Elucidation of structure–function relationships in Methanocaldococcus jannaschii RNase P, a multi-subunit catalytic ribonucleoprotein

Hong-Duc Phan1,2,3,†, Andrew S. Norris1,3,4,†, Chen Du1,3,4, Kye Stachowski1,3, Bela H. Khairunisa5,6, Vaishnavi Sidharthan1,2,3, Biswarup Mukhopadhyay5, Mark P. Foster1,2,3, Vicki H. Wysocki1,2,3,4,* and Venkat Gopalan1,2,3,*

1Department of Chemistry and Biochemistry, Columbus, OH 43210, USA, 2The Ohio State Biochemistry Program, Columbus, OH 43210, USA, 3Center for RNA Biology, Columbus, OH 43210, USA, 4Resource for Native Mass Spectrometry-Guided Structural Biology, The Ohio State University, Columbus, OH 43210, USA, 5Department of Biochemistry, Virginia Tech, Blacksburg, VA 24061, USA and 6Genetics, Bioinformatics, and Computational Biology Program, Virginia Tech, Blacksburg, VA 24061, USA

Received May 26, 2022; Editorial Decision June 20, 2022; Accepted June 27, 2022

ABSTRACT

RNase P is a ribonucleoprotein (RNP) that catalyzes removal of the 5′ leader from precursor tRNAs in all domains of life. A recent cryo-EM study of Methanocaldococcus jannaschii (Mja) RNase P produced a model at 4.6-Å resolution in a dimeric configuration, with each holoenzyme monomer containing one RNase P RNA (RPR) and one copy each of five RNase P proteins (RPPs; POP5, RPP30, RPP21, RPP29, L7Ae). Here, we used native mass spectrometry (MS), mass photometry (MP), and biochemical experiments that (i) validate the oligomeric state of the Mja RNase P holoenzyme in vitro, (ii) find a different stoichiometry for each holoenzyme monomer with up to two copies of L7Ae, and (iii) assess whether both L7Ae copies are necessary for optimal cleavage activity. By mutating all kink-turns in the RPR, we made the discovery that abolishing the canonical L7Ae–RPR interactions was not detrimental for RNase P assembly and function due to the redundancy provided by protein–protein interactions between L7Ae and other RPPs. Our results provide new insights into the architecture and evolution of RNase P, and highlight the utility of native MS and MP in integrated structural biology approaches that seek to augment the information obtained from low/medium-resolution cryo-EM models.

INTRODUCTION

The architecture, dynamics and function of ribonucleoprotein (RNP) complexes, which play a key role in control of gene expression (e.g. splicing, translation), are regulated by both RNA–protein and protein–protein interactions (1,2). RNA-binding proteins in RNPs recognize their cognate RNAs using a primary sequence, a structural motif, or a post-transcriptional modification. Questions regarding the genesis of large RNPs (e.g., ribosome, RNase P) are inspired in part by observations that the functions of these RNPs depend on the exquisite cooperation between multiple subunits that were likely accrued over time. The evolutionary trajectories of these RNPs, however, are not easily reconstructed by molecular paleontology. Here, we report on adventitious findings on archaeal RNase P, an RNP that catalyzes removal of 5′ leader from precursor transfer RNAs (pre-tRNAs) (3), to illustrate the idea that a protein could be entrenched in an RNP through protein–protein interactions and independent of its native RNA-binding properties.

The RNase P RNP includes a single catalytic RNA subunit (RNase P RNA; RPR) and variable numbers of protein subunits (RNase P Proteins; RPPs) (3–8). In bacteria, the RNase P RNP comprises an RPR and an RPP. The bacterial RPRs are classified as type A, type B and type C (5,9), based on their secondary structure. In contrast to the simple make-up of the bacterial version, eukaryotic RNase P is comprised of one RPR and up to ten RPPs (5,10–12). Archaeal RNase P is intermediate in complexity, comprising one RPR and up to five RPPs, with the RPRs classified as type A, type M, and type P (13–15). The archaeal type A

†To whom correspondence should be addressed. Tel: +1 614 292 1332; Email: gopalan.5@osu.edu
Correspondence may also be addressed to Vicki H. Wysocki. Email: wysocki.11@osu.edu

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

© The Author(s) 2022. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
RPR resembles the bacterial RPR and similarly can process the pre-tRNA in vitro even without protein subunits (16–18). The type M RPR, which lacks some of the structural elements in bacterial RPR and archaeal type A RPR, requires the cooperation of protein cofactors to process the pre-tRNA in trans (16,18,19). The type P RPR, despite being only two-thirds of the length of the type A/M RPRs, is capable of pre-tRNA cleavage in vitro (15). Unlike the ~10-subunit eukaryotic RNase P that has been difficult to reconstitute in vitro (20) for recent success with yeast RNase P, the six-subunit archaeal relative has been assembled from recombinant subunits (16,17,19,21–23). Thus, archaeal RNase P is a useful model to study protein-aided RNA catalysis and to understand how structural alterations in an RNP even within one domain of life may affect its subunit composition, assembly, and function.

Previous biochemical studies have delineated the functional contribution of each archaeal RPP. POP5 forms a heterotetramer with RPP30, where the POP5 dimer is flanked by two copies of RPP30 (24–26). The POP5-RPP30 pair enhances the RPR’s cleavage rate (16,17,19). Formation of the heterotetramer was shown to be important for the function of an archaeal type A RNase P since mutations that disrupt POP5 dimerization lowered the activity of the assembled holoenzyme (24). RPP21 forms a heterodimer with RPP29 (RPP21-RPP29), and this binary complex improves substrate binding by the catalytic RPR (16,17,19,27–29). L7Ae, which binds kink-turns in archaeal RPRs (30–32), raises the reaction’s $T_{opt}$ and brings the $k_{cat}/K_M$ close to the value obtained with the native enzyme (21,23).

While NMR, crystallography, and mass spectrometry (MS) studies provided insights into structures of individual or binary RPPs, binding sites of RPPs in RPR, and stoichiometry of RPPs in partial complexes (24,26–28,33,34), the archaeal RNase P holoenzyme structure remained intractable like other large RNPs that could not be crystallized (35). This limitation has recently been overcome by advances in cryogenic-electron microscopy (cryo-EM) as evidence from the many RNP structures (11,36–40), including a few RNase P variants (10,11,41). The archaeal RNase P from Methanocaldococcus jannaschii (Mja) with and without a tRNA was solved at resolutions of 4.6 and 4.3 Å, respectively (41). The Mja (type M) RNase P is a dimer with a two-fold symmetry mediated by the Mja POP5-RPP30 tetramer at the center, with each monomer comprising a single copy of the RPR and five RPPs; the quaternary structure in vivo is unknown. The RPPs form a continuous surface linked by protein–protein interactions and are arranged neatly on one side of the RPR (Supplementary Figure S1A) (41), with the POP5-RPP30 heterodimer anchored on the RPR’s catalytic domain (responsible for pre-tRNA cleavage) and the RPP29-RPP21-L7Ae heterotrimer bound to the RPR’s specificity domain (essential for substrate recognition). The interactions between RPP29 and RPP30 string together the RPPs bound to the RPR’s catalytic and specificity domains (Figure 1A). In the specificity domain, RPP21 is flanked by L7Ae and RPP29. L7Ae, which binds to kink-turns, promotes fold-back of P12 (P, paired region). Kink-turns are characterized by a 3′-nucleotide (nt) bulge flanked by a 5′ canonical helix and a 3′ non-canonical helix that contains tandem G-A and A-G trans Hoogsteen/sugar edge pairs (42). The kink-turn, which results in an included angle of ~5° between the two helices, is stabilized by binding of L7Ae.

The cryo-EM model ofSaccharomyces cerevisiae RNase P (Figure 1B) (10) with and without a bound pre-tRNA showed a similar placement of the five homologous RPPs in archaeal/eukaryotic RNase P. The binding of yeast RPP29-RPP21-POP3 heterotrimer to the RPR’s specificity domain is remarkable given that the RPR does not have a kink-turn. POP3 (the L7Ae paralog in yeast) barely contacts the RPR and is held largely by hydrophobic and electrostatic interactions with RPP21.

The above studies motivate some questions. First, the retention of POP3 in yeast RNase P is surprising given that its RNA recognition site is absent. What features then drive the evolving architecture of large RNPs? Second, the dimeric form of Mja RNase P has not been reported for any archaeal RNase P, and our previous native MS study (31) showed two copies of L7Ae bound to the Mja RPR in contrast to the single copy reported in the cryo-EM structure (41). What are the technical/biological reasons for this observed structural variability? Here, by combining native MS, mass photometry, functional assays, and estimation of copy number in vivo, we sought to better understand subunit stoichiometry and oligomerization as well as the functional significance of protein–protein interactions in Mja RNase P. Our results provide new insights into the structure and evolution of RNase P, and highlight the importance of complementing cryo-EM with other structural methods such as native MS and mass photometry to enhance the interpretation of features found in a cryo-EM density map at low/medium-resolution.

**MATERIALS AND METHODS**

**Preparation of Mja RPR and RPPs**

Details of in vitro synthesis of Mja RPR (wild-type and mutant derivatives) as well as the overexpression and purification of Mja RPPs are provided in the Supplementary Information. Details of all the oligonucleotide primers used in this study are listed in Supplementary Table S1.

**Size-exclusion chromatography to analyze Mja RPR samples**

One-hundred μl of an in vitro transcription (IVT) reaction was set up without or with L7Ae (~17 μM final). After incubation at 37°C for 10 h, 1 μl of DNase I (10 U/μl) was added to the IVT and the reaction was incubated at 37°C for 30 min. The IVT reaction was then centrifuged (22,000 × g, 20 min, 22°C) and the supernatant was loaded on to a Superdex 200 1000 GL (GE Healthcare) column using a 100-μl sample loop. The column was run at a flow rate of 0.5 ml/min and with 50 mM HEPES-KOH (pH 7.5), 50 mM MgCl₂, 200 mM NaCl, 500 mM KCl and 50 mM potassium acetate (KOAc) as the running buffer (41).

**Assembly of Mja RPR and Mja RPPs for native MS analyses**

*Mja RPR*. To prepare the Mja RPR sample for native MS studies, 100 μM Mja RPR in RNase-free water was incubated at 50°C for 50 min and then at 37°C for 10 min.
Next, 5X assembly buffer [250 mM HEPES–KOH (pH 7.5), 250 mM MgCl₂, 1 M NaCl, 2.5 M KCl, 250 mM KOAc] was added to the sample for a final RPR concentration of 6 μM and then further incubated at 37°C for 10 min. The sample was centrifuged (20,000 × g, 10 min, 25°C) and 100 μl of the supernatant was injected onto a Superdex 200 Increase 10/300 GL column (GE Healthcare) at a flow rate of 0.5 ml/min with the 800 mM ammonium acetate (NH₄OAc), pH 7, 2 mM Mg(OAc)₂ as the running buffer. Abs₂₈₀ was used to monitor the elution, which was collected in 50-μl fractions. The Abs₂₆₀ of these fractions was measured using a Nanodrop spectrophotometer (Thermo Fisher Scientific) and the fraction containing the highest amount of RPR (∼200 nM) was used for native MS analysis.

*Mja* RPR + L7Ae. To prepare the *Mja* RPR + L7Ae sample, the *Mja* RPR was refolded as described above and then mixed with *Mja* L7Ae (6 μM RPR and 18 μM L7Ae). The mixture was incubated at 37°C for 30 min, and then at 55°C for another 30 min. The sample was centrifuged (20,000 × g, 10 min, 25°C) and 100 μl of the supernatant was injected onto a Superdex 200 Increase 10/300 GL column using the same conditions described above. The Abs₂₆₀ of each collected fraction was measured using a Nanodrop spectrophotometer, even though we recognized that there will be a contribution from L7Ae to Abs₂₆₀. Three fractions having the highest absorbance were collected and concentrated to ∼200 nM using a Vivaspin 5000 Da-cutoff centrifugal concentrator (Sartorius).

*Mja* RPR + L7Ae + other RPPs. To prepare the assemblies with the RPP binary complexes, the concentrated *Mja* RPR + L7Ae sample was either mixed with *Mja* RPP21-RPP29, *Mja* POP5-RPP30 or both *Mja* RPP21-RPP29 and *Mja* POP5-RPP30. First, *Mja* RPP21-RPP29 and *Mja* POP5-RPP30 stocks were individually diluted to 16 μM (dimer and tetramer, respectively) in 800 mM NH₄OAc and 2 mM Mg(OAc)₂. Then the three different assembly mixtures were made to a final *Mja* RPR + L7Ae concentration of 200 nM and binary complex concentration(s) of 500 nM (2.5-fold excess). Each assembly mixture was incubated at 55°C for 10 min. For the assembly with both *Mja* RPP21-RPP29 and *Mja* POP30-POP5, *Mja* RPP21-RPP29 was added first and then incubated as above prior to the addition of *Mja* POP5-RPP30. This order of addition followed the method used to prepare the holoenzyme sample for the cryo-EM study (41).

For the studies with RPRₕ₉₂₃, all the components were dialyzed into 800 mM NH₄OAc and 2 mM Mg(OAc)₂. RPRₕ₉₂₃ was refolded as described above for the wild-type. To prepare the RPRₕ₉₂₃+RPP assemblies, 1 μM RPRₕ₉₂₃ was assembled with either 2 μM RPP21-RPP29 or with 2 μM RPP21-RPP29 and 4 μM L7Ae. The assemblies were incubated for 10 min at 55°C before native MS analysis.

**Native MS analysis of the *Mja* RNase P assemblies**

Samples (3–5 μl) that were prepared as described above were directly loaded into in-house pulled nanoESI emitters. The electrospray voltage was applied through a platinum wire directly in contact with the sample solution and the voltage was adjusted between 0.6 and 1 kV and then held constant throughout each acquisition. Analyses were carried out using a Q Exactive Ultra-High Mass Range (UHMR) mass spectrometer (Thermo Fisher Scientific) modified with a SID (surface-induced dissociation) device that was set to transmission mode unless being used for
activation. Instrument settings were the following: capillary temperature 250°C, S-lens RF level 200, S-lens voltage 21 V, injection Flapatole Offset 5 V (or 8 V only for the RPR-mK72 studies), Bent Flapatole 2 V (or 6 V only for the RPR-mK72 studies), ion transfer target high m/z (Injection Flapatole RF amplitude 698 V and frequency 538 kHz, Bent Flapatole RF amplitude 936 V and frequency 1077 kHz), Extended trapping 1, resolution 3125 or 6250 (at 400 m/z), and ultra-high vacuum 3–5 E−10 mbar. CID (collision-induced dissociation) voltages for IST (in-source trapping) and HCD (higher-energy collisional dissociation) were used to obtain the best resolved spectra. Voltages in the range of 60–120 V were applied to both IST and HCD. The quadrupole was used for charge state isolation and increased HCD voltages (150–240 V) were used for deaddingucting and dissociating complexes. Spectra with SID voltages in the range of 95–115 V were also collected (see Supplementary Information for additional details on data analysis).

Mass spectra analysis

A combination of manual calculations and UniDec (43) was used to identify charge states, to calculate masses, and to annotate spectra. The individual components were deconvolved with UniDec, and the mass of each RPP monomer was verified to ±1 Da. All RPP samples had some small amounts of adducts like sodium and possible proteoforms. For example, there was an identified POP5 proteoform that was missing C-terminal residues. While mostly minor, these forms of heterogeneity contribute to the complexity of spectra collected for the different assemblies. Unlike the RPPs, the RPR was poorly resolved even with CID. In UniDec, a large Point Smooth Width (100) was used to account for the broad RPR charge states. The determined mass was ∼500 Da larger than the expected RPR mass with a mass error between charge states of ∼300 Da and peak FHWM of ∼700 Da. Thus, the RPR is likely the key contributor to the peak broadening and accounts for most of the positive mass shift in the assembly spectra.

Mja RNase P activity assays

The activity of Mja RNase P was measured in different buffer conditions depending on the experiment. For studying the activity of Mja RNase P under the conditions used for native MS, we used 0.8 M NH4OAc, 2 mM Mg(OAc)2; for comparison, we used HEPES-N2-hydroxyethylpiperazine-N2-ethanesulfonic acid-containing buffer at pH 7 [50 mM HEPES-KOH (pH 7), 0.8 M NH4OAc, 2 mM Mg(OAc)2]. For studying the activity of kink-turn mutants and Mg2+-dependent activity, we used 50 mM HEPES-KOH (pH 8), 800 mM NH4OAc and 2 mM Mg(OAc)2. Typically, 10 nM refolded Mja RPR was assembled sequentially with 100 nM Mja L7Ae (or the specified concentrations in the case of Figure 5), 100 nM Mja RPP21-RPP29, and Mja RPP30-POP5 at 55°C for 10 min. The cleavage reaction was initiated by adding 500 nM Escherichia coli pre-tRNA 5′, a trace amount of which was labeled using 5′-γ-[32P]-ATP (PerkinElmer, Shelton, CT) and T4 PNK (New England Biolabs). At defined time points, 4 μl aliquots were withdrawn and quenched with 10 μl loading buffer [7 M urea, 20% (v/v) phenol, 0.2% (w/v) SDS, 5 mM EDTA, 0.05% (v/v) bromophenol blue and 0.05% (w/v) xylene cyanol]. Reaction products were separated on a denaturing PAGE gel [10% (v/v) polyacrylamide, 7 M urea], and visualized using the Typhoon Phosphorimager (GE Healthcare). The ImageQuant (GE Healthcare) software was employed to quantify the extent of the cleavage reaction, and the turnover number was calculated in Excel. The final gel images were prepared with Fiji (44).

Mass photometry analysis of Mja RNase P assembly

Mja RPR in RNase-free water was refolded by incubating at 55°C for 50 min and then at 37°C for 10 min. An equal amount of 2× native MS solution [1.6 M NH4OAc, 4 mM Mg(OAc)2] was subsequently added and the sample incubated at 37°C for 30 min. One μM of refolded Mja RPR was assayed with five Mja RPPs (4 μM L7Ae + 2 μM RPP21-RPP29 + 2 μM RPP30-POP5 in this order), or with two Mja RPPs (4 μM L7Ae + 2 μM RPP30-POP5 in this order), or with one Mja RPP (2 μM RPP30-POP5). These assemblies were incubated at 55°C for 10 min and then kept at 22°C until mass photometry analysis. The concentration of the final assemblies is labeled based on the concentration of the RPR used in the reconstitution. The samples were diluted to different concentrations in 0.8 M NH4OAc and 2 mM Mg(OAc)2 (same conditions as those used for the native MS) before loading on the mass photometer. The mass photometer (Refeyn OneMP) was calibrated using proteins: lens epithelium-derived growth factor, LEDGF (60.4 kDa); bovine serum albumin, BSA (68.5 kDa); HIV integrase (80.4 kDa, monomer; 160.8 kDa, dimer); dihydroorotate dehydrogenase B (237 kDa); apoferritin (480 kDa); thyroglobulin (670 kDa). Details on how to set up and load the sample onto the instrument are described elsewhere (45). The expected masses were calculated using the OligoAnalyzer for the RPR and Protparam for the RPPs (46). The data were binned into histograms and the density plot was fitted with calculated average mass and standard deviation using Refeyn DiscoverMP software. The measurements were repeated three times, and the replicate data revealed a similar trend.

RESULTS

Mja RPR obtained from in vitro transcriptions performed with or without L7Ae adopts similar structural states

Our overarching objective was to firmly establish the makeup of archaeal RNase P to parse the functional contributions of individual subunits and allow comparative studies with its eukaryotic cousin. In this regard, the cryo-EM structure of Mja RNase P (41) provides an excellent foundation. However, we first sought to understand the basis for some unexpected findings from this previous study. To assemble the Mja RNase P holoenzyme for the cryo-EM study, Wan et al. found by size-exclusion chromatography (SEC) that in vitro transcribed and purified Mja RPR formed soluble aggregates. They solved this problem by including Mja L7Ae in the IVT reaction, a change motivated
by the expectation that the chaperone action of \textit{Mja} L7Ae would promote an \textit{Mja} RPR fold suitable for further assembly with the remaining four \textit{Mja} RPPs (41). This payoff is surprising since functional RNase P holoenzymes have been reconstituted \textit{in vitro} using the archaeal type A/M RPR and \textit{only} four RPPs (POP5–RPP30 and RPP21–RPP29) (16,17,19,21–23). Since L7Ae was shown previously to be not obligatory for assembly of types A and M archaeal RNase P, we decided to take a closer look at the need for L7Ae to seed assembly.

When we compared the size-exclusion chromatographic behavior of \textit{Mja} RPR (Figure 2A) purified from \textit{in vitro} transcription reactions that were performed either with or without \textit{Mja} L7Ae (Figure 2B), we observed no differences in the peak shapes of chromatograms obtained with these two samples. \textit{Mja} RPR purified post-IVT (in the absence of L7Ae) showed a predominant bell-shaped peak at about 12 ml. Some unresolved minor peaks at \textasciitilde 9 and \textasciitilde 10.5 ml likely reflect unfolded states of the \textit{Mja} RPR. \textit{Mja} RPR purified post-IVT (in the presence of L7Ae) also showed a bell-shaped peak at about 11 ml, eluting slightly earlier than the free \textit{Mja} RPR as would be expected from binding of \textit{Mja} L7Ae to the RPR. Again, we noticed some unresolved minor peaks \textasciitilde 9 and \textasciitilde 10 ml. When we added \textit{Mja} L7Ae to the purified \textit{Mja} RPR post-transcriptionally and assessed this assembly, we obtained the same chromatogram as the \textit{Mja} RPR sample that was transcribed in the presence of \textit{Mja} L7Ae (Figure 2B).

The above observations suggest that (i) \textit{in vitro} transcribed and purified \textit{Mja} RPR is monodisperse as judged by SEC and (ii) \textit{Mja} L7Ae can be assembled with \textit{Mja} RPR either co- or post-transcriptionally. While these findings differ from those reported by Wan \textit{et al.}, we emphasize that the RPRs used in the two studies are different. While our \textit{Mja} RPR construct has a long P1 helix interrupted by a single nt bulge in the middle, the \textit{Mja} RPR in Wan \textit{et al.} (2019) is missing the first and last seven nucleotides resulting in a P1 helix shortened by 7 bp (Figure 2A). Since the termini and sequence of the native \textit{Mja} RPR have not been mapped before, we sought to obtain this information using 5′ and 3′ RACE (rapid amplification of cDNA ends) of total RNA isolated from \textit{Mja} cells. We confirmed that the native \textit{Mja} RPR has the longer P1 helix (our construct) and a few additional non-base paired nucleotides at the two termini (Figure 2A). It is conceivable that the \textit{Mja} RPR with an abbreviated P1 [used by Wan \textit{et al}. 2019, (41)] may differ from the native RPR in its aggregation properties and dependency on L7Ae.

Our previous biochemical studies (16,19) and those used in the initial stages of this investigation used an \textit{Mja} RPR lacking the 2 and 4 nt at the 5′ and 3′ termini, respectively, in the native version (Figure 2A). Therefore, to assess if either RPR version is suitable for biochemical/structural characterization, we sought to compare their respective RNase P activity. The presence of the modest extensions at the two termini had no effect on the pre-tRNA cleavage activity (Supplementary Figure S2), mirroring a similar observation made with another type M RPR (21). Since the turnover numbers of the two \textit{Mja} RPR versions are nearly identical (6.5 and 6.7 min\(^{-1}\); Supplementary Figure S2) and we had already invested some effort with the variant lacking...
the terminal extensions, this study was conducted with a “non-native” version lacking the termini but containing a complete P1 helix [in contrast to Wan et al. (41)] (Figure 2A).

Native MS confirms the presence of the Mja RNase P dimer, with up to two copies of L7Ae per monomer

The dimeric configuration found in the cryo-EM structure of Mja (type M) RNase P has not been observed for any archaean RNase P. When the stoichiometry of Pyrococcus furiosus (Pfu, type A) RPR with Pfu POP5-RPP30 and/or Pfu RPP21-RPP29 was previously investigated by native MS, only a monomeric assembly was detected (47,48). While this difference in the quaternary arrangement of type A versus M archaean RNase P suggests structural variability in an ancient enzyme, this dissimilarity may have resulted from comparing gas phase versus vitreous state studies or from other technical reasons (e.g. different concentrations/buffer conditions). We decided to investigate this question in depth.

Also, the cryo-EM structure of Mja RNase P showed a 1:1 stoichiometry for the RPR and five RPPs (41). The finding of one L7Ae copy per RPR is inconsistent with our previous native MS study showing that two copies of L7Ae bind to two kink-turns named KT2 and KT3 in the P12 region of the Mja RPR (Figure 2A) (31). The only L7Ae copy in the cryo-EM structure is modeled at the KT3 position, leaving KT2 unbound. Due to the low resolution of the structure at the periphery, which includes the P12 region and L7Ae, we examined the electron density map and found that we could model a copy of L7Ae into the density located at KT2, both in the absence or presence of substrate (Supplementary Figure S1). This observation prompted us to re-investigate L7Ae stoichiometry in the Mja RNase P holoenzyme.

To examine the stoichiometries of the Mja RPR assembly with the Mja RPPs en route to the holoenzyme, we turned to native MS. First, each component required for assembly was purified and then characterized by gel electrophoresis and native MS (Supplementary Figure S3). In native MS, we noticed that for the RPR individual peaks within the charge state distribution were broad, even when high collision-induced dissociation (CID) voltages were used for de-adducting. The calculated mass even with CID activation was ~500 Da larger than expected. However, with the available resolution, it cannot be determined if this mass shift is due either to formation of adducts with cations or to non-template addition of nucleotides to the 3’ end (49,50). Magnesium adduction seems likely because Mg2+ is present in the IVT reaction and is required for folding and activity of the RPR (48). For the Mja L7Ae monomer, a monomer was observed as expected (Supplementary Figure S3). For the Mja RPP21-RPP29 sample, the expected heterodimer was observed, however, there was some heterotrimer containing two copies of each RPP, and free Mja RPP21 (Supplementary Figure S3). The Mja POP5-RPP30 sample was observed to form the expected heterotrimer containing two copies of each RPP (Supplementary Figure S3).

Because solutions used for native MS require minimal amounts of non-volatile components, the Mja RNase P assemblies for native MS analysis were prepared in an ammonium acetate solution that does not contain buffering components like HEPES, which are typically used in the biochemical assay to measure pre-tRNA cleavage activity. To ensure that RNase P samples generated for native MS are functional assemblies, we determined the activity of the reconstituted Mja RPR + 5 RPPs in the absence and presence of HEPES with otherwise identical solutions containing 800 mM ammonium acetate and 2 mM Mg2+. We were encouraged to find that the Mja RNase P holoenzyme is functional in the native MS solution albeit ~2-fold weaker than in the assay containing HEPES (Supplementary Figure S4), and thus proceeded to the MS studies.

Mja RPR + L7Ae. First, we assembled the Mja RPR with Mja L7Ae and purified the resultant complex by SEC. The peak fractions were collected and concentrated to ~200 nM for native MS analysis. The native mass spectrum of the Mja RPR+L7Ae was not homogeneous, but rather contained free L7Ae (not shown in mass spectrum range), free RPR, RPR + L7Ae, and RPR + 2 L7Ae (Figure 3A). When RPR + 2 L7Ae, the dominant species in the spectrum, was quadrupole-isolated and subjected to surface-induced dissociation (SID), we could validate its composition (Supplementary Figure S5). The relative abundances of species in the full mass spectrum that we obtained here for Mja RPR + L7Ae closely resemble the previously published Mja RPR + L7Ae spectra (31).

Mja RPR + L7Ae + RPP21·RPP29. Next, we assembled the Mja RPR+L7Ae sample individually with the two different binary complexes for native MS analysis. In addition to the species that we observed in the Mja RPR + L7Ae spectrum, the addition of Mja RPP21-RPP29 led to new complexes primarily RPR + 2 L7Ae + RPP21 + RPP29 and RPR + 2 L7Ae + RPP21 (Figure 3B, free L7Ae not shown). This spectrum clearly showed that RPR + 2 L7Ae was the RNP platform used to generate the new complexes with RPP21-RPP29 (or RPP21). Additional validation of the complex assignments was achieved with isolation and fragmentation by CID in the HCD cell (Supplementary Figure S6).

Mja RPR + L7Ae + POP5·RPP30. Addition of Mja POP5-RPP30 to the Mja RPR + L7Ae sample followed by native MS analysis revealed RPR + 2 L7Ae + 2 POP5 + 2 RPP30 and RPR + 2 POP5 + 2 RPP30 in addition to the species observed in the Mja RPR + L7Ae sample (Figure 3C). Although we initiated the assembly with Mja RPR + L7Ae, it is possible that the RPR + 2 POP5 + 2 RPP30 complex was generated from some residual free RPR, which was present due to either incomplete SEC separation or dissociation of L7Ae. To support the assignment of RPR + 2 POP5 + 2 RPP30 (n complex), we quadrupole-isolated a single-charge state and fragmented the species by CID in the HCD cell. This approach was successful in producing a very small amount of n-POP5 complex (i.e., the n-mer from which a single copy of POP5 was dissociated) (Supplementary Figure S7).

Importantly, we observed signal in the 8500–11000 m/z range of the full mass spectrum that contained broad peaks
Figure 3. Full native mass spectra of the different *Mja* RNase P assemblies. Mass spectra are annotated with colored shapes that mark each charge state distribution and the charge state of the most abundant peak in each distribution. The oligomer assignments for each charge state distribution are displayed next to the corresponding-colored shape, with the same labeling scheme for all four spectra. Mass spectra of (A) *Mja* RPR+L7Ae (SEC fraction), (B) *Mja* RPR+L7Ae assembled with *Mja* RPP21-RPP29, (C) *Mja* RPR+L7Ae assembled POP5-RPP30 and (D) *Mja* RPR+L7Ae assembled with both RPP21-RPP29 and POP5-RPP30. In (C, D), isolation and fragmentation spectra of the 8500–11000 *m/z* region are also shown. The CID voltage is ∼150–180 V by HCD. Oligomer assignments are displayed below the spectra with the corresponding-colored shapes used to annotate each charge state distribution. The pink boxes represent the quadrupole isolation ranges used for the fragmentation spectra.

with unresolved overlapping charge states. The whole 8500–11000 *m/z* range was quadrupole isolated and subjected to increased HCD voltage (>150 V) that resulted in better resolved charge states and produced fragments that are the typical CID-type (n–1)-mer fragments (Figure 3C). Although the dissociated 1-mer itself was not detected, the difference in the *n* and (n–1)-mer was determined to be consistent with the dissociation of a single L7Ae copy. The resulting assignments for the dominant n complexes with accompanying n-L7Ae dissociation products were 2 RPR + 2 POP5 + 2 RPP30 + 4 L7Ae and 2 RPR + 2 POP5 + 2 RPP30 + 4 L7Ae. Overall, these data support the formation of a dimeric RPR complex mediated by the *Mja* POP5-RPP30 heterotetramer.

*Mja* RPR + L7Ae + POP5-RPP30 + RPP21-RPP29. Finally, we assembled the *Mja* RPR+L7Ae sample with the sequential addition of *Mja* RPP21-RPP29 and *Mja* POP5-RPP30 for native MS analysis. The resulting spectrum showed a combination of the RPR-containing complexes that were observed in the samples when *Mja* RPR + L7Ae was assembled with either *Mja* RPP21-RPP29 or *Mja* POP5-RPP30 (Figure 3D). Notably, there was no or almost no signal for the RPR + 2 L7Ae + RPP21 + RPP29 and RPR + 2 L7Ae + 2 POP5 + 2 RPP30 complexes suggesting their potential conversion to the completely assembled holoenzyme.

Similar to the spectrum from the *Mja* RPR + L7Ae + POP5-RPP30 sample, there was signal in the 8500–11 000...
Figure 4. Mass photometry (MP) analysis of *Mja* RNase P. (A) *Mja* RPR (1 μM) was refolded and diluted in 800 mM NH₄OAc and 2 mM Mg(OAc)₂ before being loaded onto the mass photometer (Refeyn). (B–D) *Mja* RPR (1 μM) was refolded and reconstituted with a 4-fold molar excess of *Mja* L7Ae (4 μM), a two-fold molar excess of *Mja* RPP21-RPP29 (2 μM) and *Mja* POP5-RPP30 (2 μM) in 800 mM NH₄OAc and 2 mM Mg(OAc)₂. The samples were diluted before MP analysis. The expected masses for the holoenzyme monomer and dimer are based on the mass of [1RPR + 1RPP21 + 1RPP29 + 1POP5 + 1RPP30 + 2L7Ae] and [2RPR + 2RPP21 + 2RPP29 + 2POP5 + 2RPP30 + 4L7Ae] species, respectively. Although we obtained triplicate measurements, the data from only a single measurement are shown here (see Supplementary Table S2 for the data from the other runs). (E, F) *Mja* RPR (1 μM) was refolded and reconstituted with either a four-fold molar excess of *Mja* L7Ae (4 μM) and *Mja* POP5-RPP30 (2 μM) or only a 2-fold molar excess of *Mja* POP5-RPP30 (2 μM). The assemblies were in 800 mM NH₄OAc and 2 mM Mg(OAc)₂. The expected masses for the monomer and dimer are based on the mass of [1RPR + 1POP5 + 1RPP30 ± 2L7Ae] and [2RPR + 2POP5 + 2RPP30 ± 4L7Ae] species, respectively. In all cases, the observed masses for each population were calculated by the DiscoverMP software (Refeyn) and the % count for each population is shown in parentheses.
in the isolated m/z range were better resolved and clearer (n−1)-mer distributions were observed. There was a clear n-L7Ae fragment that is best matched to 2 RPR + 2 POP5 + 2 RPP30 + 3 L7Ae + 2 RPP21 + 1 RPP29. However, the unique charge states for the complete n complex (containing 4 L7Ae) were not easily identified. The theoretical mass of this n complex has charge state overlap with the +35 and +34 belonging to the 2 RPR + 2 POP5 + 2 RPP30 + 4 L7Ae + 2 RPP21 + 2 RPP29 complex (red stars in Figure 3D). The replicate analyses support the presence of these main complexes but there may be other species given the presence of peaks that could not be assigned. Despite the heterogenous assembly intermediates, which are likely formed en route to the final Mja RNase P holoenzyme and whose in vivo significance is unclear, the native MS data support the presence of a dimeric RNase P holoenzyme with up to two copies of L7Ae per monomer.

**Dimerization of Mja RNase P is concentration dependent and is mediated by POP5-RPP30**

While native MS provided accurate information on the composition of Mja RNase P, it is possible that the relative abundance of the observed species may be biased by differences in their ionization efficiency and ion transmission efficiency. Therefore, to better understand the abundance and possible concentration dependence of the Mja RNase P dimer species in solution, we turned to mass photometry (MP; Figure 4 and Supplementary Table S2), which has been used effectively to determine the oligomeric state of proteins, either free or complexed to other macromolecular species (45,48,51,52).

One μM Mja RPR was assembled with 4 μM L7Ae, 2 μM POP5-RPP30, and 2 μM RPP21-RPP29, and the resultant sample was diluted to 10, 100, 250 and 500 nM (based on the initial RPR concentration) before analysis by MP. We first examined Mja RPR alone and observed a mass range of 87 ± 17 kDa, which agrees with the expected mass of 88 kDa (Figure 4A). For Mja RPR assembled with five Mja RPPs, the expected mass for the holoenzyme monomer is 183 kDa (RPR + POP5 + RPP30 + RPP21 + RPP29 + 2 L7Ae) and dimer is 366 kDa (2 RPR + 2 POP5 + 2 RPP30 + 2 RPP21 + 2 RPP29 + 4 L7Ae). At 10 nM, the holoenzyme sample showed two major species with mass ranges of 82 ± 22 kDa (36% of total counts) and 137 ± 18 kDa (43% of total counts), the latter likely representing partial assemblies. We also observed a minor species around 400 kDa but could not quantitate the same (data not shown). As the concentration was increased from 100 to 250 to 500 nM, there was a noticeable decrease in the partial assemblies and a concomitant increase in the fraction of dimer (31 to 46 to 74%) with mass ranges of 362 ± 34, 348 ± 35 and 404 ± 87 kDa, respectively (Figure 4B–D). These MP data validate the dimeric arrangement of the Mja RNase P holoenzyme in solution and suggest that its formation is concentration dependent.

We showed that Mja L7Ae is not required for producing monodisperse Mja RPR by IVT (Figure 2). However, the Mja RNase P holoenzyme assembly was initiated by forming a complex between Mja RPR with Mja L7Ae both in the earlier cryo-EM study (41) and in our native MS analysis. We detected in our native MS studies 2 RPR + 2 POP5-RPP30 (i.e., no L7Ae), an observation that suggests the ability of Mja RPR to form complexes independent of L7Ae. We tested this idea using MP (Figure 4E and F). We compared the assembly between Mja RPR with POP5-RPP30 either in the absence or presence of L7Ae. The Mja RPR+L7Ae+POP5-RPP30 sample exhibited a predominant species with a mass range of 318 ± 68 kDa (89% of total counts) consistent the expected mass of the dimer (2 RPR + 2 POP5 + 2 RPP30 + 4 L7Ae, 313 kDa). The Mja RPR+POP5-RPP30 sample showed a predominant species with a mass range of 255 ± 42 kDa (82% of total counts) in agreement with the expected mass of the dimer (2 RPR + 2 POP5 + 2 RPP30, 262 kDa). Thus, Mja POP5-RPP30 can assemble with Mja RPR to form a dimer even in the absence of L7Ae. Our findings are consistent with the earlier observation (41) based on SEC analyses that addition of POP5-RPP30 to Mja RPR+L7Ae promotes dimer formation.

**Cellular concentration of Mja RPR**

Results from our MP studies indicate that at 250 nM Mja RNase P, the holoenzyme (RPR + 5 RPPs) exists as ~50% dimer *in vitro*. Any extension of this finding to the *in vivo* scenario requires an estimate of the cellular concentration of Mja RNase P. Thus, we sought to determine the level of expression of Mja RPR (a proxy for the holoenzyme). Mja was grown under anaerobic conditions in mineral salt media supplemented with H2 + CO2 and with sodium sulfide as the reductant (see Supplementary Information). The cell density was determined using a Petroff-Hausser bacteria counter, and the diameter range was determined by scal- ing with respect to the edges of the counter grid (50 μm) of the Petroff-Hausser counter.

We harvested Mja cells at mid-log phase, where we expected the RPR copy number to be maximal during growth in culture. We isolated total RNA from these cells and determined the RPR amount experimentally using a TaqMan probe-based, quantitative RT-PCR assay. *In vitro* transcribed Mja RPR was used to generate the standard curve for calibration (Supplementary Figure S8). The qRT-PCR experiments were performed using two biological replicates (independent Mja cultures) and with three different dilutions of each cDNA batch that was generated. Since Mja cells are spherical (53), we calculated the volume based on the experimentally determined diameter values. Given the cell density, the cellular volume (based on a diameter of 1.6 μm), and the RPR amount experimentally determined, we calculated Mja RPR to be 50–63 nM (~100–120 copies per cell; Supplementary Table S3). If all these Mja RPR molecules in a cell assemble with the RPPs to form the holoenzyme, then Mja RNase P (~50 nM) is likely to exist predominantly as a monomer (as judged by our MP data; Figure 4). However, we emphasize that even a 2- to 4-fold change in this estimate would imply that at least a fraction of cellular Mja RNase P exists as a dimer. We consider a few technical and biological reasons why such a possibility merits consideration.

First, it is possible that thus far unidentified post-transcriptional RPR modifications may have decreased the amount of cDNA product formed by the reverse transcri-
tase. Second, the RPR (and RPP) copy number may change depending on growth or stress conditions. For example, the *E. coli* RPR was found to vary from 80 to 1060 molecules per cell at 0.4 and 2.7 doublings per hour, respectively; this change under different growth conditions corresponds to a shift from 250 to 680 nM (54). Finally, crowding in the cellular milieu is likely to affect macromolecular equilibria (55). Moreover, if *Mja* RNase P exists in a phase-separated microenvironment under certain physiological conditions, then our calculations based on the entire spherical volume of the *Mja* cell will need revision.

One L7Ae copy suffices for the optimal cleavage activity of the *Mja* RNase P holoenzyme

Although L7Ae is not needed to either help fold during *Mja* RNase P reconstituted with either RPR (WT) or its kink-turn mutant derivatives. Ten nM RPR (WT, mKT2, mKT3, mKT23) was assembled with 100 nM RPP21-RPP29, 100 nM POP5-RPP30, and varying amounts of L7Ae (0, 100, 250 or 500 nM). Cleavage assays were performed in 50 mM HEPES–KOH (pH 8), 800 mM (NH₄)OAc, 2 mM Mg(OAc)₂, using as substrate 500 nM *E. coli* pre-tRNA³⁵spliced with a trace amount of 5'-³²P-labeled *E. coli* pre-tRNA³⁵. Data from three independent measurements were used to determine the mean ± standard deviation for the reported turnover numbers.

One L7Ae copy suffices for the optimal cleavage activity of the *Mja* RNase P holoenzyme

Although L7Ae is not needed to either help fold during *Mja* RNase P in vitro transcription or for assembly of the RPR with other RPPs, we previously reported that *Pfu* (type A) L7Ae decreased by 3-fold Kₘg²⁺ (the Mg²⁺ concentration required for 50% optimal activity) of the *Pfu* RNase P holoenzyme (31). We first investigated if this attribute was also shared by *Mja* (a type M) RNase P. Indeed, addition of *Mja* L7Ae to *Mja* RPR + 4 RPPs at 55°C decreased Kₘg²⁺ from 9 to 1.7 mM, and increased the maximum turnover for cleavage of pre-tRNA³⁵ by 2.5-fold (Supplementary Figure S9). The fold-increase in activity is as much as 20-fold at 1 or 2 mM Mg²⁺ indicating the importance of L7Ae at near-cellular [Mg²⁺]. Since L7Ae is distal from the cleavage site and even the substrate-binding site in *Mja* RNase P (41), it is unclear how L7Ae contributes to *Mja* RNase P catalysis. A first step to unravel the mechanism of action of L7Ae is to investigate how many copies are needed for its role in RNase P catalysis.

We showed previously (21,31) that archaeal type M RPRs have two kink-turns (KT2 and KT3) that comprise an unusual double kink-turn motif because the NC helix of KT2 also serves as the C helix of KT3; these are not two tandem kink-turns (Figure 2). Since two copies of *Mja* L7Ae bind to *Mja* RPR through KT2 and KT3, we set out to understand how each KT-bound L7Ae copy affects the optimal cleavage activity of the holoenzyme in vitro. We leveraged *Mja* RPRmKT2, *Mja* RPRmKT3, and *Mja* RPRmKT23, three mutants in which the two sheared G·A pairs of KT2, KT3, and both, respectively, are replaced with two Watson-Crick G·C pairs (Supplementary Figure S10). Native MS was used previously to demonstrate that these mutations disrupt the ability of L7Ae to bind one kink-turn (mKT2, mKT3) or both (mKT23) (31). We assembled 10 nM *Mja* RPR (WT or mutants) with 100 nM of each *Mja* RPP21-RPP29, *Mja* POP5-RPP30, and four different *Mja* L7Ae concentrations (0, 100, 250 and 500 nM). The pre-tRNA⁵′ cleavage activity of these assemblies was assayed under multiple-turnover conditions at 2 mM Mg²⁺, which is the same concentration used in our native MS and MP experiments. Some interesting and unanticipated findings emerged from these assays.

As might be expected for mutations that impair binding of L7Ae and not the other four RPPs, RPRmKT2, RPRmKT3 and RPRmKT23 displayed near-wildtype activity when they were reconstituted with RPP21-RPP29 + POP5-RPP30; in fact, there was a 2-fold increase in the case of RPRmKT2, presumably from the stabilization afforded by replacing G·A with G·C pairs (Figure 5). When 10 nM *Mja* RPRWT was assembled with 4 RPPs (i.e. RPP21-RPP29 + POP5-RPP30), addition of 100, 250, and 500 nM L7Ae increased the turnover number from 0.4 ± 0.01 min⁻¹ to 3.5 ± 0.8, 4.7 ± 0.9 and 5.9 ± 0.2 min⁻¹, respectively (Figure 5). With RPRmKT2, which has the L7Ae binding site at KT3 but not KT2, addition of 100, 250 and 500 nM L7Ae increased the turnover number from 0.9 ± 0.04 min⁻¹ to 5.5 ± 0.2, 6.4 ± 0.3 and 7.2 ± 0.5 min⁻¹, respectively, suggesting that only one L7Ae copy bound to KT3 suffices for optimal cleavage activity of the *Mja* RNase P holoenzyme (Figure 5). With RPRmKT3, which has the L7Ae binding site at KT2 but not KT3, addition of 100, 250 and 500 nM L7Ae increased the turnover number from 0.5 ± 0.03 min⁻¹ to 0.9 ± 0.2, 4.4 ± 0.5 and 4.8 ± 0.4 min⁻¹, respectively (Figure 5). Thus, KT2-bound L7Ae also increases the activity of the *Mja* RNase P holoenzyme to the optimal level, although a higher L7Ae concentration is needed when compared to KT3-bound L7Ae. With RPRmKT23, which lacks both KT2 and KT3, addition of 100, 250 and 500 nM L7Ae unexpectedly increased the turnover number from 0.5 ± 0.02 min⁻¹ to 1.3 ± 0.2, 2.6 ± 0.5 and 2.9 ± 0.4 min⁻¹, respectively (Figure 5). While the activity observed with RPRmKT23 assembled with five RPPs is less than that observed with RPRWT, RPRmKT2 and RPRmKT3 under the same condition, the 6-fold enhancement is quite remarkable considering the absence of kink turns in this mutant.

The unexpected effect of L7Ae observed in the RNase P holoenzyme assembled with RPRmKT23 suggests that protein–protein interactions between L7Ae and RPP21-RPP29 might help fasten L7Ae in the RNP even in the absence of RNA–protein interactions. Consistent with previous yeast two-hybrid studies (56), we could not observe a stable interaction between RPP21-RPP29 and L7Ae in the absence of RPR (not shown). However, our native MS studies showed that RPRmKT23 is assembled to-
Figure 6. Native mass spectra of partial RNase P assemblies consisting of RPR<sub>mKT23</sub> with L7Ae, RPP21·RPP29, or RPP21·RPP29 + L7Ae. Mass spectra are annotated with colored shapes that mark each charge state distribution and the charge state of the most abundant peak in each distribution is labeled. The oligomer assignments for each charge state distribution are displayed next to the corresponding-colored shape, with the same labeling scheme for all four spectra. All reconstitutions were for 10 min at 55°C before native MS analysis. Mass spectra of (A) 1 μM RPR<sub>mKT23</sub>, (B) 1 μM RPR<sub>mKT23</sub> + 1 μM L7Ae, (C) 1 μM RPR<sub>mKT23</sub> + 2 μM RPP21·RPP29 and (D) 1 μM RPR<sub>mKT23</sub> + 2 μM RPP21·RPP29 + 1 μM L7Ae (assembled in this order). In (D), we do observe a small amount of the [RPR<sub>mKT23</sub> + RPP21·RPP29 + L7Ae]<sub>2</sub>, whose significance is unclear.

together with L7Ae and RPP21·RPP29 (Figure 6), indicating RPR-dependent protein–protein interactions.

The above surprising finding with RPR<sub>mKT23</sub> motivated us to comprehensively examine L7Ae stoichiometry with the different RPR mutants in the absence and presence of RPP21·RPP29. First, although we had data from our previous native MS studies with the Mja RPR kink-turn mutants ± L7Ae (31), we repeated these measurements under our current conditions. Our new data confirm that L7Ae stoichiometry (in the absence of other RPPs) is as expected: two copies bound to RPR<sub>WT</sub>, one copy each to RPR<sub>mKT2</sub> or RPR<sub>mKT3</sub>, and no copies to RPR<sub>mKT3</sub> (Supplementary Figures S11 and S12). Second, there are interesting changes to this trend upon addition of RPP21·RPP29. There are two L7Ae copies bound to RPR<sub>WT</sub>, one copy bound to RPR<sub>mKT2</sub> or RPR<sub>mKT3</sub>, and surprisingly two copies bound to RPR<sub>mKT3</sub> (Supplementary Figures S11 and S12). Our finding that RPR<sub>mKT1</sub> is assembled with RPP21·RPP29 and two copies of L7Ae is likely due to one L7Ae copy bound to KT2 through RNA–protein interactions and another anchored near KT3 due to protein–protein interactions with RPP21·RPP29. This explanation is entirely consistent with the stoichiometry that we observed with RPR<sub>mKT23</sub> (Figure 6).

**DISCUSSION**

Here, we leveraged native MS and mass photometry, we confirmed the dimeric arrangement of the Mja RNase P holoenzyme, which was first reported based on cryo-EM studies (41). By analyzing different assemblies made of Mja RPR and Mja RPPs, dimeric species were observed in the high m/z region (8000–11000) of the mass spectrum when the reconstitutions contained POP5·RPP30 (Figure 3 and Supplementary Figure S13). Our MP studies also revealed that dimer formation is mediated by Mja POP5·RPP30 and does not require L7Ae (Figure 4). Because the Mja POP5·RPP30 alone exists as a heterotetramer, it appeared that this complex is key to formation of the holoenzyme dimer (Supplementary Figure S13). But this expectation contrasts with what we previously observed with *Pfu* (type A) RNase P (47,48). When *Pfu* RPR was assembled with *Pfu* POP5·RPP30 in 6 mM Mg<sup>2+</sup>, we observed a complex of only 1 RPR and 1 POP5 + 1 RPP30, even though *Pfu* POP5·RPP30 alone formed a heterotetramer (47). A minor signal corresponding to a complex of 1 RPR + 2 POP5 + 2 RPP30 could be extracted from the background noise at 4 mM Mg<sup>2+</sup>. Here, we did not observe 1 RPR + 1 POP5 + 1 RPP30 with Mja RNase P at 2 mM Mg<sup>2+</sup> (Figure 3). Also, our MP studies of the *Pfu* RNase P holoenzyme revealed that it exists as a monomer (48). Together, these data suggest different higher-order structures of archaeal type A and M RNase P, although the functional implications are obscure as noted previously with some other RNPs that dimerize (57). Disrupting formation of the Mja RNase P dimer modestly reduced the activity *in vitro* [(41); not quantitated, however]. The importance and proportion of the dimeric conformation *in vivo* require further study. In this regard, approaches including *ab initio* cryo-EM reconstruction maps from par-
entially purified native complexes that reflect the structure in the cellular milieu (58) should be a high priority.

We showed previously that type M RPRs have two kink-turns (KT2 and KT3) that comprise an unusual double kink-turn motif, which is highly conserved in Methanococcales (21). Moreover, we demonstrated that two copies of Mja L7Ae bind to the Mja RPR in the absence of the other four RPPs (31). Consistent with this earlier finding, results from the native MS studies here reveal that Mja RNase P holoenzyme contains up to two copies of L7Ae per RPR. However, our functional tests with the kink-turn mutants yielded both expected and surprising results. For example, addition of L7Ae to RPRmKT2 (a mutant with intact KT3) + 4 RPPs predictably increased the turnover number, consistent with the cryo-EM model which showed that L7Ae in the RPP29-RPP21-L7Ae heterotrimeric assembly was bound to KT3. Thus, KT3 was expected to be less dispensable than KT2. Unexpectedly, we found the activity trend for RPPmKT3 (lacking KT3) mirrored RPPmKT2, although at slightly higher L7Ae concentrations. A clue to understanding this unexpected result came from our finding that addition of L7Ae to RPRmKT2 (which lacks both KTs) + 4 RPPs also boosted pre-rRNA cleavage, albeit not as much as that observed with the WT (Figure 5). Thus, formation of the RPP29-RPP21-L7Ae heterotrimer appears to override the necessity for KT3 in both RPRmKT3 and RPRmKT23. We infer that L7Ae contributes to RNase P catalysis through protein–protein interactions even when its binding site (K-turn) is absent.

While the archaeal (type M) RNase P contains up to two copies of L7Ae per RPR as expected based on the double kink-turn, one copy suffices for optimal catalytic activity and even that is not dependent on the presence of a kink-turn (Figure 5). We cannot exclude the likelihood that the binding of two copies of L7Ae might be necessary for specialized functions of archaeal RNase P, or that the copy bound to KT2 might be important for initiating assembly but is absent in the final complex. The latter possibility may even account for heterogeneity in the preparation used for cryo-EM and the weaker density observed at this P12 periphery (41). Interestingly, most eukaryotic RPRs have only one kink-turn while fungal RPRs have none (11). The reduction in the copy number of L7Ae (and kink-turns) from archaeal to eukaryotic RNase P might be a hallmark of regressive evolution, i.e., loss of nonfunctional attributes during evolution (59).

The binding of L7Ae or POP3 to kink-turns would be expected to enhance formation of the RPP29-RPP21-L7Ae/POP3 heterotrimer in an RPR-dependent fashion. Such interlacing was reported with U2 snRNP, where protein–protein contacts enhanced RNA–protein interactions (60). However, in archaeal RNase P, the RPR-L7Ae interaction seems redundant presumably due to a ratchet that has already firmly rooted L7Ae in the RNP through electrostatic and hydrophobic interactions with RPP21. Such an idea is supported by the structure of S. cerevisiae RNase P, which contains POP3 (the L7Ae paralog) despite the RPR lacking kink-turns (Figure 1) (10). While POP3 lacks the two key positively-charged residues in the canonical NExxKxxxxR motif that anchors L7Ae/paralogs to the cognate kink-turn, a large interface that holds together POP3 and RPP21 appears to have been subjected to positive selection (Supplementary Figure S14).

The advances from our study also raise new questions that remain to be addressed. For instance, we observed that the interactions between L7Ae and RPP21-RPP29 occur only in the presence of RPR, suggesting that binding of RPP21-RPP29 to the RPR somehow facilitates the interaction with L7Ae. The underpinnings for this cooperativity as well as the functional gains conferred by L7Ae, even when it does not contact the RPR, remain a mystery.

SUMMARY

The use of native MS and mass photomometry as a complement to validate and to extend a low-/medium-resolution cryo-EM structure of an RNP exemplifies the value of integrative structural biology in yielding functional and evolutionary insights. Moreover, as illustrated with archaeal RNase P, intertwining RNA- and protein–protein interactions in multi-subunit RNPs offers loss-of-function safeguards in the event of a disruption of one of these interfaces. Given this coupling between RNA- and protein–protein interactions, bioinformatic inventories of RNA-binding proteins for a target RNA based solely on the presence of RNA-recognition sites may prove incomplete.

DATA AVAILABILITY

All data described in this paper are available from the corresponding authors upon request.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are indebted to Dr Lien Lai (OSU) for providing materials, guidance for experiments, and insightful comments, Dr Marcos Sotomayor (OSU) for suggesting a closer look at the raw cryo-EM data, and Dr Dalton Snyder for help with acquisition of initial MP data. We thank Dr Richard Fishel and Dr Ross Larue (OSU) for generous consent to use their mass photometer, and Dr Jennifer Ottesen for providing access to her NanoDrop spectrophotometer.

FUNDING

NIH [GM-120582 to M.P.F., V.H.W., V.G., P41 GM128577 to V.H.W., S10OD023582 to J.E.J.]; NASA [NNX13AI05G to B.M.]; Virginia Tech Agricultural Experiment Station Hatch Program [CRIS project VA160021 to B.M.]; OSU Comprehensive Cancer Center for a Pelotonia Pre-doctoral Fellowship (to H.D.P.); Virginia Tech Genetics, Bioinformatics, and Computational Biology Doctoral Program Fellowship (to B.K.). Funding for open access charge: University funds.

Conflict of interest statement. None declared.
45. Zahurancik, W.J., Norris, A.S., Lai, S.M., Snyder, D.T., Wysocki, V.H. and Gopalan, V. (2021) Purification, reconstitution, and mass analysis of archaeal RNase P, a multisubunit ribonucleoprotein enzyme. *Methods Enzymol.*, **659**, 71–103.

46. Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.C., Williams, K.L., Appel, R.D. and Hochstrasser, D.F. (1999) Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.*, **112**, 531–552.

47. Ma, X., Lai, L.B., Lai, S.M., Tanimoto, A., Foster, M.P., Wysocki, V.H. and Gopalan, V. (2014) Uncovering the stoichiometry of *Pyrococcus furiosus* RNase P, a multi-subunit catalytic ribonucleoprotein complex, by surface-induced dissociation and ion mobility mass spectrometry. *Angew. Chem. Int. Ed. Engl.*, **53**, 11483–11487.

48. Marathe, I.A., Lai, S.M., Zahurancik, W.J., Poirier, M.G., Wysocki, V.H. and Gopalan, V. (2021) Protein cofactors and substrate influence Mg2+-dependent structural changes in the catalytic RNA of archaeal RNase P. *Nucleic Acids Res.*, **49**, 9444–9458.

49. Milligan, J.F., Groebe, D.R., Witherell, G.W. and Uhlenbeck, O.C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.*, **15**, 8783–8798.

50. Krupp, G. (1988) RNA synthesis: strategies for the use of bacteriophage RNA polymerases. *Gene*, **72**, 75–89.

51. Soltermann, F., Foley, E.D.B., Pagnoni, V., Galpin, M., Benesch, J.L.P., Kukura, P. and Struve, W.B. (2020) Quantifying protein–protein interactions by molecular counting with mass photometry. *Angew. Chem. Int. Ed. Engl.*, **59**, 10774–10779.

52. Sonn-Segev, A., Belaïci, K., Bodrug, T., Young, G., VanderLinden, R.T., Schuman, B.A., Schimpf, J., Friedrich, T., Dip, P.V., Schwartz, T.U *et al.* (2020) Quantifying the heterogeneity of macromolecular machines by mass photometry. *Nat. Commun.*, **11**, 1772.

53. Mukhopadhyay, B., Johnson, E.F. and Wolfe, R.S. (2000) A novel pH2 control on the expression of flagella in the hyperthermophilic strictly hydrogenotrophic methanarchaeum *Methanococcus jannaschii*. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 11522–11527.

54. Dong, H., Kirsebom, L.A. and Nilsson, L. (1996) Growth rate regulation of 4.5 S RNA and M1 RNA the catalytic subunit of *Escherichia coli* RNase P. *J. Mol. Biol.*, **261**, 303–308.

55. Rivas, G. and Minton, A.P. (2016) Macromolecular crowding in vitro, in vivo, and in between. *Trends Biochem. Sci.*, **41**, 970–981.

56. Kifusa, M., Fukuhara, H., Hayashi, T. and Kimura, M. (2005) Protein-protein interactions in the subunits of ribonuclease P in the hyperthermophilic archaeon *Pyrococcus horikoshii* OT3. *Biosci. Biotechnol. Biochem.*, **69**, 1209–1212.

57. Bleichert, F. and Baserga, S.J. (2010) Ribonucleoprotein multimers and their functions. *Crit. Rev. Biochem. Mol. Biol.*, **45**, 331–350.

58. Ho, C.M., Li, X., Lai, M., Terwilliger, T.C., Beck, J.R., Wohlschlegel, J., Goldberg, D.E., Fitzpatrick, A.W.P. and Zhou, Z.H. (2020) Bottom-up structural proteomics: cryoEM of protein complexes enriched from the cellular milieu. *Nat. Methods*, **17**, 79–85.

59. Jeffery, W.R. (2009) Regressive evolution in Astyanax cavefish. *Annu. Rev. Genet.*, **43**, 25–47.

60. Scherly, D., Dathan, N.A., Boelens, W., van Venrooij, W.J. and Mattaj, I.W. (1990) The U2B+ RNP motif as a site of protein–protein interaction. *EMBO J.*, **9**, 3675–3681.