subunit, whereas the di-RNPs contain two of each component. All crystal structures of archaeal box C/D RNPs are mono-RNPs, including that of a fully assembled active Ss box C/D RNP, and so is the recently determined U3 snoRNP as the core complex of the processome (Figs. 2, 5; Xue et al. 2010; Lin et al. 2011; Komprobst et al. 2016; Chaker-Margot et al. 2017; Sun et al. 2017). However, among the mono-RNP structures, the NTD of Nop55/Nop56p/Nop58p is observed in multiple positions that were believed to be a combined effect of crystal packing, RNA constructs used in crystallization, and the different native species where the RNPs were from. These structural studies seem to suggest that the functional form of the methylation box C/D RNP is of mono-RNP. Note that the U3 snoRNP does not catalyze pre-RNA methylation. The first surprising deviation from the mono-RNP quaternary structure came when the in vitro reconstituted Methanocaldococcus jannaschii (Mj) box C/D RNPs, with wild-type box C/D RNAs, revealed a pair of face-to-face mono-RNPs,
or a di-RNP architecture under an electron microscope (Bleichert et al. 2009) that are later refined to a higher resolution (Yip et al. 2016). In the di-RNP, two box C/D RNAs are each bound across the two sets of box C/D proteins (Figs. 2, 5). This architecture extends to a range of distant-related archaeal box C/D RNPs including that from Ss (Bower-Phipps et al. 2012). Analysis of the di-RNP of a higher resolution revealed a dependence of the box C/D RNA stem in di-RNP formation (Yip et al. 2016), which may explain why some box C/D s(no)RNPs can or cannot form diRNP. Interestingly, the crystal structure of Pf mono-RNP that is considered inactive due to the well-separated distance between the target RNA and fibrillarin can be docked into the di-RNP EM electron density with minor adjustments. In the hybrid crystal and EM Pf di-RNP model, the fibrillarin located in one half of the di-RNP is positioned to methylate the target RNA bound with the other half of the di-RNP (Xue et al. 2010). This suggests that the conformation observed in the Pf box C/D mono-RNP may represent a functional state for di-RNP assembly and catalysis.

More support for di-RNP assembly came from nuclear magnetic resonance (NMR) and small angle scattering studies of in vitro assembled Pf box C/D RNPs bound with or without target RNA (Lapinaite et al. 2013). NMR measurements with $^{13}$C-labeled cofactor or fibrillarin (Lapinaite et al. 2013; Graziadei et al. 2016) revealed an asymmetry in binding and methylation of the two substrate RNA, arguing for a different enzyme-bound environment between them. The substrate complementary to the D-adjacent guide is methylated by the box C/D RNP only if the D′-adjacent guided RNA is bound. Small angle X-ray and neutron scattering (SAXS and SANS) studies of the Pf box C/D RNP with and without the two substrate RNAs revealed an estimated molecular weight consistent with that of the di-RNP. Interestingly, the conformation of the SAXS/SANS-derived substrate-bound di-RNP model differs drastically from that of the apo RNP observed by either cryoEM or SAXS/SANS but is consistent with the observed asymmetry in activity (Figs. 2, 5; Lapinaite et al. 2013). Note that the asymmetry in methylation efficiency between the D and D′ substrates may be a result of specific guide sequences between the two substrates, rather than a difference in environment, a condition generally true for box C/D RNPs. Due to limited resolution of SAXS/SANS methods, structures of the substrate-bound di-RNP at an atomic resolution remain unavailable.

Recently, a yeast box C/D snoRNP structure as part of the pre-ribosome 90S processome became available through cryoEM studies by three different groups (Fig. 5; Barandun et al. 2017; Cheng et al. 2017; Sun et al. 2017). In the processome, the U3 snoRNP forms a mono-RNP but is surrounded by many processome factors and the precursor rRNA (pre-rRNA). Whereas the structure of U3 snoRNP is clearly mono-RNP, the fact that it is stabilized extensively by processome factors and pre-rRNA suggests a propensity of the mono-RNP form of the box C/D snoRNPs in requiring binding partners. In the absence of binding partners such as in the nonprocessing snoRNPs,

**FIGURE 5.** Comparison of currently known box C/D RNP structures. All box C/D RNP structures are superimposed similarly as in Figure 4 and displayed separately in the same orientation. The Pf apo box C/D di-RNP was a model based on cryoEM density map EMD-8146 (Yip et al. 2016). The colors used are the same in all figures. Fibrillarin is colored in yellow, Nop5/56p/58p in teal, RNA box CD RNA in red, and substrate RNA (if present) in black. The relationship between any pair of structures, if known, is indicated.
how box C/D snoRNPs are assembled is completely unknown at this time.

Regardless of the quaternary structure of the box C/D snoRNPs, the structure of Nop5(56/58) CTD appears to be well conserved. Superimposed CTD of all available Nop56/58 structures revealed only small changes in the conformation around the hinge C between the yeast and archaeal Nop proteins (Figs. 3, 4). In contrast, the conformation around the hinge N remains the most diverse (Figs. 3, 4). This comparison reaffirms that the hinge C is relatively stable whereas the hinge N is not. The fact that the box C/D core in the yeast U3 snoRNP does not match the overall conformation of the active Ss box C/D RNP or that of the Pf box C/D di-RNP is consistent with the idea that snoRNPs of different functions could have different conformation/oligomerization.

Despite the overwhelming evidence for the existence of both mono- and di-RNPs, there has not been a case where both di- and mono-RNPs are captured for a single box C/D RNP of the same composition. For instance, the crystal structure of the active Ss box C/D mono-RNP was obtained when a synthetic box C/D and a target RNA with two symmetric halves were used (Lin et al. 2011), whereas its di-RNP form was observed when the target RNA was not present and with native box C/D RNA (Bower-Phipps et al. 2012). Similarly, the crystal structure of the monomeric Pf box C/D RNP contains a half-mer box C/D RNA bound with a target RNA (Xue et al. 2010), whereas its di-RNP form was observed with a native box C/D RNA without a bound target RNA (Lapinaite et al. 2013). Furthermore, addition of the target RNA did change the conformation but not the oligomeric state (Lapinaite et al. 2013). These results seem to suggest that each box C/D RNP adopts one oligomeric state suitable for its function without switching between the two.

WHY DO BOX C/D RNPs HAVE MULTIPLE CONFORMATIONS?

Nop5/56/58 and box C/D RNA are the major components responsible for the flexibility in box C/D snoRNP assembly. Changes in Nop5/56/58 conformation or in box C/D RNA could result in box C/D RNPs adopting different conformation or oligomeric states. The box C/D snoRNPs can have mono-, di-, RNP, or processome-like assemblies that vary in fibrillarina placements. Further adding to the discussion of snoRNP structural diversity is regulation by chemical modifications in box C/D snoRNAs. Many human box C/D motif sequences contain N6-methylation of adenine, which was recently shown to disrupt K-turn formation and thus 15.5 kD binding (Huang et al. 2017). Depending on the oligomeric state of individual snoRNPs, N6-methylation-facilitated regulation may have different effects. The question remains if the diverse box C/D snoRNP assembly is merely a result of experimental conditions or a reflection of functional diversity. We propose that the intrinsic flexibilities in box C/D RNA and proteins contribute to the assembly diversity and that the different assemblies serve to stabilize the RNPs, regulate methylation activity, bind different protein partners, or act as molecular scaffolds.

The collective efforts of several groups provided solid evidence for the existence of both mono- and di-RNPs. Rather than designating one as the functional form of the box C/D snoRNPs, we suggest that both exist because, together, they better account for the observed functional diversity. It is difficult to imagine that the unique snoRNPs assisting in acetylation, splicing, or leukemogenesis all assemble similarly as that of U3 snoRNP or that responsible for 2’-O-methylation. By regulating the box C/D RNA stem or by changing Nop56/58 conformation, each snoRNP may form its function-specific assembly. For instance, though the work of Yang and coworkers elegantly demonstrated that the in vitro assembled Ss mono-box C/D RNP contains a tight and narrow target-binding channel allowing for a constant pairing of 10-nt duplex (Yang et al. 2016), the work by van Nues and Watkins showed that the C’/D’ motifs in yeast snR13, snR48, and U18 could guide methylation of two nucleotide positions of the target RNA, suggesting a structural plasticity in snoRNPs especially around the C’/D’ motif (van Nues and Watkins 2017). However, it remains to be shown experimentally how these snoRNPs accommodate the plasticity.

Though box C/D snoRNPs are typically associated with ribosome biogenesis, as many as half of currently known human box C/D snoRNPs have no known functions. These are called orphan SNORDs for which no target RNA could be identified (Soeno et al. 2010; Falaleeva et al. 2017). Significantly, several such snoRNPs are shown to be associated with genetic diseases in human. For instance, loss of expression of the 15q11.2-q13.1 region consisting of four protein-coding genes and six orphan SNORDs leads to Prader-Willi syndrome (PWS) (Cavaillé et al. 2000; DuBose et al. 2012). Others have been implicated in cancer (Williams and Farzaneh 2012; Gong et al. 2017; Zhou et al. 2017). The function of these SNORDs remains unknown but their ability to form the conserved box C/D motifs raises the possibility that they could be assembled as either similarly or yet other diverse classes of box C/D RNPs. The structural diversity of the box C/D RNPs may thus serve their multifaceted physiological roles in cells. Allocation of the di- versus mono-RNP or other assemblies with varying Nop5/56/58 and box C/D RNP structures could be the molecular basis that underlies the diverse function of box C/D RNPs.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grants R01 GM124622 and R01 GM099604 to H.L.
REFERENCES

Aitaleb M, Rashid R, Chen Q, Palmer JR, Daniels CJ, Li H. 2003. Structure and function of archaeal box C/D snRNP core proteins. Nat Struct Biol 10: 256–263.

Allmang C, Henry Y, Morrissey JP, Wood H, Petfalski E, Tollervey D. 1996. Processing of the yeast pre-rRNA at sites A2 and A3 is linked. RNA 2: 63–73.

Barandun J, Chaker-Margot M, Hunziker M, Molloy KR, Chait BT, Klinge S. 2017. The complete structure of the small-subunit processome. Nat Struct Mol Biol 24: 944–953.

Beltrame M, Tollervey D. 1995. Base pairing between U3 and the pre-ribosomal RNA is required for 18S rRNA synthesis. EMBO J 14: 4350–4356.

Bleichert F, Gagnon KT, Brown BA 2nd, Maxwell ES, Leschziner AE, Wang HW, Baserga SJ. 2012. The box C/D snRNP dimeric architecture is conserved across domain Archaea. RNA 18: 1527–1540.

Cahill NM, Friend K, Speckmann W, Li ZH, Terns RM, Terns MP, Steitz JA. 2002. Site-specific cross-linking analyses reveal an asymmetric protein distribution for a box C/D snRNP. EMBO J 21: 3816–3828.

Cavalli J, Buiting K, Kieffmann M, Lalande M, Brannan CI, Hors-themke B, Bacherlere JP, Brosius J, Huttenhofer A. 2000. Identification of brain-specific and imprinted small nuclear RNA genes exhibiting an unusual genomic organization. Proc Natl Acad Sci USA 97: 14311–14316.

Chao JA, Prasad GS, White SA, Stout CD, Williamson JR. 2003. Inherent protein structural flexibility at the RNA-binding interface of L30e. J Mol Biol 326: 999–1004.

Cheng X, Roberts RJ. 2001. AdoMet-dependent methyltransferase, DNA methyltransferases and base flipping. Nucleic Acids Res 29: 3784–3795.

Cheng J, Kellner N, Bernhausen O, Hurt E, Beckmann R. 2017. 3.2-Å-resolution structure of the 90S preribosome before A1 pre-rRNA cleavage. Nature Struct Mol Biol 24: 954–964.

Decatur WA, Fournier MJ. 2003. RNA-guided nucleotide modification of ribosomal and other RNAs. J Biol Chem 278: 695–698.

Dez C, Tollervey D. 2004. Ribosome synthesis meets the cell cycle. Curr Opin Microbiol 7: 631–637.

Drago F, Gallagher JE, Compagnone-Post PA, Mitchell BM, Ponwan-cher KA, Wehner KA, Wormsley S, Settiage RE, Shabanowit J, Osheim Y, et al. 2002. A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. Nature 417: 967–970.

DuBose AJ, Smith EY, Johnstone KA, Resnick JL. 2012. Temporal and developmental requirements for the Prader-Willi imprinting center. Proc Natl Acad Sci USA 109: 3446–3450.

Dunbar DA, Baserga SJ. 1998. The U14 snRNA is required for 2'-O-methylation of the pre-18S rRNA in Xenopus oocytes. RNA 4: 195–204.

Erales J, Marchand V, Panthu B, Gillot S, Belin S, Ghayad SE, Garcia M, Laforêt F, Marcel V, Baudin-Baillieu A, et al. 2017. Evidence for rRNA 2'-O-methylation plasticity: control of intrinsic translational capabilities of human ribosomes. Proc Natl Acad Sci USA 114: 12934–12939.
Nop domain to a composite RNA-protein platform in U4 snRNP. Science 316: 115–120.

Lowe TM, Eddy SR. 1999. A computational screen for methylation guide snoRNAs in yeast. Science 283: 1168–1171.

Martin JL, McMillan FM. 2002. SAM (dependent) IAM: the S-adenosylmethionine-dependent methyltransferase fold. Curr Opin Struct Biol 12: 783–793.

Massenet S, Bertrand E, Verheggen C. 2017. Assembly and trafficking of box C/D and H/ACA snoRNPs. RNA Biol 14: 680–692.

Matera AG, Terns RM, Terns MP. 2007. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. Nat Rev Mol Cell Biol 8: 209–220.

Maxwell ES, Fournier MJ. 1995. The small nucleolar RNAs. Annu Rev Biochem 64: 897–934.

Moore T, Zhang Y, Fenley MO, Li H. 2004. Molecular basis of box C/D RNA-protein interactions; cocystal structure of archaeal L7Ae and a box C/D RNA. Structure 12: 807–818.

Nottrott S, Hartmuth K, Fabrizio P, Urlaub H, Vidovic I, Ficner R, Lührmann R. 1999. Functional interaction of a novel 15.5kD [U4/U6·U5] tri-snRNP protein with the 5′ stem–loop of U4 snRNA. EMBO J 18: 6119–6133.

Oruganti S, Zhang Y, Li H. 2005. Structural comparison of yeast snoRNP and splicingosomal protein Snu13p with its homologs. Biochem Biophys Res Commun 333: 550–554.

Oruganti S, Zhang Y, Li H, Robinson H, Terns MP, Terns RM, Yang W, Li H. 2007. Alternative conformations of the archaeal Nop56/58-Mth2 ribonuclein complex imply flexibility in box C/D R Nep. J Mol Biol 371: 1114–1150.

Phipps KR, Charette J, Baserga SJ. 2011. The small subunit processes in ribosome biogenesis—progress and prospects. Wiley Interdiscip Rev RNA 2: 1–21.

Rashid R, Attaleab M, Chen Q, Spiegel K, Demeler B, Li H. 2003. Functional requirement for symmetric assembly of archaeal box C/D small ribonucleoprotein particles. J Mol Biol 333: 295–306.

Schultz A, Nottrott S, Hartmuth K, Lührmann R. 2006a. RNA structural requirements for the association of the spliceosomal hPp31 protein with the U4 and U4atac small nuclear ribonucleoproteins. J Biol Chem 281: 28278–28286.

Schultz A, Nottrott S, Watkins NJ, Lührmann R. 2006b. Protein–protein and protein–RNA contacts both contribute to the 15.5K-mediated assembly of the U4/U6 snRNP and the box C/D snoRNPs. Mol Cell Biol 26: 5146–5154.

Sharma S, Yang J, van Nues R, Watzinger P, Kött P, Lafontaine D, Grimm, S, Entian KD. 2017. Specialized box C/D snoRNPs act as antisense guides to target RNA base acetylation. PLoS Genet 13: e1006804.

Sloan KE, Leisegang MS, Doebele C, Ramirez AS, Simm S, Saferth C, Kretschmer J, Schorge T, Markouts S, Haag, S, et al. 2015. The association of late-acting snoRNPs with human preribosomal complexes requires the RNA helicase DDX21. Nucleic Acids Res 43: 553–564.

Soeno Y, Taya Y, Stasyk T, Huber LA, Aoba T, Hüttenhofer A. 2010. Identification of novel ribonucleoprotein complexes from the brain-specific snoRNA MBII-52. RNA 16: 1293–1300.

Sun Q, Zhu X, Qi J, An W, Lan P, Tan D, Chen R, Wang B, Zheng S, Zhang C, et al. 2017. Molecular architecture of the 90S small subunit pre-ribosome. Elife 6: e22086.

Szwczuk LB, DeGregorio SJ, Strobel SA, Steitz JA. 2002. Exclusive interaction of the 15.5kD protein with the terminal box C/D motif of a methylation guide snoRNP. Chem Biol 9: 1095–1107.

Thomson E, Ferreira-Cerca S, Hurt E. 2013. Eukaryotic ribosome biogenesis at a glance. J Cell Sci 126: 4815–4821.

Tomkuviené M, Ličytė J, Olendraitė I, Liutkiviciūtė Z, Clouet-d’Orval B, Klimašauskas S. 2017. Archaeal fibrillarin-Nop5P heterodimer 2–O-methylates RNA independently of the C/D guide RNP particle. RNA 23: 1329–1337.

Tumer B, Melcher SE, Wilson TJ, Norman DG, Lilley DM. 2005. Induced fit of RNA on binding the L7Ae protein to the kink-turn motif. RNA 11: 1192–1200.

van Nues RW, Watkins NJ. 2017. Unusual C/D’ motifs enable box C/D snoRNPs to modify multiple sites in the same rRNA target region. Nucleic Acids Res 45: 2016–2028.

van Nues RW, Granneman S, Kudla G, Sloan KE, Chicken M, Tollervey D, Watkins NJ. 2011. Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing. EMBO J 30: 2420–2430.

Vidovic I, Nottrott S, Hartmuth K, Lührmann R, Ficner R. 2000. Crystal structure of the spliceosomal 15.5 kD protein bound to a U4 snRNA fragment. Mol Cell 6: 1331–1342.

Watkins NJ, Bohnsack MT. 2012. The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. Wiley Interdiscip Rev RNA 3: 397–414.

Watkins NJ, Segault V, Charpentier B, Nottrott S, Fabrizio P, Bach, A, Wilm M, Rosbash M, Branlant C, Lührmann R. 2000. A common core RNP structure shared between the small nucleolar box C/D RNPs and the spliceosomal U4 snRNP. Cell 103: 457–466.

Williams GT, Farzaneh F. 2012. Are snoRNAs and snoRNA host genes new players in cancer? Nat Rev Cancer 12: 84–88.

Woolford JL Jr, Baserga SJ. 2013. Ribosome biogenesis in the yeast Saccharomyces cerevisiae. Genetics 195: 643–681.

Xue S, Wang R, Yang F, Terns RM, Terns MP, Zhang X, Maxwell ES, Li H. 2010. Structural basis for substrate placement by an archaeal box C/D ribonucleoprotein particle. Mol Cell 39: 939–949.

Yang Z, Lin J, Ye K. 2016. Box C/D guide RNAs recognize a maximum of 10 nt of substrates. Proc Natl Acad Sci 113: 10878–10883.

Ye K, Jia R, Lin J, Ju M, Peng J, Xu A, Zhang L. 2009. Structural organization of box C/D RNA-guided RNA methyltransferase. Proc Natl Acad Sci 106: 13808–13813.

Yip WS, Shigematsu H, Taylor DW, Baserga SJ. 2016. Box C/D snoRNPs stem ends act as stabilizing anchors for box C/D di-sRNPs. Nucleic Acids Res 44: 8976–8989.

Yu YT, Meier UT. 2014. RNA-guided isomerization of uridine to pseudouridine—pseudouridylation. RNA Biol 11: 1483–1494.

Zhou F, Liu Y, Rohde C, Pauli C, Geroff D, Köhn M, Misiak D, Bäumer N, Cui C, Gällner S, et al. 2017. AML1-ETO requires enhanced C/D box snoRNA/RNP formation to induce self-renewal and leukaemia. Nat Cell Biol 19: 844–855.