High-resolution structure of eukaryotic Fibrillarin interacting with Nop56 N-terminal domain

Simone Höfler[1], Peer Lukat[3], Wulf Blankenfeldt[3], Teresa Carlomagno[1,2,*]

[1]Leibniz University Hannover, Institute for Organic Chemistry and Centre for Biomolecular Drug Research (BMWZ) – Schneiderberg 38 – D-30167 Hannover – Germany

[2]Helmholtz Centre for Infection Research, Group of NMR-based Structural Chemistry – Inhoffenstrasse 7 – D-38124 Braunschweig – Germany

[3]Helmholtz Centre for Infection Research, Department of Structure and Function of Proteins – Inhoffenstrasse 7 – D-38124 Braunschweig – Germany

* Correspondence to: teresa.carlomagno@oci.uni-hannover.de

ABSTRACT

Ribosomal RNA (rRNA) carries extensive 2'-O-methyl marks at functionally important sites. This simple chemical modification is thought to confer stability, promote RNA folding and contribute to generate a heterogenous ribosome population with a yet-uncharacterized function. 2'-O-methylation occurs both in archaea and eukaryotes and is accomplished by the Box C/D RNP enzyme in an RNA-guided manner. Extensive and partially conflicting structural information exists for the archaeal enzyme, while no structural data is available for the eukaryotic enzyme. The yeast Box C/D RNP consists of a guide RNA, the RNA-primary binding protein Snu13, the two scaffold proteins Nop56 and Nop58 and the enzymatic module Nop1. Here we present the high-resolution structure of the eukaryotic Box C/D methyltransferase Nop1 from *Saccharomyces cerevisiae* bound to the N-terminal domain of Nop56. We discuss similarities and differences between the interaction modes of the two proteins in archaea and eukaryotes and demonstrate that eukaryotic Nop56 recruits the methyltransferase to the Box C/D RNP through a protein–protein interface that differs substantially from the archaeal orthologues. This study represents a first achievement in understanding the evolution of the structure and function of these proteins from archaea to eukaryotes.

INTRODUCTION

Methylation of the 2'-O-ribose position (2'-O-Me) is the simplest and second most abundant modification of ribosomal RNA (rRNA) (Ayadi et al. 2019). Like many other rRNA modifications, 2'-O-Me marks are transferred to the RNA co-transcriptionally and post-transcriptionally in the nucleolus and are important for ribosome biogenesis and translational accuracy (Baudin-Baillieu et al. 2009; Kimura and Suzuki 2010; Siibak and Remme 2010; Arai et al. 2015; Ishiguro et al.; Sloan et al. 2017). 2'-O-Me has been shown to promote RNA folding and increase chemical stability, thereby influencing the association of ribosomal proteins in a space- and/or time-dependent manner (Green and Noller 1996; Polikanov et al. 2015; Arai et al. 2015; Natchiar et al. 2017; Ishiguro et al.).

In archaea and eukaryotes, this modification is catalyzed by multi-component ribonucleoprotein (RNP) complexes. The RNPs are assembled around guide RNAs named box C/D s/snoRNAs (archaea/eukaryotes) after their distinct sequence motifs box C (5'-RUGAUGA) and box D (5'-CUGA). The box C/D motif folds in a characteristic structure termed K-turn (Klein 2001; Špačková et al. 2010; Shi et al. 2016). L7Ae or Snu13, in archaea and yeast, respectively, bind to the K-turn motif of the guide RNA. A similar, but less conserved, box C'/D' motif is present in all archaeal guide RNAs and has been proposed to exist in eukaryotic guide RNAs as well (van Nues et al. 2011; Qu et al. 2011). This motif forms a characteristic structure named K-loop (Nolivos et al. 2005). A second copy of L7Ae binds the box C'/D' motif in archaea; in contrast, Snu13 has been found unable to recognize putative box C'/D' motifs in yeast (Cahill et al. 2002). The substrate RNA is recruited by base-pairing with complementary sequences, located upstream of box D and D', respectively (Tran et al. 2003, 2005; Appel and Maxwell 2007). The guide RNA–L7Ae/Snu13 interaction mediates the recruitment of the scaffolding proteins Nop5 (archaea) or Nop56 and Nop58 (eukaryotes). In archaea the Nop5 protein enters the complex as a homodimer; in eukaryotes $Nop5_2$ is substituted by the Nop56–Nop58 heterodimer (Figure 1A, B). The SAM-dependent methyltransferase Fibrillarin (Fib, in archaea and *H. sapiens*) or Nop1 (in fungi) is integrated

into the Box C/D RNP complex via interaction with the N-terminal domain of Nop5 (archaea) or Nop56 and Nop58 (eukaryotes) (Tollervey et al. 1991; Lyman et al. 1999; McKeegan et al. 2009; Lechertier et al. 2009a; Ye et al. 2009; Quinternet et al. 2015; Barandun et al. 2017; Paul et al. 2019). Two copies of the methyltransferase are recruited to the RNP by the Nop protein dimers.

Components of the Box C/D RNP enzymes have been linked to disease phenotypes (Dimitrova et al. 2019). In particular, the methyltransferase Fibrillarin, which shows a high degree of conservation in most domains of life, has been associated with autoimmune diseases and cancer: the protein is frequently overexpressed in tumor cells and has been related to a poor survival rate in breast cancer patients (Marcel et al. 2013). Thus, understanding the functional mechanism and regulation of eukaryotic snoRNPs is important both to elucidate the biology of the cell and in a disease context.

The structure-function relationships of the archaeal Box C/D sRNPs have been extensively studied in the past years (Ye et al. 2009; Lin et al. 2011; Lapinaite et al. 2013; Graziadei et al. 2016; Yang et al. 2016; Yu et al. 2018; Graziadei et al. 2020), leading to two structural models of the enzyme loaded with the substrate RNAs (Figure 1A). In the first model, the enzyme is a mono-RNP and comprises one copy of guide RNA, two substrate RNAs, and two copies of each protein (L7Ae, Nop5, and Fib) (Lin et al. 2011). In the second model, the enzyme is a di-RNP and comprises two copies of guide RNA, four copies of substrate RNAs, and four copies of each protein (Lapinaite et al. 2013). The oligomerization state of the enzyme has been shown to depend on the sequence of the guide RNA and has consequences on the regulation of methylation efficiency at the different substrate sites (Graziadei et al. 2020; Yip et al. 2016).

Conversely, little is known about protein–protein and RNA–protein interactions in the eukaryotic Box C/D snoRNP at high resolution, as the *in vitro* reconstruction of functional eukaryotic Box C/D snoRNPs has so far yielded complexes of heterogenous composition (Peng et al. 2014; Yang et al. 2020). Here, we determine the atomic details of the complex of

the RNA 2'-O-methyltransferase Nop1 with the Nop56 N-terminal domain from *S. cerevisiae* (*Sc*). We find that the interaction interface between the two proteins has significantly evolved from archaea to eukaryotes. As a result, archaeal and eukaryotic proteins cannot complement each other. Our high-resolution structure reveals the key interaction features of Nop1 and Nop56 within the eukaryotic complex and suggests that evolutionary pressure has caused structural and potentially functional divergence of the Nop56 and Nop58 N-terminal domains from the archaeal Nop5 counterparts.

RESULTS

Archaeal RNA 2'-O-methyltransferase cannot complement its eukaryotic ortholog in yeast. In this study we used proteins from *S. cerevisiae* as representatives of eukaryotic snoRNPs. The primary sequences of Nop1 and Nop56 share an overall 73.9% and 58.5% similarity with their respective human orthologs (Supplementary Figures 1 and 2). When compared to the archaeal Fibrillarin, Nop1 has an additional 80 amino acids long, RGG-rich, N-terminal domain, shared by all eukaryotic 2'-O-methyltransferases; this domain is predicted to be disordered, may play a role in nuclear localization and is likely involved in RNA binding (Chong et al. 2018; Smith et al. 2020). Beyond this RGG-rich domain, the primary sequence of the catalytic domain of *P. furiosus* (*Pf*) Fibrillarin has 60.9% and 47.9% similarity with yeast Nop1 and human Fibrillarin, respectively (Figure 2A and Supplementary Figure 1). On the other hand, the primary sequence of the N-terminal domain of yeast Nop56 shares only 31.3% similarity and 17.6% identity with the archaeal ortholog (Figure 2B and Supplementary Figure 2), posing the question as to whether the interaction mode of the eukaryotic 2'-O-methyltransferase with the N-terminal domains of the Nop56 and Nop58 proteins is similar to that of archaeal Fibrillarin with the Nop5 N-terminal domain.

To answer this question, we first tested whether the archaeal and eukaryotic methyltransferases could complement each other. We used recombinant Nop1$^{83-327}$ (lacking the N-terminal RGG-rich domain), Nop56$^{1-166}$ (the N-terminal domain of Nop56), *Pf* Fibrillarin

and *Pf* Nop5$^{1-123}$ (the N-terminal domain of Nop5) and analyzed their ability to form stable complexes with each other using size-exclusion chromatography. We found that the N-terminal domain of *Pf* Nop5 (Figure 3A and 3C) cannot form a stable complex with Nop1$^{83-327}$. Similarly, archaeal Fibrillarin does not form a stable complex with the N-terminal domain of Nop56 (Figure 3B and 3C). Control experiments showed that *Sc* Nop1$^{83-317}$ and Nop56$^{1-166}$ as well as *Pf* Fibrillarin and Nop5 form stable complexes (Figure 3D). We conclude that the interaction mode between the methyltransferase and the N-terminal domains of the scaffold proteins Nop5, Nop56, and Nop58 have evolved to become specific for archaea and eukaryotes. Thus, we set out to characterize the Nop56–Nop1 complex in eukaryotes at atomic resolution.

Structure of the Nop56$^{1-166}$–Nop1$^{83-327}$ complex. We used X-ray crystallography to solve the structure of *Sc* Nop1$^{83-327}$ in complex with *Sc* Nop56$^{1-166}$ to a resolution of 1.7 Å (Figure 4, Table 1). The structure of Nop1$^{83-327}$ is very similar to that of human Fibrillarin (PDB-ID: 2IPX) with an average root mean square deviation (RMSD) of 0.6 Å; a comparable degree of similarity is found also with the structures of different archaeal Fibrillarins (Aittaleb et al., 2003; Gagnon et al., 2012; Oruganti et al., 2007; Ye et al., 2009) (Figure 2A). The N-terminal lobe (comprising amino acids 83-146) consists of five anti-parallel β-sheets; the substantially larger C-terminal lobe consists of seven β-sheets sandwiched between six α-helixes and adopts the Rossmann-fold typical for S-adenosylmethionine (SAM)-dependent methyltransferases (Figure 2A and 4A) (Wang et al. 2000).

The N-terminal domain of Nop56 consists of five β-sheets, with β-sheets 1–3 (β1-3) and β-sheets 4–5 (β4–5) being in an anti-parallel and parallel orientation, respectively (Figure 2B and 4A); the β-sheets core is surrounded by seven α-helices. An insertion of 26 amino acids between β2 and β3 forms α-helix 1 (α1) and 2 (α2); this insertion is present in both eukaryotic Nop56 and Nop58 proteins but absent in archaeal Nop5 proteins (Figure 2B and Supplementary Figure 2).

Nop1$^{83-327}$ engages α3–5 and the C-terminal tail in the interaction with α1, α6 and α7, β1 and the loop region β4–α5 of Nop56$^{1-166}$ (Figure 4A). Complex formation is driven by the high charge complementarity, with Nop1 presenting an overall positive surface charge that fits snuggly the predominantly negative surface charge of the Nop56 N-terminal domain (Figure 4B).

Three main interaction regions hold Nop1 and Nop56 together (Figure 4C–E). First, Nop56$^{1-166}$ α1 and α6 contact the C-terminal tail of Nop1 (Figure 4C). Hydrogen bonds are formed between Nop1-S323 and Nop56$^{1-166}$ α1-Q35 and Nop1-R322 and Nop56$^{1-166}$ α6-D126, while Nop56$^{1-166}$ α1-K32 has a polar contact with the carbonyl of Nop1-Ser323. These interactions lock the C-terminal tail of Nop1 in a stable conformation. Additional hydrogen bonds at this site of the complex are formed between Nop1-D263 (located in a short loop between α5 and β10) and Nop56$^{1-166}$ α1-N39 as well as Nop1 α6-R297 and Nop56$^{1-166}$ α6-E122.

Second, Nop56$^{1-166}$ α7 contacts Nop1-α3, β7, and β8 (Figure 4D). Nop1-Y195, which is highly conserved in eukaryotes and archaea, forms a hydrogen bond with Nop56-Q151, the backbone of Nop1-β8 has hydrogen bonds with the side chain of Nop56-S159 and Nop1-K169 and K205 have polar contacts with the side chain of Nop56-E148 and the backbone carbonyl of Nop56-V163, respectively.

Third, an extensive network of hydrogen bonds is formed between residues of Nop1-α4 (Q228 and R231) and the preceding loop (E222) with Nop56-S101, E10, and Y158, respectively (Figure 4E). This interaction area is driven by a strong charge complementary between Nop1-R231 and a negatively charged pocket on Nop56, containing E10 and E11.

All in all, the complex between Nop1 and the Nop56 N-terminal domain is supported almost exclusively by polar interactions and hydrogen bonds.

Charge complementarity is partially conserved from archaea to eukaryotes. After solving the first high-resolution structure of a eukaryotic complex between the Nop56 N-terminal domain and the Nop1 methyltransferase, we compared the binding mode of the

eukaryotic proteins with that of archaeal proteins. We chose three representative structures of archaeal Nop5–Fib complexes in *P. furiosus (Pf)* (PDB-ID: 2NNW) (Oruganti et al. 2007), *A. fulgidus (Af)* (PDB-ID: 1NT2) (Aittaleb et al. 2003) and *S. solfataricus (Ss)* (PDB-ID: 3ID5) (Ye et al. 2009), which are members of the three evolutionarily different archaeal classes Thermococci, Archaeoglobi, and Thermoprotei, respectively (Brochier-Armanet et al. 2011).

The 2'-O-methyltransferase is quite well conserved among archaeal classes (average primary sequence similarity, 63.7%) and archaea and eukaryotes (average primary sequence similarity of archaeal Fibrillarins with *Sc* Nop1, 59.1%) (Supplementary Figure 1). The structure of the catalytic domain is also well conserved with RMSD values between the structures of the archaeal proteins and that of Nop1 ranging from 1.6 to 2.9 Å (Figure 2A). The presence of the N-terminal RGG-rich domain is the main feature differentiating the eukaryotic methyltransferase from its archaeal orthologs (Rodriguez-Corona et al. 2015). On the other hand, the eukaryotic Nop56 N-terminal domain diverges substantially from that of the archaeal Nop5 (average primary structure similarity of archaeal Nop5 N-terminal domains with *Sc* Nop56$^{1\text{-}166}$, 26.2% (Supplementary Figure 2). Structurally, the N-terminal domains of Nop5 and Nop56 proteins are also divergent: a core fold, which is conserved across *Pf*, *Ss* and *Sc*, is accompanied by two additional helices (α1 and α2) in the eukaryotic proteins (Supplementary Figure 2 and Figure 2B), one of which (α1) provides an interaction area with the methyltransferase. It is important to notice that the Nop5 N-terminal domain is not uniformly conserved across archaeal species either: in *Af*, this domain is significantly smaller than in *Pf* and *Ss*, misses three α-helixes and two β-sheets and is not able to fold in the absence of Fibrillarin *in vitro* (Aittaleb et al. 2003).

Despite these differences, the formation of the complex between the Nop5 N-terminal domain and Fibrillarin in *Pf* and *Ss* is driven by charge complementarity as well (Figure 5A–C). A predominantly negatively charged surface of Nop5 binds a predominantly positively charged surface of Fibrillarin. Notably, the charge distribution of the solvent-accessible surface of *Ss* Nop5$^{1\text{-}132}$ is more similar to that of *Sc* Nop56$^{1\text{-}166}$ than it is the one of *Pf* Nop5$^{1\text{-}}$

^{123}. Nevertheless, the *Ss* Fibrillarin surface interacting with Nop5 is considerably less charged than the corresponding surface of *Sc* Nop1. *Pf* Nop5$^{1\text{-}123}$ displays a higher number of aromatic residues on the interaction surface with Fibrillarin, limiting the accessibility to the charged residues. Thus, while the driving force for the formation of the complex is conserved from *Pf* and *Ss* to *Sc*, the details of the interactions differ both among archaeal species and between archaea and eukaryotes.

Different from all other complexes, the interacting surfaces of the *Af* proteins show little charge complementarity to each other, suggesting that complex formation is driven by hydrophobic contacts (Figure 5D). This fact may explain how the presence of Fibrillarin promotes folding of the *Af* Nop5 N-terminal domain, as the folding process requires hydrophobic interactions.

The eukaryotic Nop56–Nop1 complex lacks conserved aromatic interactions present in the archaeal orthologs. Next, we analyzed the differences in the interaction interfaces of eukaryotic and archaeal complexes in atomic details. Nop1-Y195, a conserved Tyr in β7 of the methyltransferase, forms a hydrogen bond with Nop56$^{1\text{-}166}$-α7 Q151 and stabilizes the complex via a polar interaction (Figure 6A, F). This Nop56$^{1\text{-}166}$ Gln is well conserved in eukaryotes (Figure 6G), while in archaea the same position is occupied by a hydrophobic amino acid and a hydrophobic cluster is built around the conserved Tyr. *Pf* Nop5-α6 W104, F105 and Y108 surround *Pf* Fib-Y102 (Figure 6B); *Ss* Fib-Y105 is part of a cluster comprising *Ss* Nop5-Y113 and Y114, with the hydrophobic contacts being reinforced by *Ss* Nop5-L117 (Figure 6C); even in *Af*, despite the peculiar characteristics of the *Af* Nop5–Fib complex, this hydrophobic cluster is conserved (Figure 6D). The structure of the Nop5–Fib complex of *M. janaschii* (*Mj*) is not available; nevertheless, docking of the crystal structures of Nop5 and Fib with the ClusPro 2.0 Webserver (Kozakov et al. 2013, 2017; Vajda et al. 2017) shows that two aromatic residues of *Mj* Nop5-α6 (Y96 and F99) could build a hydrophobic cluster with the conserved *Mj* Fib-Y102 (Figure 6E, H). The presence of at least two aromatic residues on Nop5-α6 is characteristic of archaeal proteins, while eukaryotic

Nop56 proteins have a polar, negatively charged residue at the central position of the aromatic cluster (Figure 6G–H).

Next, we tested whether the difference in the nature of the interactions of the conserved Tyr is crucial in determining the selectivity of the *Sc* Nop56 N-terminal domain for *Sc* Nop1 versus Pf Fibrillarin. We generated the *Sc* Nop56$^{1\text{-}166}$-L147W/E148F mutant (Nop56$^{1\text{-}166}$-WF) and tested its ability to bind Nop1$^{83\text{-}327}$ (Figure 3E) and *Pf* Fib (Figure 3F). The double mutation did not perturb the binding preferences of Nop56$^{1\text{-}166}$ suggesting that this interaction area is not crucial for the formation of a stable protein–protein complex.

Eukaryotic proteins Nop56 and Nop58 evolutionarily acquired an additional α-helix at the binding interface with Nop1.

Eukaryotic Nop56 and Nop58 proteins have a 23–26 amino acids long insertion between β2 and β3, which folds into helices α1 and α2 (Figure 7A). α1 is at the interface with Nop1 and enables Nop56 to interact with its extended C-terminal tail (Figure 7 B). In archaea, where the two additional α-helices are missing, the Fibrillarin C-terminal end cannot interact with the Nop5 N-terminal domain and remains shorter than in eukaryotes (Figure 7C–E).

The primary sequence of α1 is reasonably well conserved across eukaryotes but is slightly more hydrophobic in vertebrates than in fungi (Figure 7F). On the other hand, the amino acid composition of the C-terminal tail of Nop1/Fib diverges (Figure 7E): in vertebrates, it contains a poly-proline stretch, which makes it more rigid and hydrophobic, while in fungi the prolines are substituted by polar residues. Despite these differences, all eukaryotic proteins share a predominantly negatively charged surface on Nop56-α1, which matches the total positive charge of the Nop1/Fib C-terminal tail. Thus, the interaction between the Nop56-α1 and the Nop1/Fib C-terminal tail contributes to the overall charge complementarity of the binding interface in all eukaryotes, with vertebrates displaying a higher contribution of hydrophobic contacts.

In archaea, the interaction interface ends at Nop5-α5 (corresponding to Nop1-α6), which presents mostly polar side chains to Fibrillarin (Figure 7C–D, G); thus, the electrostatic nature of the protein–protein interactions is conserved at this site from archaea to eukaryotes.

To test whether the interaction between the extended C-terminal tail of Nop1$^{83-327}$ and Nop56-α1 is pivotal to the formation of a stable Nop56–Nop1 complex, we generated Nop1$^{83-321}$, lacking the C-terminal tail, and the Nop56$^{1-166}$-Q35A/N39A mutant. We then used size-exclusion chromatography to test the interaction between Nop1$^{83-321}$ and wild-type Nop56$^{1-166}$, as well as between Nop1$^{83-327}$ and Nop56$^{1-166}$-Q35A/N39A. We found that complex formation was not compromised by the mutations (Figure 8A, left panel, 8D). Nevertheless, the dissociation constant (K_D) of the Nop56$^{1-166}$-Q35A/N39A–Nop1$^{83-327}$ complex, measured by isothermal titration calorimetry, was one order of magnitude larger than that of the Nop56$^{1-166}$–Nop1$^{83-327}$ complex (Figure 8A, right panel, 8D and Supplementary Figure 3), confirming that the contacts between the Nop1 C-terminal tail and Nop56-α1 contribute to the binding affinity.

Nop1/Fib helix α4 is central to the formation of the complex with Nop56 and Nop5 in both eukaryotes and archaea. At the core of the Nop56–Nop1 complex, Nop1-α4 inserts between Nop56-α6 and α7, with Nop1-R231 forming a hydrogen bond with Nop56-β1 E10 (Figure 7H). In all available structures of archaeal Fibrillarin in complex with Nop5, Fib-α4 also inserts between Nop5-α5 and α6. This interaction site has the highest charge complementarity both in archaea and eukaryotes. As in *Sc*, *Pf* Fib-R138 forms a hydrogen bond with Nop5-E7 (Figure 7I). This Fib-α4 Arg is conserved among all eukaryotes but is more variable in archaea (Figure 7L). In *Ss,* the K141 found at this position does not penetrate as deep as to reach Nop5-β1 but instead forms hydrogen bonds with E66 and the backbone of S90 on Nop5-α5 (Figure 7J). Despite preserving a network of hydrogen bonds, the interaction at the same site of the *Af* Nop5–Fib complex involves a completely different set of amino acids (Figure 7K).

To verify the relevance of this interaction site for the formation of a stable Nop56–Nop1 complex, we generated the Nop1$^{83-327}$-R231E and Nop56$^{1-166}$-E10K/E11K mutants and used size-exclusion chromatography to test their ability to form a complex with wild-type Nop56$^{1-166}$ and Nop1$^{83-327}$, respectively. None of the two mutants bound the wild-type interaction partners (Figure 8B, D), confirming that the surface charge complementarity is the central driving force of complex formation between Nop1$^{83-327}$ and Nop56$^{1-166}$.

Repulsion forces between archaeal and eukaryotic proteins. After establishing the similarities and differences in the nature of the attraction forces between the 2'-O-methyltransferase and the scaffold protein of the Box C/D 2'-O-methylation enzyme in eukaryotes and archaea, we analyzed whether any repulsion force actively counteracts the formation of complexes between *Sc* Nop1 and *Pf* Nop5 or *Pf* Fib and *Sc* Nop56. We compared the structure of the *Sc* Nop56$^{1-166}$–Nop1$^{83-327}$ complex determined here with that of the *Pf* Nop5–Fib complex from (Xue et al. 2010) (PDB-ID: 3NMU) and found two major hot spots where the nature of the side-chains of the ortholog binding partner would cause either severe steric clashes or charge repulsion, thus preventing complex formation (Figure 8C). First, if Fib were substituted by *Sc* Nop1 in the structure of the *Pf* Nop5–Fib complex, E136 of *Pf* Fib-α4 would be replaced by Nop1-K229, resulting in an electrostatic repulsion with *Pf* Nop5-K40 and R122 (Figure 8C, left panel). Second, if Nop1 were substituted by *Pf* Fib in the structure of the Nop56$^{1-166}$– Nop1$^{83-327}$ complex, Nop1-D263 and M236 at the interface with Nop56$^{1-166}$-α6 would be replaced by *Pf* Fib-R170 and K143, respectively. R170 could clash with Nop56$^{1-166}$-α1, which is absent in the orthologous *Pf* Nop5 protein; K143 would introduce a positive charge that could cause repulsion with Nop56$^{1-166}$-R129 (Figure 8C, middle panel), while *Pf* Nop5 carries a negatively charged Glu at this position.

To test whether these two areas of predicted clashes and repulsion contribute to the incompatibility between archaeal and eukaryotic proteins, we generated three mutants. First, we engineered the Nop56$^{1-166}$-E71K/Y158K to mimic the predicted effect of Nop5-K40 and R122 and tested its interaction with wild-type Nop1$^{83-327}$ via size-exclusion chromatography.

The Nop56$^{1-166}$-E71K/Y158K mutant did not form a stable complex with Nop1$^{83-327}$ (Figure 8C, right panel, 8D), confirming that *Pf* Nop5-K40 and R122 have a disruptive effect on the formation of the complex between *Pf* Nop5$^{1-123}$ and *Sc* Nop1$^{83-327}$.To enhance charge repulsion, we also generated the Nop1$^{83-327}$-E222K mutant, which, as predicted, failed to form a complex with Nop56$^{1-166}$-E71K/Y158K (Figure 8C, right panel, 8D). Last, we generated the Nop1$^{83-327}$-D263R/M236K mutant, to mimic the predicted effect of *Pf* Fib-R170 and K143, and tested its ability to form a complex with wild-type Nop56$^{1-166}$. Nop1$^{83-327}$-D262R/M236K did not bind wild-type Nop56$^{1-166}$, confirming the disruptive effect of *Pf* Fib-R170 and K143 on the formation of the complex between *Sc* Nop56$^{1-166}$ and *Pf* Fib (Figure 8C, right panel, 8D).

In conclusion, we identified and confirmed two interface regions where the nature of the amino acid side chain actively disturbs the formation of a cross-species complex and subsequently cause the incompatibility of Nop1$^{83-327}$ with *Pf* Nop5$^{1-123}$ and of *Pf* Fib with Nop56$^{1-166}$.

DISCUSSION

The ability of Nop1 to form a stable complex with the N-terminal domains of Nop56 and Nop58 is essential for its recruitment to the Box C/D snoRNP enzyme. Nevertheless, while the C-terminal domain of the Nop56 and Nop58 proteins is highly conserved across species and even from archaea to eukaryotes, the N-terminal domain of the same proteins is highly variable. Similarly, the methyltransferase is very well conserved in its catalytic site and RNA binding motif (85.7% and 79.5% average similarity between Nop1 and archaeal Fibrillarins, calculated for residues in a range of 4 Å from either the cofactor and substrate-ribose or the RNA, respectively) but more variable in the Nop5 or Nop56/Nop58 interacting surface (53.7% average similarity between Nop1 and archaeal Fibrillarins, calculated for residues in a range of 4 Å from the binding partner). As a result, the proteins from *S. cerevisiae* cannot bind the proteins of *P. furiosus*.

The variability of Fibrillarin in regions other than the catalytic site and the RNA binding surface may be a consequence of the additional roles acquired by the protein in eukaryotes with respect to archaea. In *S. cerevisiae* and humans, Fibrillarin methylates Q105 and Q104 of histone H2A, respectively (Iyer-Bierhoff et al. 2018; Tessarz et al. 2014). In this role, the binding partners of Fibrillarin are unknown; however, they are likely to engage the same surface as the Nop56/Nop58 N-terminal domains. Fibrillarin has also been found to interact with nucleophosmin (NPM1) in an RNA-independent manner (Nachmani et al. 2019). Also, in this case the structural details of the interaction are unknown.

Similarly, the Nop56 and Nop58 proteins have acquired a number of additional binding partners in eukaryotes with respect to archaea. While some of these interactions have been found to engage the C-terminal domain (such as the AAA+ ATPase R2TP complex recruiting the C-terminal unstructured tail of Nop58 (Yu et al. 2019)), the binding mode of many other partners is unknown, as for example that of the Drosophila protein hoip, which binds both Nop56 and Nop58 (Murata et al. 2008), or of the component of the R2TP complex Nop17, which binds Nop58 during the assembly of the Box C/D snoRNP (Prieto et al. 2015). In addition, Nop56 has been suggested to act in Burkitt's lymphoma associated with c-Myc mutations (Cowling et al. 2014) through a yet-uncharacterized mechanism. Thus, both Fibrillarin and the N-terminal domains of Nop56 and Nop58 may have coevolved to ensure an efficient mutual interaction while still supporting the binding to other partners.

To date, the only structural information available for an eukaryotic snoRNP has been obtained by cryogenic electron microscopy for a complex assembled around snoRNA U3, which is not a methylation guide RNA (Barandun et al. 2017; Cheng et al. 2019). In this complex two copies of Snu13 bind to two kink-turn motifs in snoRNA U3; conversely, guide RNAs involved in 2'-O-Me contain only one K-turn motif and bind only one copy of Snu13. The architecture of the U3 snoRNP is similar to that of archaeal Box C/D mono-RNPs (Lin et al. 2011); however, it remains unclear whether this architecture can be assumed also for methylation-competent Box C/D snoRNPs. In the U3 snoRNP structure, the N-terminal

domains of both Nop56 and Nop58 bind the SAM-dependent methyltransferase Nop1: yet, due to its limited resolution, the structure gives only incomplete information on the binding interfaces. Here we solve the first high-resolution structure of the Nop56–Nop1 complex of *S. cerevisiae* and reveal the atomic details of the protein–protein interaction surface.

When comparing our high-resolution structure of the *Sc* Nop56$^{1-166}$–Nop1$^{83-327}$ complex with the 3.8 Å-resolution cryo-EM structure of the same protein complex integrated in the U3 snoRNP bound to the 90S pre-ribosome (Barandun et al. 2017) (PDB-ID: 5WLC), we observe that Nop1$^{83-327}$ keeps its conformation in the multicomponent complex (Figure 9A), despite engaging in interactions with several other proteins, such as Utp11, Fcf2, Sas10, Utp24 and Bud21 (Barandun et al. 2017; Cheng et al. 2019). The average backbone root mean square deviation (RMSD) of the two Nop1$^{83-327}$ structures is 1.22 Å. In contrast, the N-terminal domain of Nop56 undergoes conformational changes upon interaction with the other binding partners in the 90S pre-ribosome, resulting in an average backbone RMSD of 2.5 Å in the Nop56$^{1-166}$ structures. The conformational changes are localized to the residues comprised between β2 and β3 (region 1), which form α-helices 1 and 2, and the Lys-rich loop between α4 and β4 (region 2) (Figure 9A). Both regions are eukaryotic-specific extensions of Nop56 and are absent in archaeal Nop5. The conformational changes occurring in Nop56-α1 and α2 result in an altered surface charge distribution (Figure 9B) and are correlated with the recognition of the protein binding partner Sof1 (Figure 9C, E, F, Supplementary Figure 4A); similarly, the Lys-rich loop rearranges to interact with the 5' ETS RNA in the 90S pre-ribosome (Figure 9D, G, H).

In the high-resolution structure of the Nop56$^{1-166}$–Nop1$^{83-327}$ complex, the amino acid sequence between Nop56$^{1-166}$-β2 and β3 adopts a conformation that optimizes intramolecular interactions (Figure 9E) with hydrophobic residues L22, I27, R30 and L31 oriented towards to core of the protein and hydrophilic or charged residues, such as Q23, D25, D26 and S29, exposed to the solvent. Upon interaction with Sof1, R30 flips outwards by more than 180°, dragging the hydrophobic residues I17, L22, L31 and L33 towards the interaction interface

with Sof1. At the same time, α1 shifts upwards towards the Nop1 interface (Figure 9F), causing a conformational change of the Nop1 C-terminal tail. This alters the interface between Nop1$^{83-327}$ and Nop56$^{1-166}$ at this site (Supplementary Figure 4A). The rest of the interactions between Nop1$^{83-327}$ and Nop56$^{1-166}$ are well conserved in the two structures (Supplementary Figure 4B-C). However, the better resolution of our crystallographic structure allows determining the position of all side chains with a higher level of accuracy.

The comparison between the structures of Nop56–Nop1 in isolation and as part of the U3 snoRNP in the 90S pre-ribosome shows that the eukaryotic-specific segments of both Nop56 and Nop1 may adopt different conformations to adapt to multiple interaction partners. This observation is also in agreement with our results demonstrating that the interaction between the Nop1 C-terminal tail and Nop56-α1 is not essential for the formation of the Nop56–Nop1 complex. This contact site must likely remain flexible to allow adaptation to other interaction partners, and thus cannot be pivotal to the formation of the Nop56–Nop1 complex.

In the cryo-EM structure of the U3 snoRNP bound to the 90S pre-ribosome, the area corresponding to the Nop58–Nop1 complex is poorly defined. Thus, to verify whether the forces driving the formation of the Nop58–Nop1 complex could be the same as those driving the formation of the Nop56–Nop1, we used MODELLER (Webb and Sali 2016) to generate an homology model of the N-terminal domain of Nop58 based on the structure of Nop56$^{1-166}$ in the Nop56$^{1-166}$–Nop1$^{83-327}$ complex (Figure 10A). The surface charge distribution of Nop58$^{1-155}$ at the interface with Nop1 is very similar to that of Nop56$^{1-166}$ and conserves the central negatively charged cavity that interacts with Nop1-α4 (Figure 10B). The residues that are crucial for the interaction of Nop56 with Nop1 are conserved in Nop58 (Figure 10C). These include Nop56-E10 (Nop58-E7) in β1, Nop56-E71 (Nop58-E65) in α3, Nop56-D102 (Nop58-E91) in the loop between α5 and β5, Nop56-D126 (Nop58-115), Nop56-R129 (Nop58-R118) and Nop56-R132 (Nop58-121) in α6, as well as the C-terminal residues of both N-terminal domains (Figure 10C, Supplementary Figure 2). Based on these observations it is likely, that

the driving forces that govern the formation of the Nop1-Nop56 complex are conserved in the Nop1-Nop58 complex.

Interestingly, Nop58$^{1-155}$ and Nop56$^{1-166}$ diverge most significantly in the sequence of the eukaryotic-specific insertion between β2 and β3, which we identified as an adaptable hub for protein binding in Nop56$^{1-166}$ (Supplementary Figure 2 and Figure 10C). This fact underlines that the eukaryotic-specific insertion has emerged from the need to support interactions with binding partners other than Nop1 and that these binding partners differ between Nop56 and Nop58. The amino acid composition of Nop58-α1 may also influence the binding affinity of Nop58$^{1-155}$ for Nop1, as we have shown for the Nop56$^{1-166}$–Nop1 complex (Figure 8A) and could explain the better affinity of Nop1 for Nop56 than for Nop58 (Gautier et al. 1997; Lechertier et al. 2009b).

In conclusion, we have determined the first high-resolution structure of the eukaryotic N-terminal domain of Nop56 bound to the 2'-O-methyltransferase. We have demonstrated that Nop56 has additional structural elements with respect to its archaeal orthologue Nop5, which are involved in the interaction with the eukaryotic methyltransferase and form an adaptable hub for eukaryotic specific binding partners, such as Sof1 in the 90S pre-ribosome. In addition, while the nature of the interaction surface is electrostatic in both eukaryotic and archaeal complexes, the atomic details of the binding are different and reveal coevolution of the two proteins.

The structure and functional mechanism of the eukaryotic Box C/D snoRNP are still unknown, while the structure of the substrate loaded archaeal Box C/D RNP, as well as the regulation of its function, are still a matter of debate. Unclear is also the degree of similarity between the archaeal and eukaryotic enzymes. The structural information presented here is a first step to understand the evolution of the structure and function of the Box C/D RNP responsible for 2'-O-methylation from archaea to eukaryotes.

MATERIAL AND METHODS

Cloning

The genes of Nop1$^{83-327}$ and Nop56 from *S. cerevisiae* were optimized for codon-usage in *E. coli* and ordered from Invitrogen (by Thermo Fisher Scientific). All primers were purchased from Sigma-Aldrich. Dry genes were spun down and resuspended in sterile high LC-MS grade water (Merck 7732-18-5) to a concentration of ~ 0.25 to ~ 0.1 µg/µl. Genes were amplified via PCR using Phusion® high-fidelity DNA Polymerase (New England Biolabs M0530S). All forward and reverse primers contained the cleavage site for NcoI-HF (New England Biolabs R3193S) and NotI-HF (New England Biolabs R3189S), respectively. Amplified genes were purified using the QIAquick PCR purification kit (Qiagen 28104), cleaved with NcoI-HF and NotI-HF and re-purified with the same kit. The vector (pETM-11) was cleaved using the above-mentioned restriction enzymes and purified in the same way. Ligation used T4 DNA Ligase (New England Biolabs M0202S) with the standard protocol provided by the manufacturer; vector and inserts were mixed in a 1:3 ratio using 60 ng of pETM-11. After ligation, the reaction was transformed into chemically competent *E. coli* Top10 cells by transferring to complete ligation mixture (20 µl) to Top10 cells (50 µl) and incubating the mixture for 30 min on ice followed by a 1 min long heat-shock at 42°C. After the heat-shock, 500 µl of LB medium (Carl Roth®) was added and the reaction was incubated for 1 hour at 37 °C. Cells were plated on LB-Agar plates containing Kanamycin (50 µg/ml). The insertion of the gene into to vector was verified by sequencing (Eurofins). Plasmids for Fibrillarin, Nop5, and Nop5$^{1-123}$ subcloned into pETM-11 vectors were available in the lab from previous work (Lapinaite et al. 2013).

Mutagenesis

Mutants Nop1$^{83-321}$, Nop1$^{83-327}$-R231E, Nop1$^{83-327}$-E222K, Nop1$^{83-327}$-D263R/M236K, Nop56$^{1-166}$-L147W/E148F (WF), Nop56$^{1-166}$-E10K/E11K, Nop56$^{1-166}$-E71K/Y158K and Nop56$^{1-166}$-Q35A/N39A were generated using the Pfu Plus! DNA polymerase (Roboklon) and the protocol provided by the manufacturer. PCR products were cleared from the starting material

by digestion with DpnI (New England Biolabs R0176S); the enzyme was heat-inactivated before transformation into *E. coli* OmniMax cells. Positive mutants were verified by sequencing (Eurofins).

Expression and purification

For expression the plasmids were transformed into *E. coli* BL21(DE3) cells; positive transformants were selected by kanamycin resistance. Cells were grown in LB medium at 37 °C until an OD600 of 0.6-0.8 was reached; cells were then shifted to 16 °C except for the cells expressing Fibrillarin and Nop5$^{1-123}$, which were shifted to 20°C. Expression of Nop1$^{83-327}$, Nop56$^{1-166}$, Nop1$^{83-321}$, Nop1$^{83-327}$-R231E, Nop1$^{83-327}$-E222K, Nop1$^{83-327}$-D263R/M236K, Nop56$^{1-166}$-L147W/E148F (WF), Nop56$^{1-166}$-E10K/E11K, Nop56$^{1-166}$-E71K/Y158K and Nop56$^{1-166}$-Q35A/N39A was induced by adding IPTG (Carl Roth® 367-93-1) to a final concentration of 0.5 mM. Expression of Fibrillarin and Nop5$^{1-123}$ was induced by the addition of a final concentration of 1 mM IPTG. All proteins were expressed individually. Cells were harvested 18-20 hours after induction by centrifugation at 4500 rpm at 4°C.

Cells were resuspended in lysis buffer (buffer A: 50mM Tris-HCl, 1 M NaCl, 10% glycerol, 10 mM imidazole, pH 7.5). One tablet of cOmplete, EDTA-free protease inhibitor cocktail (Roche) and 1 mg of lysozyme (Carl Roth® 12650-88-3) were added for 30 min on ice. After incubation, the cells were lysed using sonication for 30 min. The lysate was then centrifuged at 18500 rpm for 1 hour at 4 °C and the supernatant carefully decanted. For Fibrillarin and Nop5$^{1-123,}$ the supernatant was heated for 15 min at 80 °C and again centrifuged as described above.

All proteins were purified from the supernatant using 5 ml HisTrap™ FF columns (Cytiva) connected to Äkta pure or Äkta start systems (GE Healthcare). The supernatant was loaded using lysis buffer (buffer A, described above). The column was then washed three times with three columns volumes of buffer B (50 mM Tris-HCl, 1 M NaCl, 10% glycerol, 10 mM imidazole, 1 M LiCl, pH 7.5) and the protein was eluted using a gradient of buffer C (50 mM Tris-HCl, 1 M NaCl, 10% glycerol, 1 M imidazole, pH 7.5). The collected proteins were buffer-

exchanged in buffer A using a HiPrep 26/10 Desalting column (Cytiva). To remove the N-terminal His-tag all proteins were incubated with His-tagged TEV protease (made in-house) overnight at 4 °C. TEV protease and cleaved His-Tag were removed from the proteins using a 5 ml HisTrap™ FF column (Cytiva).

Complex assembly

Purified Nop1$^{83-327}$ and Nop56$^{1-166}$, or mutants thereof, were mixed in a 1:1 ratio and incubated for 15 min at room temperature before further purification steps. All complexes were purified using size-exclusion chromatography (SEC) on an Äkta pure system at room temperature with running buffer D (50 mM sodium phosphate, 100 mM NaCl, 10 mM β-mercaptoethanol, pH 7.0) or running buffer E for crystallization (50 mM Tris-HCl, 50 mM NaCl, 1 mM DTT, pH 7.5). A Superdex 200 Increase 10/300 GL column (Cytiva) or, for the preparation of the crystallization sample, a HiLoad 16/600 Superdex 75 pg column (Cytiva) were used. Purity was assessed using SDS gel-electrophoresis. For crystallization, the purified complex was concentrated using Amicon® Ultra-15 10K centrifugal filters (Merck).

Isothermal Titration Calorimetry measurements

Isothermal titration calorimetry measurements of Nop1$^{83-327}$ with Nop56$^{1-166}$ and of Nop1$^{83-327}$ with Nop56$^{1-166}$-Q35A/N39A were performed using a TA Instruments Nano ITC calorimeter with a cell temperature of 20 °C applying a stirring rate of 300 rpm. Due to the limited solubility of Nop1$^{83-327}$ in the absence of Nop56$^{1-166}$ under low salt conditions, both proteins were kept in 50 mM sodium phosphate, 500 mM NaCl, 10 mM β-mercaptoethanol, pH 7.0, rather than in running buffer D, which was instead used for size-exclusion chromatography. Nop1$^{83-327}$ was placed in the measurement cell at either 100 µM or 85 µM, while Nop56$^{1-166}$ and Nop56$^{1-166}$-Q35A/N39A were used as titrant at concentrations of either 970 µM or 917 µM. Blank measurements were performed for both titration series. The data was analyzed with the NanoAnalyze software provided by TA Instruments and plotted with OriginPro 2020. The exact values of the K_D and the other binding parameters reported in Supplementary

Figure 3 are however inaccurate, as some Nop1$^{83-327}$ precipitated during the measurements due to stirring.

Crystallization

A concentrated solution of ~ 10 mg/ml of Nop56$^{1-166}$-Nop1$^{83-327}$ in crystallization buffer was used for crystallization by sitting drop vapor diffusion. Initial crystallization screens were set up with a Crystal Pheonix crystallization robot (Art Robbins Instruments) using NeXtal DWBlock Suites (Qiagen); JCSG Core I Suite, JCSG Core II Suite, JCSG Core II Suite, JCSG Core IV Suite, Protein Complex Suite, Nucleix Suite, PEG Suite, and PEG II Suite. The drop solution was equilibrated against 200 µl of reservoir solution at 12 °C. Crystals appeared in the Protein Complex Suite screen as well as in the PEG Suite screen after one week. Based on the two best hit conditions, grid screens were set up and incubated at 12 °C. After one week crystals were obtained in multiple conditions, the best of which was 0.1 M HEPES (pH 7.5), 0.1 M LiCl, and 40% PEG400. Cryo-protection was achieved by the addition of 10% (2*R*, 3*R*)-2,3-butanediol before flash-freezing.

Data collection and processing

Data were collected at beamline P11 of PETRA III, DESY (Deutsches Elektronen-Synchroton, Hamburg, Germany) (Burkhardt et al. 2016). The datasets were recorded at 100K (λ = 1.03 Å) and processed using the AutoPROC toolbox (Vonrhein et al. 2011) executing XDS (Kabsch 2010) followed by Pointless (Evans 2006) and Aimless (Evans and Murshudov 2013) from the CCP4 program suite (Winn et al. 2011). The high-resolution cut off was determined using a signal/noise ratio (I/σ(I)) of 2.0.

Structure determination and representation

The crystal structure of Nop56$^{1-166}$-Nop1$^{83-327}$ was solved by molecular replacement executing Balbes (Long et al. 2008) from the CCP4 suite (Winn et al. 2011). For Nop1$^{83-327}$ human Fibrillarin (PDB-ID: 2IPX, sequence identity: 74%) was found as a working search model. For Nop56$^{1-166}$ a model was built using the Phyre2 web server (Kelley et al. 2015) and energy

minimized using Maestro from the Schrödinger2018 suite (Schrödinger, LLC). This model was used as a search model using Phaser (McCoy et al. 2007) from the Phenix suite (Adams et al. 2010) including the fixed partial solution containing only the model for Nop1$^{83-327}$. Initial structural models were built with AutoBuild (Terwilliger et al. 2008) from the Phenix software package (Adams et al. 2010). The models were analyzed and completed by iterative model-building and refinement cycles using Coot (Emsley et al. 2010) and Phenix.refine (Afonine et al. 2012), including TLS-refinement and the addition of hydrogens in riding positions. Data collection and refinement statistics are summarized in Table 1.

Surface electrostatics for Figure 4 and 5 were calculated using the Adaptive Poisson-Boltzmann Solver (APBS) 2.1 (Jurrus et al. 2018).

Sequence alignments

All alignments shown in Figures 6 and 7 were done manually based on available structures or models and formatted with Espript 3.0 (Robert and Gouet 2014). (Sc – *S. cerevisiae*, Sp – *S. pombe*, Hs – *H. sapiens*, Mm – *M. musculus*, Xl – *X. leavis*, Dr – *D. rerio*, Pf - *P. furiosus*, Ph – *P. horikoshi*, Ss – *S. solfataricus*, Af – *A. fulgidus*, Mj – *M. janaschii*).

Data Availability

Structural data have been deposited in the Protein Data Bank (https://www.rcsb.org) with the PDB code 6ZDT.

ACKNOWLEDGMENTS

We thank DESY (Hamburg, Germany), a member of the Helmholtz Association HGF, for the provision of experimental facilities. Parts of this research were carried out at PETRA III using beamline P11. The authors thank Ute Widow and Susanne zur Lage, HZI, for excellent technical assistance in the wet lab. This work has been funded by the Deutsche Forschungsgemeinschaft through a project grant to TC in the frame of the SPP “Chemical Biology of Natural Nucleic Acids Modifications” (CA 294/11-1).

REFERENCES

Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, et al. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst D* **66**: 213–221.

Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. 2012. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**: 352–367.

Aittaleb M, Rashid R, Chen Q, Palmer JR, Daniels CJ, Li H. 2003. Structure and function of archaeal box C/D sRNP core proteins. *Nature Structural & Molecular Biology* **10**: 256–263.

Appel CD, Maxwell ES. 2007. Structural features of the guide:target RNA duplex required for archaeal box C/D sRNA-guided nucleotide 2'-O-methylation. *RNA* **13**: 899–911.

Arai T, Ishiguro K, Kimura S, Sakaguchi Y, Suzuki T, Suzuki T. 2015. Single methylation of 23S rRNA triggers late steps of 50S ribosomal subunit assembly. *Proc Natl Acad Sci USA* **112**: E4707–E4716.

Ayadi L, Galvanin A, Pichot F, Marchand V, Motorin Y. 2019. RNA ribose methylation (2'-O-methylation): Occurrence, biosynthesis and biological functions. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1862**: 253–269.

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.

Baudin-Baillieu A, Fabret C, Liang X, Piekna-Przybylska D, Fournier MJ, Rousset J-P. 2009. Nucleotide modifications in three functionally important regions of the Saccharomyces cerevisiae ribosome affect translation accuracy. *Nucleic Acids Res* **37**: 7665–7677.

Brochier-Armanet C, Forterre P, Gribaldo S. 2011. Phylogeny and evolution of the Archaea: one hundred genomes later. *Current Opinion in Microbiology* **14**: 274–281.

Burkhardt A, Pakendorf T, Reime B, Meyer J, Fischer P, Stübe N, Panneerselvam S, Lorbeer O, Stachnik K, Warmer M, et al. 2016. Status of the crystallography beamlines at PETRA III. *Eur Phys J Plus* **131**: 56.

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

Cheng J, Baßler J, Fischer P, Lau B, Kellner N, Kunze R, Griesel S, Kallas M, Berninghausen O, Strauss D, et al. 2019. Thermophile 90S Pre-ribosome Structures Reveal the Reverse Order of Co-transcriptional 18S rRNA Subdomain Integration. *Mol Cell* **75**: 1256-1269.e7.

Chong PA, Vernon RM, Forman-Kay JD. 2018. RGG/RG Motif Regions in RNA Binding and Phase Separation. *J Mol Biol* **430**: 4650–4665.

Cowling VH, Turner SA, Cole MD. 2014. Burkitt's lymphoma-associated c-Myc mutations converge on a dramatically altered target gene response and implicate Nol5a/Nop56 in oncogenesis. *Oncogene* **33**: 3519–3527.

Dimitrova DG, Teysset L, Carré C. 2019. RNA 2'-O-Methylation (Nm) Modification in Human Diseases. *Genes* **10**: 117.

Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**: 486–501.

Evans P. 2006. Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**: 72–82.

Evans PR, Murshudov GN. 2013. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **69**: 1204–1214.

Gagnon KT, Biswas S, Zhang X, Brown BA, Wollenzien P, Mattos C, Maxwell ES. 2012. Structurally Conserved Nop56/58 N-terminal Domain Facilitates Archaeal Box C/D Ribonucleoprotein-guided Methyltransferase Activity. *J Biol Chem* **287**: 19418–19428.

Gautier T, Bergès T, Tollervey D, Hurt E. 1997. Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis. *Molecular and Cellular Biology* **17**: 7088–7098.

Graziadei A, Gabel F, Kirkpatrick J, Carlomagno T. 2020. The guide sRNA sequence determines the activity level of box C/D RNPs eds. L.E. Kay and P.A. Cole. *eLife* **9**: e50027.

Graziadei A, Masiewicz P, Lapinaite A, Carlomagno T. 2016. Archaea box C/D enzymes methylate two distinct substrate rRNA sequences with different efficiency. *RNA* **22**: 764–772.

Green R, Noller HF. 1996. In vitro complementation analysis localizes 23S rRNA posttranscriptional modifications that are required for Escherichia coli 50S ribosomal subunit assembly and function. *RNA* **2**: 1011–1021.

Ishiguro K, Arai T, Suzuki T. Depletion of S-adenosylmethionine impacts on ribosome biogenesis through hypomodification of a single rRNA methylation. *Nucleic Acids Research*. https://academic.oup.com/nar/article/47/8/4226/5364157?rss=1.

Iyer-Bierhoff A, Krogh N, Tessarz P, Ruppert T, Nielsen H, Grummt I. 2018. SIRT7-Dependent Deacetylation of Fibrillarin Controls Histone H2A Methylation and rRNA Synthesis during the Cell Cycle. *Cell Rep* **25**: 2946-2954.e5.

Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, Brookes DH, Wilson L, Chen J, Liles K, et al. 2018. Improvements to the APBS biomolecular solvation software suite. *Protein Science* **27**: 112–128.

Kabsch W. 2010. XDS. *Acta Crystallogr D Biol Crystallogr* **66**: 125–132.

Karplus PA, Diederichs K. 2012. Linking crystallographic model and data quality. *Science* **336**: 1030–1033.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10**: 845–858.

Kimura S, Suzuki T. 2010. Fine-tuning of the ribosomal decoding center by conserved methyl-modifications in the Escherichia coli 16S rRNA. *Nucleic Acids Res* **38**: 1341–1352.

Klein DJ. 2001. The kink-turn: a new RNA secondary structure motif. *The EMBO Journal* **20**: 4214–4221.

Kozakov D, Beglov D, Bohnuud T, Mottarella SE, Xia B, Hall DR, Vajda S. 2013. How good is automated protein docking? *Proteins* **81**: 2159–2166.

Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, Beglov D, Vajda S. 2017. The ClusPro web server for protein-protein docking. *Nat Protoc* **12**: 255–278.

Lapinaite A, Simon B, Skjaerven L, Rakwalska-Bange M, Gabel F, Carlomagno T. 2013. The structure of the box C/D enzyme reveals regulation of RNA methylation. *Nature* **502**: 519–523.

Lechertier T, Grob A, Hernandez-Verdun D, Roussel P. 2009a. Fibrillarin and Nop56 interact before being co-assembled in box C/D snoRNPs. *Exp Cell Res* **315**: 928–942.

Lechertier T, Grob A, Hernandez-Verdun D, Roussel P. 2009b. Fibrillarin and Nop56 interact before being co-assembled in box C/D snoRNPs. *Experimental Cell Research* **315**: 928–942.

Lin J, Lai S, Jia R, Xu A, Zhang L, Lu J, Ye K. 2011. Structural basis for site-specific ribose methylation by box C/D RNA protein complexes. *Nature* **469**: 559–563.

Long F, Vagin AA, Young P, Murshudov GN. 2008. BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr* **64**: 125–132.

Lyman SK, Gerace L, Baserga SJ. 1999. Human Nop5/Nop58 is a component common to the box C/D small nucleolar ribonucleoproteins. *RNA* **5**: 1597–1604.

Marcel V, Ghayad SE, Belin S, Therizols G, Morel A-P, Solano-Gonzàlez E, Vendrell JA, Hacot S, Mertani HC, Albaret MA, et al. 2013. p53 acts as a safeguard of translational control by regulating fibrillarin and rRNA methylation in cancer. *Cancer Cell* **24**: 318–330.

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Crystallogr* **40**: 658–674.

McKeegan KS, Debieux CM, Watkins NJ. 2009. Evidence that the AAA+ proteins TIP48 and TIP49 bridge interactions between 15.5K and the related NOP56 and NOP58 proteins during box C/D snoRNP biogenesis. *Mol Cell Biol* **29**: 4971–4981.

Murata T, Suzuki E, Ito S, Sawatsubashi S, Zhao Y, Yamagata K, Tanabe M, Fujiyama S, Kimura S, Ueda T, et al. 2008. RNA-binding protein hoip accelerates polyQ-induced neurodegeneration in Drosophila. *Biosci Biotechnol Biochem* **72**: 2255–2261.

Nachmani D, Bothmer AH, Grisendi S, Mele A, Bothmer D, Lee JD, Monteleone E, Cheng K, Zhang Y, Bester AC, et al. 2019. Germline NPM1 mutations lead to altered rRNA 2'-O-methylation and cause dyskeratosis congenita. *Nat Genet* **51**: 1518–1529.

Natchiar SK, Myasnikov AG, Kratzat H, Hazemann I, Klaholz BP. 2017. Visualization of chemical modifications in the human 80S ribosome structure. *Nature* **551**: 472–477.

Nolivos S, Carpousis AJ, Clouet-d'Orval B. 2005. The K-loop, a general feature of the Pyrococcus C/D guide RNAs, is an RNA structural motif related to the K-turn. *Nucleic Acids Res* **33**: 6507–6514.

Oruganti S, Zhang Y, Li H, Robinson H, Terns MP, Terns RM, Yang W, Li H. 2007. Alternative Conformations of the Archaeal Nop56/58-Fibrillarin Complex Imply Flexibility in Box C/D RNPs. *Journal of Molecular Biology* **371**: 1141–1150.

Paul A, Tiotiu D, Bragantini B, Marty H, Charpentier B, Massenet S, Labialle S. 2019. Bcd1p controls RNA loading of the core protein Nop58 during C/D box snoRNP biogenesis. *RNA* **25**: 496–506.

Peng Y, Yu G, Tian S, Li H. 2014. Co-Expression and Co-Purification of Archaeal and Eukaryal Box C/D RNPs ed. L. Randau. *PLoS ONE* **9**: e103096.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**: 1605–1612.

Polikanov YS, Melnikov SV, Söll D, Steitz TA. 2015. Structural insights into the role of rRNA modifications in protein synthesis and ribosome assembly. *Nature Structural & Molecular Biology* **22**: 342–344.

Prieto MB, Georg RC, Gonzales-Zubiate FA, Luz JS, Oliveira CC. 2015. Nop17 is a key R2TP factor for the assembly and maturation of box C/D snoRNP complex. *BMC Mol Biol* **16**: 7.

Qu G, van Nues RW, Watkins NJ, Maxwell ES. 2011. The Spatial-Functional Coupling of Box C/D and C'/D' RNPs Is an Evolutionarily Conserved Feature of the Eukaryotic Box C/D snoRNP Nucleotide Modification Complex. *Molecular and Cellular Biology* **31**: 365–374.

Quinternet M, Rothé B, Barbier M, Bobo C, Saliou J-M, Jacquemin C, Back R, Chagot M-E, Cianférani S, Meyer P, et al. 2015. Structure/Function Analysis of Protein–Protein Interactions Developed by the Yeast Pih1 Platform Protein and Its Partners in Box C/D snoRNP Assembly. *Journal of Molecular Biology* **427**: 2816–2839.

Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**: W320–W324.

Rodriguez-Corona U, Sobol M, Rodriguez-Zapata LC, Hozak P, Castano E. 2015. Fibrillarin from Archaea to human: Review on fibrillarin. *Biol Cell* **107**: 159–174.

Shi X, Huang L, Lilley DMJ, Harbury PB, Herschlag D. 2016. The solution structural ensembles of RNA kink-turn motifs and their protein complexes. *Nature Chemical Biology* **12**: 146–152.

Siibak T, Remme J. 2010. Subribosomal particle analysis reveals the stages of bacterial ribosome assembly at which rRNA nucleotides are modified. *RNA* **16**: 2023–2032.

Sloan KE, Warda AS, Sharma S, Entian K-D, Lafontaine DLJ, Bohnsack MT. 2017. Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biology* **14**: 1138–1152.

Smith D-L, Erce MA, Lai Y-W, Tomasetig F, Hart-Smith G, Hamey JJ, Wilkins MR. 2020. Crosstalk of Phosphorylation and Arginine Methylation in Disordered SRGG Repeats of Saccharomycescerevisiae Fibrillarin and Its Association with Nucleolar Localization. *J Mol Biol* **432**: 448–466.

Špačková N, Réblová K, Šponer J. 2010. Structural Dynamics of the Box C/D RNA Kink-Turn and Its Complex with Proteins: The Role of the A-Minor 0 Interaction, Long-Residency Water Bridges, and Structural Ion-Binding Sites Revealed by Molecular Simulations. *J Phys Chem B* **114**: 10581–10593.

Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, Read RJ, Adams PD. 2008. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* **64**: 61–69.

Tessarz P, Santos-Rosa H, Robson SC, Sylvestersen KB, Nelson CJ, Nielsen ML, Kouzarides T. 2014. Glutamine methylation in histone H2A is an RNA-polymerase-I-dedicated modification. *Nature* **505**: 564–568.

Tollervey D, Lehtonen H, Carmo-Fonseca M, Hurt EC. 1991. The small nucleolar RNP protein NOP1 (fibrillarin) is required for pre-rRNA processing in yeast. *EMBO J* **10**: 573–583.

Tran E, Zhang X, Lackey L, Maxwell ES. 2005. Conserved spacing between the box C/D and C'/D' RNPs of the archaeal box C/D sRNP complex is required for efficient 2'-O-methylation of target RNAs. *RNA* **11**: 285–293.

Tran EJ, Zhang X, Maxwell ES. 2003. Efficient RNA 2'-O-methylation requires juxtaposed and symmetrically assembled archaeal box C/D and C'/D' RNPs. *EMBO J* **22**: 3930–3940.

Vajda S, Yueh C, Beglov D, Bohnuud T, Mottarella SE, Xia B, Hall DR, Kozakov D. 2017. New Additions to the ClusPro Server Motivated by CAPRI. *Proteins* **85**: 435–444.

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: Extended snoRNA-rRNA base-pairing. *The EMBO Journal* **30**: 2420–2430.

Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, Womack T, Bricogne G. 2011. Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr D Biol Crystallogr* **67**: 293–302.

Wang H, Boisvert D, Kim KK, Kim R, Kim S-H. 2000. Crystal structure of a fibrillarin homologue from Methanococcus jannaschii, a hyperthermophile, at 1.6 Å resolution. *The EMBO Journal* **19**: 317–323.

Webb B, Sali A. 2016. Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Bioinformatics* **54**: 5.6.1-5.6.37.

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, et al. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**: 235–242.

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 USA* **113**: 10878–10883.

Yang Z, Wang J, Huang L, Lilley DMJ, Ye K. 2020. Functional organization of box C/D RNA-guided RNA methyltransferase. *Nucleic Acids Res*.

https://academic.oup.com/nar/article/doi/10.1093/nar/gkaa247/5820881 (Accessed April 16, 2020).

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. *PNAS* **106**: 13808–13813.

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

Yu G, Zhao Y, Li H. 2018. The multistructural forms of box C/D ribonucleoprotein particles. *RNA* **24**: 1625–1633.

Yu G, Zhao Y, Tian S, Rai J, He H, Spear J, Sousa D, Fan J, Yu H-G, Stagg SM, et al. 2019. Yeast R2TP Interacts with Extended Termini of Client Protein Nop58p. *Sci Rep* **9**: 20228.

TABLES

Table 1. Crystallographic data collection & refinement statistics

Structure	Nop56$^{1-166}$-Nop1$^{83-327}$
PDB-ID:	**6ZDT**
Data collection	
Beamline	P11, PETRA III, DESY
Wavelength (Å)	1.03
Space group	$P2_12_12$
Cell dimensions	
a, *b*, *c* (Å)	69.04, 118.29, 48.83
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)[a]	1.71-48.83 (1.71-1.74)
R_{merge} (%)[a,1]	7.7 (121.4)
R_{pim} (%)[a,2]	2.1 (35.5)
$I/\sigma I$[a]	22.2 (2.1)
Completeness (%)[a]	100 (100)
Redundancy[a]	13.2 (12.6)
$CC_{1/2}$ (%)[a] (Karplus and Diederichs 2012)	100 (85.6)
Refinement	
Resolution (Å)	1.71-48.83
No. reflections	44211
R_{work}/ R_{free} (%)	16.28/20.49
No. atoms	3870
Protein	3503
Ligand/ion	-
Water	367
B-factors (Å^2)	33.20
Protein	32.56
Ligand/ion	-
Water	39.36
R.m.s deviations	
Bond lengths (Å)	0.005
Bond angles (°)	0.686
Ramachandran statistics (%)	
Favored	98.18
Allowed	1.82
Outliers	0.00

[a] Values for the highest resolution shell are shown in parentheses.

[1] $R_{merge} = \Sigma_h \Sigma_i |\langle I_h \rangle - I_{h,i}| / \Sigma_h \Sigma_i I_{h,i}$, where h enumerates the unique reflections and i are their symmetry-equivalent contributions

[2] $R_{pim} = \Sigma_h [1/(n_h - 1)]^{1/2} \Sigma_i |\langle I_h \rangle - I_{h,I}| \Sigma_h \Sigma_i I_{h,I}$, where h enumerates the unique reflections and i are their symmetry-equivalent contributions

FIGURE LEGENDS

Figure 1. Comparison of archaeal and eukaryotic Box C/D enzymes. (A) Two different structural models have been proposed for the archaeal Box C/D enzyme loaded with substrate RNA: the mono-RNP model and the di-RNP model. (B) The eukaryotic Box C/D enzyme has been found to assemble as a mono-RNP in its substrate-free form. Snu13, of which only one copy is present in Box C/D snoRNPs capable of methylation (upper panel), has been found to bind snoRNA with a 2:1 stoichiometry in snoRNPs that are not involved in 2'-O-methylation (lower panel).

Figure 2. Structural comparison of *Sc* Nop1$^{83-327}$ and Nop56$^{1-166}$ to their archaeal orthologs Fibrillarin and Nop5. (A) Left, structure of Nop1$^{83-327}$ with the small N-terminal and central domains highlighted. The active site is indicated with red dashed lines. The structure is part of the *Sc* Nop56$^{1-166}$–Nop1$^{83-327}$ complex structure determined here (see Figure 4). Right, superposition of the structure of Nop1$^{83-327}$ with those of archaeal orthologs. *Sc* Nop1$^{83-327}$ is in slate, *Pf* Fibrillarin (PDB-ID: 2NNW) in green, *Ss* Fibrillarin (PDB-ID: 3ID5) in forest and *Af* Fibrillarin (PDB-ID: 1NT2) in smudge. (B) Left, structure of Nop56$^{1-166}$ with secondary structure elements indicated. The structure is part of the *Sc* Nop56$^{1-166}$ –Nop1$^{83-327}$ complex structure determined here (see Figure 4). Right, superposition of the structure of *Sc* Nop56$^{1-166}$ with those of Nop5 N-terminal domains from archaea. *Sc* Nop56$^{1-166}$ is in greencyan, *Pf* Nop5$^{1-123}$ (PDB-ID: 2NNW) in hotpink, *Ss* Nop5$^{1-132}$ (PDB-ID: 3ID5) in magenta and *Af* Nop5$^{1-71}$ (PDB-ID: 1NT2) in violet-purple. RMSD values were calculated between all heavy atoms.

Figure 3. Analysis of cross-species complex formation between eukaryotic and archaeal Fibrillarin and Nop56/Nop5 proteins. (A) Overlay of size-exclusion chromatograms of *Sc* Nop1$^{83-327}$ (blue), *Pf* Nop5$^{1-123}$ (pink), and *Pf* Nop5$^{1-123}$ together with *Sc* Nop1$^{83-327}$ (black). Peak 1 corresponds to *Sc* Nop1$^{83-327}$ and peak 2 to *Pf* Nop5$^{1-123}$. (B) Overlay of size-exclusion chromatograms of *Sc* Nop56$^{1-166}$ (cyan), *Pf* Fibrillarin (green), and *Sc* Nop56$^{1-166}$ together with

Pf Fibrillarin (black). (C) 4-15% Mini-PROTEAN® TGX Precast Protein Gel (BioRad) of the chromatograms shown in panel A–D. The color code above the lanes indicates the proteins present in the injection mixture: *Sc* Nop1$^{83-327}$ (blue), *Sc* Nop56$^{1-166}$ (cyan), *Sc* Nop56$^{1-166}$-WF (orange), *Pf* Fibrillarin (green) and *Pf* Nop5$^{1-123}$ (pink). Blue/pink lanes represent fractions collected from peak 1 and peak 2 of the black chromatogram of Figure 3A. The fraction loaded in the middle lane corresponds to the region between peaks 1 and 2. The green/cyan lane represents the elution peak of the black chromatogram of Figure 3B. The blue/orange lane represents the elution peak from the black chromatogram of Figure 3D. The green/orange lane represents the elution peak from the black chromatogram of Figure 3F. Lane M contains the protein standard (D) Overlay of size-exclusion chromatograms of *Pf* Fibrillarin (green), *Pf* Nop5$^{1-123}$ (pink), *Sc* Nop1$^{83-327}$ (blue), *Sc* Nop56$^{1-166}$ (cyan), *Pf* Fibrillarin with *Pf* Nop5$^{1-123}$ (grey) and *Sc* Nop1$^{83-227}$ with *Sc* Nop56$^{1-166}$ (black). (E) Overlay of size-exclusion chromatograms of *Sc* Nop1$^{83-327}$ (blue), *Sc* Nop56$^{1-166}$-WF mutant (orange) and the mixture of *Sc* Nop1$^{83-327}$ and *Sc* Nop56$^{1-166}$-WF mutant (black). (F) Overlay of size-exclusion chromatograms of *Pf* Fibrillarin (green), *Sc* Nop56$^{1-166}$-WF mutant (orange), and the mixture of *Pf* Fibrillarin and Nop56$^{1-166}$-WF mutant (black).

Figure 4. Crystallographic structure of *Sc* Nop1$^{83-327}$ in complex with *Sc* Nop56$^{1-166}$. (A) Overview of the structure with Nop56$^{1-166}$ in greencyan and Nop1$^{83-327}$ in slate with secondary structure elements indicated. (B) Solvent accessible surface electrostatics of the interaction surfaces of Nop1$^{83-327}$ and Nop56$^{1-166}$. Blue, red and white represent positive, negative and neutral charges, respectively. (C–E) Details of the intermolecular interactions between Nop56$^{1-166}$ and Nop1$^{83-327}$ in three different areas of the interface. Hydrogen bonds are depicted as yellow dashed lines.

Figure 5. Comparison of the solvent-accessible surface electrostatics at the protein–protein interface in archaeal and eukaryotic complexes. (A) *Sc* Nop56$^{1-166}$–Nop1$^{83-327}$ complex; (B) *Pf*

Nop5$^{1-123}$–Fibrillarin (PDB-ID: 2NNW); (C) *Ss* Nop5$^{1-132}$–Fibrillarin (PDB-ID: 3ID5); (D) *Af* Nop5$^{1-71}$–Fibrillarin (PDB-ID:1NT2). Blue, red and white represent positive, negative, and neutral charges, respectively.

Figure 6. The interactions of a conserved tyrosine residue differ between archaea and eukaryotes. (A) Interactions of Nop1$^{83-327}$-Y195 involving residues of Nop56$^{1-166}$-α7. (B–E) Conserved hydrophobic clusters around the conserved tyrosine in the complexes of archaeal Fibrillarin and Nop5: (B) *Pf* Nop5$^{1-123}$–Fibrillarin (PDB-ID: 2NNW); (C) *Ss* Nop5$^{1-132}$–Fibrillarin (PDB-ID: 3ID5)*;* (D) *Af* Nop5$^{1-71}$–Fibrillarin (PDB-ID: 1NT2); (E) docking model of the complex between the Nop5 N-terminal domain and Fibrillarin from *M. janaschii* (PDB-ID: 3T7Z and 1G8S). (F) Multiple sequence alignment of the fragment from eukaryotic and archaeal Fibrillarins containing the conserved tyrosine residue (highlighted in green). Residue numbers refer to the first line of the alignment. (G) Multiple sequence alignment of residues in eukaryotic Nop56-α7 that interact with the conserved tyrosine residue in Nop1/Fibrillarin. Conserved interacting residues are highlighted in turquoise. (H) Multiple sequence alignment of residues in archaeal Nop5-α6 that interact with the conserved tyrosine residue in Fibrillarin. Conserved interacting aromatic residues are highlighted in pink.

Figure 7. Comparison of interaction hot-spots in the eukaryotic Nop56$^{1-166}$–Nop1$^{83-327}$ complex and ortholog archaeal complexes. (A) Expansion of the region comprising Nop56-α1 and the C-terminal tail of Nop1 of the *Sc* Nop56$^{1-166}$–Nop1$^{83-327}$ complex overlayed with archaeal orthologous Nop5 –Fib complexes. *Sc* Nop1$^{83-327}$, *Pf* and *Ss* Fibrillarin are in slate, green and forest, respectively; *Sc* Nop56$^{1-166}$, *Pf* and *Ss* Nop5 are in greencyan, hotpink and magenta, respectively. The *Pf* and *Ss* structures are from PDB-IDs 2NNW and 3ID5, respectively. (B) Contacts between Nop56$^{1-166}$-α1 and -α6 and Nop1. (C) Contacts between *Pf* Nop5-α5 and *Pf* Fibrillarin. (D) Contacts between *Ss* Nop5-α5 and *Ss* Fibrillarin. Hydrogen bonds are depicted as yellow dashed lines. (E) Multiple sequence alignment of the C-terminal residues of eukaryotic and archaeal Fibrillarin; conserved residues are highlighted

in slate. (F) Multiple sequence alignment of Nop56-α1 in eukaryotes. Interacting residues in *S. cerevisiae* are highlighted in greencyan. (G) Multiple sequence alignment of eukaryotic Nop56-α6 and archaeal Nop5-α5. Interacting residues are highlighted in greencyan. (H) Contacts between Nop1$^{83-327}$-α4 and Nop56$^{1-166}$. (I) Contacts between *Pf* Fib-α4 and *Pf* Nop5. (J) Contacts between *Ss* Fib-α4 and *Ss* Nop5 (K) Contacts between *Af* Fib-α4 and *Af* Nop5. (L) Multiple sequence alignment of α4 residues of eukaryotic and archaeal Fibrillarins; conserved and interacting residues are highlighted in slate. Residue numbers refer to the first line of the alignment.

Figure 8. Experimental verification of the Nop56$^{1-166}$–Nop1$^{83-327}$ binding interface. (A) Left panel: overlay of size-exclusion chromatograms of Nop56$^{1-166}$ with Nop1$^{83-327}$ (black), Nop56$^{1-166}$ with Nop1$^{83-321}$ (green) and Nop56$^{1-166}$-Q35A/N39A with Nop1$^{83-327}$ (pink). Right panel: isothermal titration calorimetry binding curves of Nop1$^{83-327}$ with Nop56$^{1-166}$ (black) and Nop1$^{83-327}$ with Nop56$^{1-166}$-Q35A/N39A (pink). The approximate K_D values determined for each complex are given in the respective color. Only approximate K_D values can be provided, as stirring caused partial precipitation of Nop1$^{83-327}$ during the measurements, thus impeding an accurate estimate of the Nop1$^{83-327}$ concentration. The K_D of the Nop1$^{83-321}$– Nop56$^{1-166}$ complex could not be measured due to the poor solubility of Nop1$^{83-321}$. (B) Overlay of size-exclusion chromatograms of Nop56$^{1-166}$ with Nop1$^{83-327}$ (black), Nop56$^{1-166}$ with Nop1$^{83-327}$-R231E (magenta) and Nop56$^{1-166}$-E10K/E11K with Nop1$^{83-327}$ (red). (C) Left panel: overlay of the structures of the *Sc* Nop1$^{83-327}$–Nop56$^{1-166}$ complex and *Pf* Nop5–Fib complex showing the potential charge repulsion between *Pf* Nop5$^{1-123}$-R122 and K40 (pink) and *Sc* Nop1$^{83-327}$-K229 (blue) in a putative cross-species complex. *Sc* Nop56$^{1-166}$ is in greencyan and *Pf* Fib is in green. Middle panel: overlay of the structures of the *Sc* Nop1$^{83-327}$–Nop56$^{1-166}$ complex and *Pf* Nop5–Fib complex showing the steric clash between *Pf* Fib-R170 (green) and Nop56$^{1-166}$-α1 (greencyan) as well as the charge repulsion between *Pf* Fib-K143 and Nop56$^{1-166}$-R129 and R140 in a putative cross-species complex. Right panel: overlay of size-exclusion chromatograms showing that Nop56$^{1-166}$-E71K/Y158 and Nop1$^{83-327}$-D263R/M263K do not

form a complex with wild-type Nop1$^{83-327}$ and Nop56$^{1-166}$, respectively. Nop56$^{1-166}$-E71K/Y158K is also unable to interact with of Nop1$^{83-327}$-E222K, as expected. All chromatograms in A–C show the absorbance at 280 nm and use the following color code: Nop1$^{83-327}$ alone, blue; Nop56$^{1-166}$ alone, greencyan; Nop1$^{83-327}$ with Nop56$^{1-166}$, black; Nop1$^{83-327}$ with Nop56$^{1-166}$-Q35A/N39A, hotpink; Nop1$^{83-321}$ with Nop56$^{1-166}$, green; Nop1$^{83-327}$ with Nop56$^{1-166}$-E10K/E11K, red; Nop1$^{83-327}$-R231E with Nop56$^{1-166}$, pink; Nop1$^{83-327}$ with Nop56$^{1-166}$-E71K/Y158K, orange; Nop1$^{83-327}$-E222K with Nop56$^{1-166}$-E71K/Y158K, olive; Nop1$^{83-327}$-D263R/M263K with Nop56$^{1-166}$, purple. (D) 15% SDS polyacrylamide gels showing the content of all elution peaks from the chromatograms in A–C. Lanes belonging to a specific chromatogram are labeled in the same color as the corresponding chromatogram and as in the legend. Multiple lanes from the same chromatogram represent multiple fractions with increasing elution volumes. The faint band at ~ 23 kDa corresponds to N-terminal His-tagged Nop56$^{1-166}$, present because of incomplete cleavage.

Figure 9. The Nop56–Nop1 complex undergoes structural changes upon interaction with other proteins. (A) Comparison of the high-resolution X-ray structure of *Sc* Nop1$^{83-327}$ in complex with the *Sc* Nop56 N-terminal domain (left) and the same complex as part of the U3 snoRNP in the cryo-EM structure of the 90S pre-ribosome solved by at 3.8 Å resolution (PDB: 5WLC) (right). Highlighted are the two regions displaying conformational changes: region 1 comprises of the eukaryotic specific insertion between Nop56-β2 and β3 and region 2 comprises of the Lys-rich loop between α4 and β4. (B) Different surface charge distributions in region 1 of Nop56$^{1-166}$ in the two structures of panel A. (C) The conformational changes in region 1 of the Nop56 N-terminal domain within the 90S pre-ribosome are induced by the binding to Sof1. (D) The conformational changes in the exposed Lys-rich loop in region 2 of the Nop56 N-terminal domain within the 90S pre-ribosome are induced by the binding to the 5' ETS RNA (E) Structural details of region 1 of Nop56$^{1-166}$ in the structure of the Nop1$^{83-327}$–Nop56$^{1-166}$ complex. (F) Detailed view of the conformational changes within region 1 of Nop56$^{1-166}$ upon interaction with Sof1 in the 90S pre-ribosome structure. (G)

Structural details of region 2 of Nop56$^{1-166}$ in the structure of the Nop1$^{83-327}$–Nop56$^{1-166}$ complex. (H) Structural details of region 2 of Nop56$^{1-166}$ interacting with the 5' ETS RNA in the 90S pre-ribosome.

Figure 10. Homology model of *Sc* Nop58 N-terminal domain. (A) Homology model generated with MODELLER (Webb and Sali 2016) using the high-resolution structure of Nop56$^{1-166}$ in the Nop56$^{1-166}$–Nop1$^{83-327}$ as template. The secondary structure elements are indicated. (B) Surface electrostatics of the Nop1-binding interface of the Nop58$^{1-155}$ homology model (top) and the Nop56$^{1-166}$ structure (bottom). White, blue and red indicate neutral, positive or negative surface charges. (C) Residues of the Nop1-binding interface conserved in Nop56 and Nop58 mapped on the Nop58$^{1-155}$ homology model. The eukaryotic-specific insertion between β2 and β3 of the Nop58 N-terminal domain is in beige.