Quaternary arrangement of an active, native group II intron ribonucleoprotein complex revealed by small-angle X-ray scattering

Kushol Gupta[1,*], Lydia M. Contreras[2,3], Dorie Smith[4], Guosheng Qu[4], Tao Huang[3,5], Lynn A. Spruce[6], Steven H. Seeholzer[1,6], Marlene Belfort[3,4] and Gregory D. Van Duyne[1]

[1]Department of Biochemistry & Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104-6059, USA, [2]Department of Chemical Engineering, University of Texas at Austin, Austin, TX 78712, USA, [3]Wadsworth Center, NYS Department of Health, Albany, NY 12201, USA, [4]Department of Biological Sciences and RNA Institute, University at Albany, State University of New York, Albany, NY 12222, USA, [5]SUNY Downstate Medical Center, University Hospital, Brooklyn, NY 11203, USA and [6]Children's Hospital of Philadelphia Research Institute, Philadelphia, PA 19104, USA

ABSTRACT

The stable ribonucleoprotein (RNP) complex formed between the _Lactococcus lactis_ group II intron and its self-encoded LtrA protein is essential for the intron's genetic mobility. In this study, we report the biochemical, compositional, hydrodynamic and structural properties of active group II intron RNP particles (+A) isolated from its native host using a novel purification scheme. We employed small-angle X-ray scattering to determine the structural properties of these particles as they exist in solution. Using sucrose as a contrasting agent, we derived a two-phase quaternary model of the protein–RNA complex. This approach revealed that the spatial properties of the complex are largely defined by the RNA component, with the protein dimer located near the center of mass. A transfer RNA fusion engineered into domain II of the intron provided a distinct landmark consistent with this interpretation. Comparison of the derived +A RNP shape with that of the previously reported precursor intron (ΔA) particle extends previous findings that the loosely packed precursor RNP undergoes a dramatic conformational change as it compacts into its active form. Our results provide insights into the quaternary arrangement of these RNP complexes in solution, an important step to understanding the transition of the group II intron from the precursor to a species fully active for DNA invasion.

INTRODUCTION

Group II introns are catalytic RNAs found in all three domains of life (1,2). This group of mobile genetic elements provides an important model system to understand RNA catalysis because of the diversity of the chemical reactions performed in partnership with a helper protein (3). Interest in this class of introns also derives from their inferred ancestral relationship to nuclear spliceosomal introns (4–9) and retrotransposons (10) and their evolutionary impact on host genomes. Mobile group II introns specify intron-encoded proteins (IEPs) with reverse transcriptase and maturase activity, which facilitates the proper folding of the intron precursor RNA (Figure 1A) (11,12). Upon splicing, a free intron lariat is excised and remains associated with its IEP, which for the _Lactococcus lactis_ group II intron _Ll.LtrB_ exists with a stoichiometry of 2:1 protein–RNA in a stable ribonucleoprotein (RNP) assembly (Figure 1B) (13). This protein serves a central role in facilitating the genetic mobility of the intron, by virtue of a DNA endonuclease domain (14,15). Although no experimental structure is currently available for any IEP, a homology model of the 70 kD LtrA has been derived (16) and its interactions with the intron have been mapped out _in vitro_ (17). The most significant interactions between the protein and the RNA occur at domain IV, a region that is predicted to project away from the rest of the RNA core (18–20).

These findings raise fundamental questions about the role of the maturase–RNA interactions in RNP assembly and function, and the overall structural properties of group II introns. Although well-characterized on the genetic and biochemical level, biophysical and structural

---

*To whom correspondence should be addressed. Tel: +1 215 573 7260; Fax: +1 215 573 4764; Email: kgupta@mail.med.upenn.edu
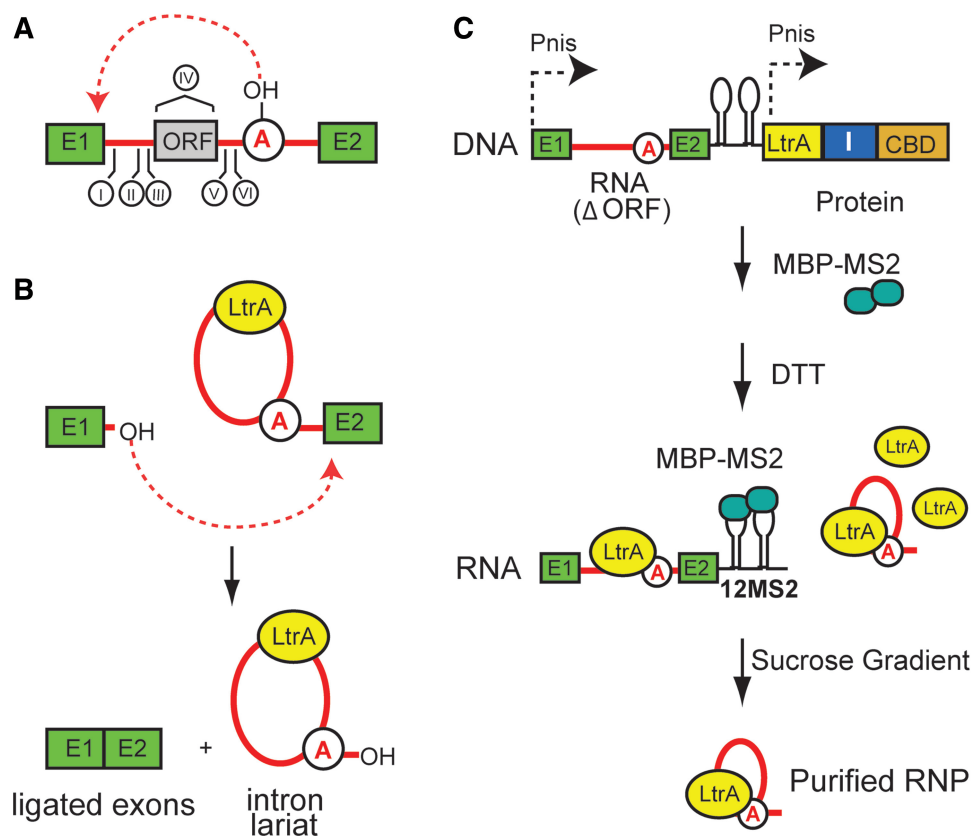
Figure 1. Domain structure, activity and purification of native group II introns. (**A**) Domain structure of a group II intron. Rendered as squares are the exons (E1 and E2) flanking the group II intron sequence. Denoted with Roman numerals are the six structural domains of the intron. Domain IV contains an ORF that encodes an IEP, also referred to as LtrA. Domain VI contains the catalytic adenosine nucleophile (circled) required for splicing. (**B**) Splicing reaction carried out by group II introns. Group II introns are excised from flanking exons through a two-step process that is catalyzed by the intron RNA. (**C**) Purification strategy. Lysates from nisin-induced cells are passed over a chitin resin column to capture LtrA–intein–CBD protein fusion and associated RNA, via a C-terminal chitin-binding domain in the LtrA fusion. I = intein. To facilitate separation (see below), MBP-MS2 (turquoise balls) is added to bind the MS2 tag in the precursor at the 3′ exon (black lollipops). Cleavage of LtrA from the intein with DTT) releases the RNP from the chitin column in both active and precursor forms. Assuming a 1:1 stoichiometry between MBP-MS2 protein (~50 kD) and its RNA-binding site (21), addition of the MS2 binding protein further increases the size differential between the 12xMS2-containing precursor (~1.4 MDa) and the spliced lariats form (~430 kD) of the RNP. This large size differential allows for successful final isolation of the active RNP from the precursor species by sedimentation, using a sucrose cushion gradient.

insights into the intact active RNP in its native form are limited. Structural studies of these large macromolecular assemblies have been challenged by the quantities of native particles typically needed for classical methods such as X-ray crystallography. While there are now experimental crystal structures of the bacterial group IIC intron RNA in different catalytic states (22,23), no structures have been reported for an intact intron RNP complex. The models available for the group IIA *Ll.LtrB* intron (20) and its IEP (16) are based on site-directed mutagenesis, chemical crosslinking and homology modeling and our only direct insight into the structure of the assembled group IIA intron have been provided by electron microscopy [EM (24)]. We recently succeeded in trapping the *Ll.LtrB* intron RNP complex in its precursor form by deleting the catalytic adenosine nucleophile in domain VI that initiates splicing. This particle was studied by sedimentation velocity, size-exclusion chromatography and cryo-EM, revealing that the intron RNP precursor (ΔA) is a large, loosely packed structure (25). The hydrodynamic properties determined for the ΔA intron contrast with those ascertained for more compactly spliced introns in a

catalytically active state (+A), suggesting that a major conformational change underlies this transition.

Fundamental parameters such as size, shape and inherent flexibility of the native, active group II intron in solution have not been reported. Here, we implement a novel purification strategy to isolate homogeneous preparations of native group II intron particles from *L. lactis*. Biochemical and biophysical analyses confirm the composition 2:1 stoichiometry and monodispersity of these isolated nucleoproteins. We combine analytical ultracentrifugation, size-exclusion chromatography in-line with multi-angle light scattering (SEC-MALS), tandem mass spectrometry (LC-MS/MS) and small-angle X-ray scattering (SAXS) with contrast variation to determine the composition and structural properties of the active group II intron RNP (+A). SAXS analyses reveal an oblate ellipsoidal particle where the RNA surrounds a centrally positioned protein component. A transfer RNA (tRNA) fusion engineered into domain II of the intron provides a landmark that corroborates our interpretation of the compositional distribution of the intron in SAXS experiments. Finally, direct comparison of the SAXS-derived

+A reconstruction with the previously reported ΔA EM result indicates that the precursor RNP undergoes a major conformational change adopting a more compact, ellipsoidal structure in the active form.

MATERIALS AND METHODS

Expression and purification of group II intron RNP particles

Details on the creation of the plasmid constructs used in this study can be found in Supplementary Methods. Transformed *L. lactis* IL1403 was grown in M17 Difico media (Becton Dickinson and Company, Franklin Lakes, NJ, USA) containing 0.5% glucose, at 30°C without aeration to an OD_{600} of 0.5–0.6 in 10 μg/ml chloramphenicol. Expression was induced with 10 ng/ml nisin for 2–3 h. The culture was then centrifuged and the pellet was washed in a buffer containing 10 mM Tris–HCl, pH 7.5, 150 mM NaCl and 1 mM ethylenediaminetetraacetic acid (EDTA). The washed cell pellet was re-suspended in 20 mM Tris–HCl, pH 8.0, 500 mM NaCl, 0.1 mM EDTA and 1 mM phenylmethanesulfonylfluoride. This cell suspension was flash-frozen in liquid N_2 and stored at −80°C until further purification.

The purification of the active RNP using the p+A(12MS2)/LtrA (p+A) and p+A(12MS2-DIItRNA)/ LtrA (p+A-DII(tRNA) constructs is similar to that reported previously for the precursor ΔA particle (25), with modifications based on MS2 tag–protein interactions to separate the precursor from the spliced intron. The MBP-MS2 binding protein (∼50 kD) used in this purification was purified as previously described (21). After the cell lysate was passed over chitin resin (New England BioLabs, Ipswitch, M.A., U.S.A.) by gravity at 4°C, the resin was washed with two column volumes of wash buffer (20 mM Tris–HCl, pH 8.0, 500 mM NaCl, 0.1 mM EDTA, 0.1% NP-40). The MBP-MS2 protein was then introduced to the resin-bound material. The resin was then washed with eight additional column volumes of wash buffer before the addition of a column volume of buffer supplemented with 40 mM dithiothreitol (DTT) to catalyze intein self-cleavage. The columns were then allowed to incubate at 4°C for 16 h to complete the reaction. The liberated group II intron RNP was eluted isocratically from the chitin resin and fractionated. To determine which fractions contained LtrA protein (and hence, intact RNP), SDS–PAGE analysis with Coomassie staining was performed. The fractions containing LtrA were pooled, concentrated and buffer-exchanged into 40 mM Tris–HCl, pH 7.5, 450 mM NaCl and 5 mM $MgCl_2$ using 50 kD MWCO Amicon Ultra-15 concentrators (EMD Millipore, Darmstadt, Germany). This material was then loaded onto an isotonic 5–20% sucrose gradient cushion and centrifuged in a SW41 rotor at 4°C and 27 000 rpm for 16 h to isolate intact RNP. Additional details on the LC-MS/MS analyses of these purified particles are provided in Supplementary Methods. The composition and calculated properties of the +A components and derivatives are provided in Supplementary Table S3.

Primer extension analysis

Primer extension was performed as described previously (26), using primer IDT1073, which is complementary to the 5′ region of the *Ll.LtrB* group II intron RNA. Primer extension products were separated by electrophoresis on denaturing 10% polyacrylamide (29:1)/8M urea/1x Tris–Borate–EDTA (TBE) gels, dried and visualized by phophorimager analysis using ImageQuant 5.2. (G.E. Healthcare Life Sciences, Piscataway, NJ, USA).

In vitro reverse splicing assay

The double-stranded DNA (dsDNA) substrate was prepared as described (27). A 1 μl aliquot of purified group II intron RNP (∼100 ng/μl) was incubated with the ^{32}P-labeled dsDNA (∼50 000 cpm per reaction) in a buffer containing 50 mM Tris–HCl pH 7.5, 10 mM KCl, 10 mM $MgCl_2$ and 5 mM DTT (with or without 0.1% NP-40 and 0.1% Tween-20) at 37°C for various times. Reactions were terminated by phenol/chloroform extraction and ethanol precipitation. Precipitated reverse splicing products were separated by electrophoresis on denaturing 5% polyacrylamide (29:1)/8M urea/1x TBE gels, dried and visualized by phophorimager analysis using ImageQuant 5.2 (G.E. Healthcare Life Sciences).

Size-Exclusion Chromatography in-line with Multi-angle Light Scattering

Absolute molecular weights of +A were determined by multi-angle light scattering coupled with refractive interferometric detection (Wyatt Technology Corporation, Santa Barbara, CA, USA) and a Superdex 200 10/300 GL column (G.E. Healthcare Life Sciences). The column was equilibrated in 20 mM Tris pH 7.4, 200 mM KCl and 5 mM $MgCl_2$, at room temperature. This analysis and Stokes radius (R_s) determinations were performed as previously described (28). For determination of molecular weight of the RNP complex, a mass-averaged dn/dc of 0.182 ml/g was used to account for the composition of the two components of the particle.

Analytical ultracentrifugation

Sedimentation velocity ultracentrifugation experiments were performed at 4°C with an XL-A analytical ultracentrifuge (Beckman-Coulter, IN, USA) and a TiAn60 rotor with two-channel charcoal-filled epon centerpieces and quartz windows. Samples were analyzed at concentrations of 7–14 ng/μl, as estimated by UV absorbance. Complete sedimentation velocity profiles were recorded every 30 s at ABS_{260} nm for 50–200 boundaries at 28 000 rpm. Data were fit using the Lamm equation as implemented in the program SEDFIT (29). After optimizing meniscus position and fitting limits, the sedimentation coefficient (S) and best-fit frictional ratio (f/f_0) was determined by iterative least squares analysis. Analyses were carried out in buffer composed of 20 mM Tris pH 7.4, 200 mM KCl and 5 mM $MgCl_2$. The solvent density (ρ = 1.0086 g/ml) and viscosity (η = 0.01007 poise) were derived from chemical composition by the program SEDNTERP (30). A partial-specific volume (v) value of 0.630 cm^3/g was used

for the +A particle, as previously estimated (25). In these experimental conditions, the pressure encountered by the particles range from ~25 to 50 bar between the middle and bottom of the cell, as calculated by SEDFIT. Theoretical data for simulation of particles of given mass but different shapes were generated using the SEDFIT program.

SAXS data collection

X-ray scattering data were measured at the F2 and G1 beam lines at the Cornell University High-Energy Synchrotron Source (CHESS, Ithaca, NY, USA), the BIOCAT beam line at the Advanced Photon Source (APS, Argonne, IL, USA) and beam line X9 at the National Synchrotron Light Source (NSLS, Upton, NY, USA). Specific details of the experimental set-up and procedures specific to each location are provided in the Supplementary Methods. In all cases, the forward scattering from the samples studied was recorded on a CCD or Pilatus (Dectris Ltd., Baden, Switzerland) detector and circularly averaged to yield one-dimensional intensity profiles as a function of q ($q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle in units of Å^{-1}). For final data collection, samples were exchanged into 20 mM Tris pH 7.4, 200 mM KCl and 5 mM $MgCl_2$ at 4°C using 10 kD MWCO Amicon Ultra-15 concentrators (EMD Millipore) and centrifuged through 0.45 μm Whatman spin filters (Fisher Scientific, Pittsburgh, PA, USA) before 0.5–30 s exposures were recorded at 4°C. Scattering from a matching buffer solution was subtracted from the data and corrected for the incident intensity of X-rays. Beam attenuation was employed and replicate exposures were examined carefully for evidence of radiation damage by Guinier and Kratky plot analysis. Silver behenate powder was used to calibrate the sample-to-detector distance and beam center. Additional details on SAXS data analysis and shape reconstruction are available in Supplementary Methods.

RESULTS

Purification of native LtrB group II intron RNP

We designed an affinity purification scheme to specifically isolate the LtrB group II intron RNP in its active state (+A) from its native *L. lactis* host. The overall strategy is modified from that employed previously for purification of the precursor particle [ΔA,(25)], by employing MS2 tag–protein interactions to purify spliced intron away from the precursor. In this scheme, the intron and the IEP (LtrA) components are expressed under the control of independent nisin A promoters (Figure 1C), allowing for formation of the RNP in *L. lactis*. To facilitate purification, a C-terminal chitin binding domain (CBD) was fused to LtrA via an intein. A truncated version of the intron RNA missing 1590 of the 1797-nt LtrA open reading frame (ΔORF) was expressed upstream of the LtrA-intein-CBD cassette. Additionally, the fusion transcript was tagged with a 12xMS2 tag (31) on its 3′ exon (Figure 1C). Upon activation with thiol reagents, the intein cleaves at its N-terminus, releasing LtrA and its tightly associated *Ll.LtrB* group II intron from the chitin resin. *In vivo* expression of this active construct

yields a mixture of RNPs: both unprocessed precursor and spliced lariats (Figure 1C). The addition of the 12xMS2 tag (~1000 nt) to the exon in the presence of MS2-maltose-binding protein (MBP) provides a strategic size differential between the precursor and the excised lariat for separation using a sucrose gradient. SDS–PAGE analysis of the protein content of the RNP captured on the chitin resin (Figure 2A, lanes 2–6) shows LtrA, the fusion MBP–MS2 (Figure 2A, lane 1) and minor contaminants. The final RNP product (Figure 2A, lane 8) showed a single band for the 70-kD LtrA protein after fractionation on the sucrose gradient.

The intron RNA lariats (902 nt) co-purify with 16S ribosomal RNA from the 30S ribosome (Figure 2B) as was the case in the purification of the ΔA precursor particle (25). Fractionation by sucrose gradient allows for isolation of +A particle both alone (Figure 2B, fractions 8–10) and in complex with this 30S ribosomal species (Figure 2B, fractions 12–30). The purity and identity of the protein and nucleic acid components of the +A preparations were confirmed by western blot analysis for the LtrA component and northern blot analysis of the intron component (data not shown). Tandem mass spectrometry analysis confirms the presence of intact LtrA protein in RNP preparations (Supplementary Figure S1). The high-affinity association between the intron RNPs and ribosomes protects the intron from ribonuclease degradation (32).

To confirm that the purified RNPs did not contain a mixture with precursor particles we conducted a primer extension termination assay on all the fractions collected from the sucrose gradient (Figure 2C). Here, we separated cDNA products of the RNAs by denaturing PAGE and determined that the main peak fractions collected during purification (Figure 2B, fractions 8–10) contained an abundance of cDNA corresponding to excised intron (41 nt) and were virtually free of the cDNAs corresponding to precursor (69 nt) (Figure 2C). These analyses confirmed that the purification scheme employed was successful in isolating homogeneous particles of the active RNP.

Functional assays confirm that the RNP is an active particle

We next confirmed the activity of the product of our expression constructs both *in vivo*, before purification, and *in vitro*, after purification. Using the primer-extension assay in Figure 2C, we assayed RNA splicing from actively growing cells and showed that the splicing efficiency of the construct with an LtrA fusion to the intein and CBD was equivalent to that with unfused LtrA (Figure 3A). The precursor–product ratios were similar in the two cases (*cf.* lanes 1 and 3), whereas the non-splicing ΔA constructs showed no products (lanes 2 and 4). Likewise, in a genetic assay that measures retrohoming into a plasmid recipient, the group II intron was inherited at ~60% of the wild-type from a donor containing the LtrA fusion, despite the potential for steric hindrance by the fusion protein (data not shown).
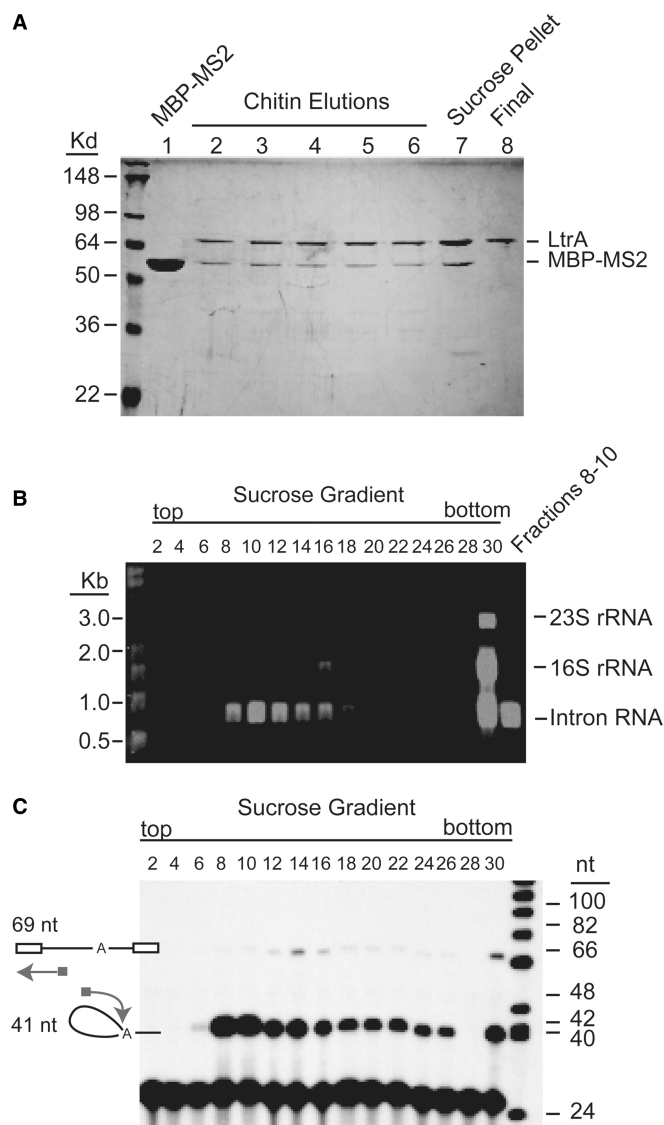
Figure 2. Purification of the active intron RNP. (**A**) SDS–PAGE analysis. Purified intron RNP contains a single protein band at an apparent mass of ~70 kD corresponding to the IEP, LtrA (lane 8) (27). Lane 1: purified MS2–MBP fusion protein, lanes 2–6: chitin elution fractions; lane 7: re-suspension of the pellet from sucrose cushion centrifugation; lane 8: purified RNP. (**B**) RNA analysis of the purified intron RNP. The RNA content of alternating sucrose gradient fractions was analyzed on a 1.2% agarose gel containing 2.2 M formaldehyde. Purified intron RNP in fractions 8–10 was pooled and concentrated. The RNA content of the pooled chitin elution and the re-suspension of the sucrose cushion pellet were also analyzed, with a band of ~902 nt corresponding to the size of the LtrB intron lariat (data not shown). (**C**) Primer extension assay. A primer was designed to target the 5′-end of the LtrB intron. A 69-nt primer extension product was generated from LtrB intron precursor and a 41-nt primer extension product was generated from LtrB intron lariat. Pure lariats (in the absence of the precursor) were isolated in the collected peak fractions 8–10.

We next analyzed the activity of the purified particles, by testing the ability of the intron RNA in the native +A RNP to reverse splice into a 129-bp internally labeled DNA substrate, which contains the exon1–exon2 junction of the *ltrB* gene (Figure 3B and C). The RNP particles collected from each fraction of the sucrose gradient were analyzed for cleavage products by denaturing PAGE analysis (Figure 3C). Endonuclease activity was evidenced by the cleavage of the antisense strand of the dsDNA substrate at position +9, yielding two cleavage products of 64 and 65 nt. Additionally, the slower migrating bands at the top of the gel reflect complete reverse splicing of the lariats on the sense strand of the dsDNA.

Monodispersity, stoichiometry and hydrodynamic properties of +A preparations

We first addressed the concentration-dependent behavior and self-association properties of the +A particle. As determined by SEC-MALS (Figure 4A), the purified +A particles are monodisperse. The +A particle elutes as a single species with an apparent mass of 483 kD against a standard curve derived from globular protein standards. Both this apparent mass and the determined mass (M_w) by MALS of 404 kD $\pm 0.6\%$ across the peak half-height are consistent with a calculated theoretical mass of 432 kD for a complex comprising a 140 kD LtrA dimer bound to a 902 nt RNA (Supplementary Table S3). These results corroborate our previous determinations of mass for these preparations (25) and prior reports of a 2:1 protein–RNA stoichiometry (13). As determined by refractive index, the eluted material was at very low molar concentration, consistent with previous reports of low nanomolar to picomolar binding affinities between the protein and RNA components *in vitro* (13,18,33).

These measurements were complemented by determination of the hydrodynamic properties of the purified +A RNP using sedimentation velocity analytical ultra-centrifugation (Figure 4B). Measurements performed at 7–14 ng/µl concentration indicated single species, consistent with previous analyses of +A and ΔA particles alone in comparable sample conditions (25). The absorbance data collected for the +A RNP was readily fit to the Lamm equation across all the time boundaries collected, yielding a single peak in the $c(S)$ distribution with an $S_{20,w}$ value of 13.3 and a frictional coefficient (f/f_o) of 2.1. This latter value indicates a non-spherical particle, with calculated axial ratios (a/b) of 18–20 for oblate or prolate ellipsoid models, respectively. For reference, a spherical particle ($f/f_o \sim 1.2$) of this composition and mass would appear as a ~30S species (Figure 4B). Using the experimentally determined parameters from sedimentation velocity and a mass-averaged partial specific volume, an estimated mass of 453 kD was derived, again in line with the predicted mass of a 2:1 complex.

Shape and volumetric properties of native group II intron particles determined by SAXS

We extended our analysis to include SAXS, a technique well-suited to characterize large native RNP particles available in modest quantities. The technique has found extensive application to the study of the ribosome (34,35). Although inherently low in spatial resolution, the technique provides the gross structural and compositional properties of particles in solution in a model-independent fashion. Due
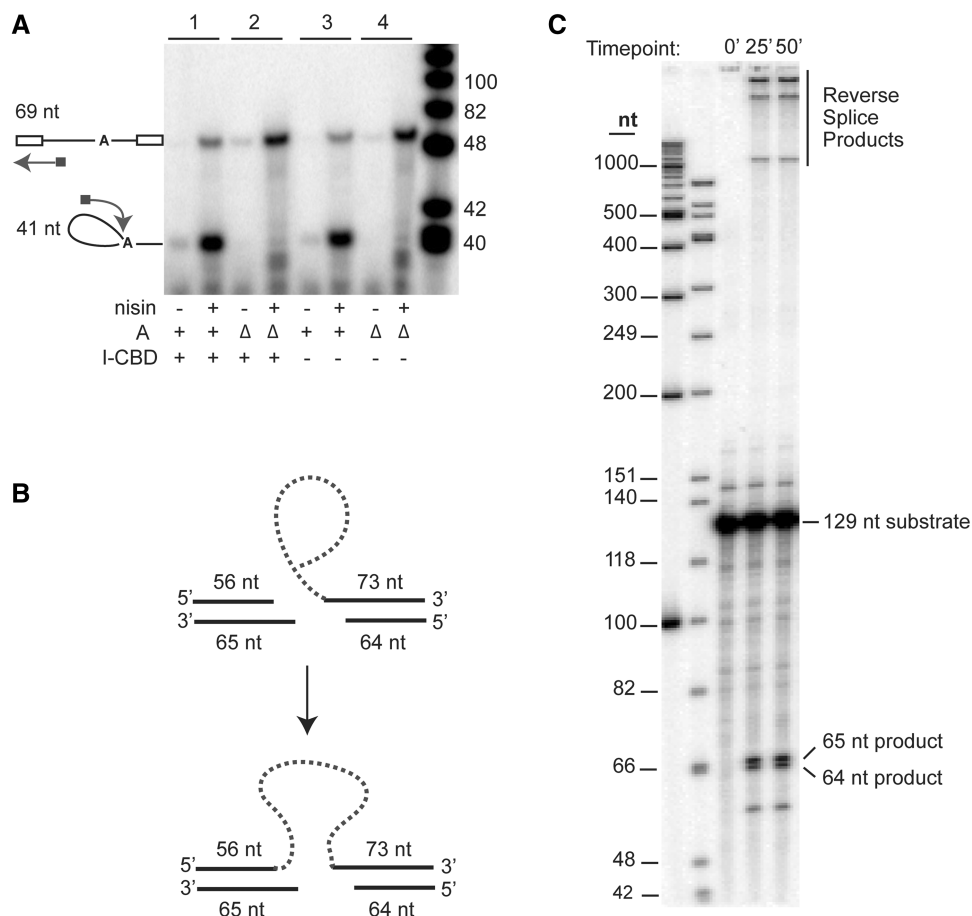
Figure 3. Activity assays. (**A**) Primer extension assay. Assays were performed as in Figure 2C. Samples in each pair of lanes were uninduced (−) or induced (+) with nisin. The splicing-proficient RNPs (+A, samples 1 and 3) were compared with splicing-defective RNPs (ΔA, samples 2 and 4). Fusion constructs containing the intein and CBD (I-CBD) are in samples 1 and 2. (**B**) Schematic description of the *in vitro* reverse splicing assay. (**C**) Reverse splicing assay. Purified intron RNP was incubated with internally ^{32}P-labeled dsDNA substrate containing a 129-bp homing site, at 37°C for varying times, quenched using 1:1 phenol–chloroform and separated on a 5% denaturing PAGE gel. The radiolabeled 129-nt substrate and the 65- and 64-nt product bands are indicated. The highest molecular weight bands (>1000 nt) are reverse splicing products of the lariats on the sense strand of the dsDNA.

to the limited microgram quantities of purified +A particles available, several experimental considerations were made and optimized to facilitate this study (Supplementary Text). Using synchrotron radiation, the +A particle had a radius of gyration (R_g) of ∼73 Å and a maximum dimension (D_{max}) of ∼235 Å (Figure 5A and Table 1), with structural parameters derived from classical Guinier analysis and the indirect Fourier transform yielding comparable results. These SAXS-derived parameters are comparable to those previously determined for 30S ribosomes. For example, the R_g and D_{max} of an oblate ellipsoidal 489 kD 30S ribosomal particle from *Sulfolobus solfataricus* (∼59% nucleic acid) are ∼69 Å and 230 Å, respectively, based on SAXS analysis (36). In contrast, the SAXS-derived R_g and D_{max} of the smaller (∼200 kD) mononucleosome particle (an oblate ellipsoid) that is ∼50% nucleic acid is ∼44 Å and ∼129 Å, respectively (37). Hence, the structural parameters determined by SAXS for the +A RNP are consistent with a nucleoprotein complex of its mass and composition. As determined by Kratky and Porod-Debye analyses (Supplementary Figure S3), some inherent flexibility in the +A particle is indicated.

The asymmetric shape distribution (Pr) profile for the +A RNP is consistent with an ellipsoidal shape (Figure 5B). Consistent with this interpretation, we found that the experimental data at lower scattering angles was well described with an ellipsoid form factor with dimensions of ∼218 × ∼218 × ∼34 Å and a hydrated volume of ∼855 000 Å^3 (χ = 1.93, data not shown), indicating an oblate ellipsoidal shape. This oblate character is further supported by the decay of scattering intensity observed in the experimental profile. For oblate particles, intensity follows approximate q^{-2} dependence, whereas for prolate ellipsoids, the decay is better described by a q^{-1} relationship (38). In the +A RNP data, the decay varies with ∼$1/q^{2.7}$ (Figure 5A). Thus, SAXS and SV analyses both indicate that the +A particle is an ellipsoid of oblate character.

Contrast variation analysis

The SAXS analyses described above assume that the particle is homogeneous. However, the +A RNP is a composite particle consisting of both protein and RNA, where the RNA contributes more scattering power per unit mass.
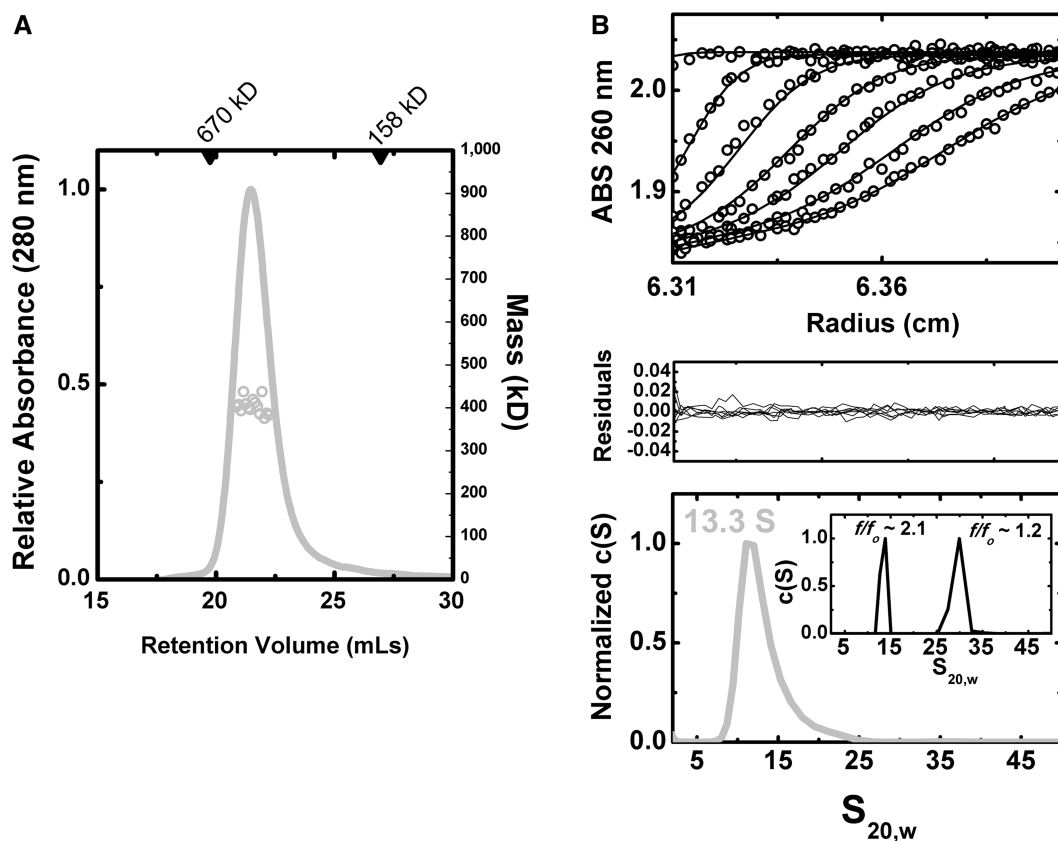
Figure 4. Biophysical analysis of RNP particles. (**A**) Representative SEC-MALS analysis. Shown as a gray line is the absorbance profile of sample, as a function of elution time from a Superdex-200 10/300 column at room temperature (left axis). Corresponding gray circles denote molecular masses determined by in-line light scattering (right axis). Across the half-peak, the determined M_w was 404 kD $\pm 0.6\%$ using a mass-averaged *dn/dc*, consistent with a predicted molecular weight of 432 kD. (**B**) Sedimentation velocity analytical ultracentrifugation. c(S) distributions (lowest panel) were derived from the fitting of the Lamm equation to the experimental data collected, as implemented in the program SEDFIT (29). Shown in the inset of the bottom panel are simulated c(S) distributions for a particle of identical mass and composition to +A but with different frictional coefficients, corresponding to spherical (*f/f_o* ~1.2) and elongated (*f/f_o* ~2.1) forms. In the uppermost panel, fits of the experimental data (black circles) to the Lamm equation are shown as black lines; in middle panel the residuals from this fitting are shown. Measurements were performed at relatively low molar concentrations (~10 ng/μl) and indicated a single species. Consistent with previous measurements, the +A particle showed a single peak with an $S_{20,w}$ value of 13.3 and a frictional coefficient (*f/f_o*) of 2.1.

To evaluate the relative contributions of protein versus RNA to the overall shape of the +A particle, we combined contrast variation with SAXS analysis. Contrast variation using small-angle neutron scattering (SANS) is typically performed by selective deuteration, or more simply, by modification of the scattering density of the solvent using D_2O. The limited quantity of +A RNP particles available precluded the application of SANS, in our study, but contrast variation can be achieved in X-ray scattering by addition of electron-rich additives such as glycerol, salts or sucrose. As the RNP particles in this study have proven to be stable in the high percentage sucrose gradients used in their preparation, we performed contrast variation with X-ray scattering using sucrose.

Theoretical calculations indicated that increasing sucrose concentrations would be expected to amplify the contribution of the RNA component and diminish the contribution of the protein component to overall R_g (which represents the root mean square distance of all of the interatomic vectors relative to the center of mass). The

trend in R_g as a function of contrast provides direct insight into the spatial arrangement of the components within the complex. We hypothesized that if LtrA is positioned on the outside of the complex, away from the center of mass, we would expect R_g to decrease with increasing contrast. Alternatively, if LtrA is located near the center of mass, we would expect R_g to increase, as the protein's contribution to the smallest interatomic vectors within the total particle would be diminished.

The data obtained by this approach are shown in Figure 5C, Table 1 and Supplementary Figure S2. With increasing sucrose concentrations, R_g increased dramatically, with a plateau ~115 Å, indicating that the lower-density protein component of the RNP lies near the center of mass of the particle. The contrast variation results also indicate that the RNA is distributed over a large volume, possibly a consequence of wrapping around a central protein core. Indeed, the apparent R_g provided by the RNA component is surprisingly large, given that Flory's law predicts an R_g of 53 Å for a compactly folded RNA of this composition (39).
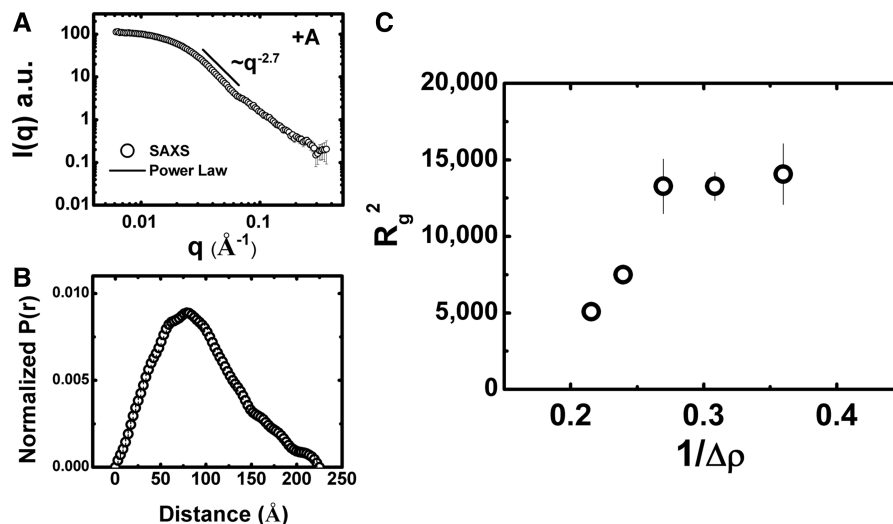
Figure 5. SAXS Analysis of +A RNPs. (**A**) Representative SAXS data from +A (black open circles). The recorded intensity is shown as a function of q ($q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle). Parameters derived from this analysis are summarized in Table 1. Error bars represent plus and minus the combined standard uncertainty of the data collection. Depicted with a black line is slope of the decay in the middle q regime. This region decays with $\sim q^{2.7}$, indicative of an oblate ellipsoid. (**B**) Pr analysis. This analysis is derived from the inverse Fourier transform of the primary scattering data, providing a shape distribution function depicting the distribution of the interatomic vectors in the particle. (**C**) Contrast variation analysis. Stuhrmann plot of the R_g values determined for the +A RNP as a function of the reciprocal of contrast (as conferred by increasing amounts of sucrose). The experimental data points are shown as open circles. The error bars represent the statistical uncertainty associated with measurement.

Table 1. Parameters derived from SAXS

		Guinier			GNOM			
Sample	Conc.[a] (ng/μl)	qR_g	R_g (Å)	I(0) (a.u.)	q[b]	R_g (Å)	I(0)	D_{max} (Å)
APS BioCAT								
+A Group II Intron	100	0.98–1.3	73.5 ± 0.8	324.0 ± 3.0	0.006–0.37	74.9	326.2	250
	50	0.47–1.3	75.3 ± 0.9	139.0 ± 1.0	0.006–0.37	72.7	133.0	225
	40	0.45–1.3	73.0 ± 1.0	120.2 ± 1.0	0.006–0.31	72.4	116.0	225
	Inf. dil.	0.57–1.2	73.3 ± 0.9			73.3		235
NSLS X9								
+A Group II Intron (DII-tRNA)	100	1.0–1.4	87.3 ± 1.2	131.7 ± 2.1	0.010–0.25	95.1	125.0	284
CHESS F2[c]								
+A Group II Intron								
0% Sucrose	40	0.71–1.3	71.2 ± 4.4	0.53 ± 0.03	0.0097–0.15	70.7	0.52	235
16.5% Sucrose	60	0.89–1.7	86.6 ± 3.9	0.70 ± 0.04	0.0097–0.15	83.8	0.64	235
32.5% Sucrose	50	1.09–1.9	105.8 ± 8.3	0.52 ± 0.06	0.0097–0.10	92.9	0.42	235
48.8% Sucrose	100	1.19–2.1	115.2 ± 5.6	0.87 ± 0.07	-n.d.-	-n.d.-	-n.d.-	
65% Sucrose	50	1.22–2.2	118.6 ± 11.8	0.41 ± 0.07	-n.d.-	-n.d.-	-n.d.-	

[a]Where $q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle, in units of Å^{-1}.
[b]As determined by U.V. absorbance.
[c]I(0) in units of (cm^{-1})

The variation in observed R_g with contrast ($\Delta\rho$) shown in Figure 5C can be described by the Stuhrmann equation (40):

$$R_g^2 = R_c^2 + \alpha/\Delta\rho + \beta/\Delta\rho^2$$

Where R_c is the R_g at infinite contrast, the constant α describes the inhomogeneity of the particle and β the displacement of the center of the scattering length density from the center of mass. While the limited range of contrast data obtained precludes rigorous quantitative analysis by this relationship (including a determination of the R_g at infinite contrast), the trends in the data clearly indicate a positive value for the α term. This result is similar to that observed for the nucleosome, where the more electron dense component is located near the periphery of the particle (41). A non-zero value for the β term, indicated by the nonlinear relationship between R_g and contrast, suggests an asymmetric arrangement of protein and nucleic acid components relative to the center of mass of the +A particle.

Introduction of a tRNA landmark into domain II of the +A particle

To further probe the structural arrangement of the +A particle, we engineered tRNAs into selected locations into domains I, II, IV and VI of the intron (Figure 6A). All four of the ΔORF-12MS2 tRNA derivatives had splicing activity (Supplementary Figure S4A), but only the domain II construct [p+A-DII(tRNA)] was well behaved during purification (Figure 6B). The free intron peaked in fractions 12–14 of the sucrose gradient in a similar range to its parental +A counterpart. In both DNA-binding assays (Supplementary Figure S4B) and reverse splicing assays (Supplementary Figure S4C) the +A-DII(tRNA) construct had similar activity to its +A counterpart.

At 20°C, the +A-DII(tRNA) particles were not sufficiently monodisperse for hydrodynamic and scattering analyses (Supplementary Figure S5A). However, at 4°C sedimentation velocity analysis indicated comparable behavior to the native +A particle during the initial stages of centrifugation (Figure 6C). As the centrifugation experiment continued, the +A-DII(tRNA) particles were susceptible to pressure-induced dissociation (Supplementary Figure S5B). This perturbation of the innate stability of the +A particle by the cell pressures associated with high rotor speeds is consistent with the view that domain II lies proximal to the protein–RNA interface (42,43) and possible weakening of the interaction by the tRNA insertion.

Accordingly, our analyses of the RNP particles in this study were performed at 4°C. Based on SAXS analysis, the +A-DII(tRNA) particle features spatial properties significantly larger than those observed for the +A particle alone, as evidenced by R_g and D_{max} (90 Å and 270 Å, Table 1). Differences in the Pr distributions occur at the largest interatomic vectors, consistent with a distal placement of the tRNA component relative to the center of mass of the complex (Figure 6D).

Reconstruction of the low resolution structure of the +A group II intron from solution scattering

To model the relative positions of the protein and RNA components of the +A particle, modeling was performed using a multi-phase bead approach to shape restoration (44). With this method, two phases of different scattering length densities were derived using the 0, 16 and 32% sucrose experimental data and incorporating calculated scattering length densities. Higher angle scattering data, in general contain contributions of internal particle structure. As this approach does not attempt to model this internal structure and assume more uniform densities for the two phases, only the highest signal-to-noise data in the low q regime of data were employed (45).

The results for mutiphase reconstruction are shown in Figure 7 and Supplementary Figure S2. Ten independent calculations yielded similar results, with χ values ranging from 1.18 to 1.39 for each of the three scattering profiles used. Reconstructions of the +A particle featured normalized spatial discrepancy (NSD) values ranging from 0.69 to 0.82, which for an anisotropic particle indicates a reasonably stable solution. These 10 calculations were combined to yield the final averaged and filtered reconstruction (Supplementary Figure S2).

The +A RNP bead model is an oblate ellipsoid with dimensions of 244 × 244 × 79 Å, with calculated sedimentation coefficient and Stokes radius values consistent with those determined experimentally (Table 2). The protein component of this two-phase model is seen to occupy a position at the center of mass, in line with contrast variation results, whereas the RNA phase is seen to assume an extended 'bow-tie'-like arrangement, in which two larger lobes of RNA density are separated by long extension. This overall arrangement is reminiscent of a similarly described 'chain-link' class of shapes observed by cryo-EM with the ΔA particle (25). The protein phase of +A is seen to lie immediately adjacent to a single long extension of the RNA phase which interconnects the remaining two larger lobes. By virtue of its high-affinity primary interaction with LtrA (13,18,33), it can be inferred that this region of the RNA corresponds to domain IV.

While SAXS measurements without contrast variation lack the ability to discern protein from nucleic acid, it is expected that relative comparisons remain informative and are generally representative of gross structural properties. To ascertain the position of the tRNA insert in the +A-DII(tRNA), solution shapes for both +A and the DII(tRNA) variant particles were each reconstructed to the same nominal scattering angle. Thirty independent reconstructions resulted in very similar shapes (Supplementary Figure S6). Although both particles have generally consistent overall properties, the +A-DII(tRNA) features an additional pronounced lobe on the distal end of the particle, extending its effective length (Figure 7B). Manual superposition of these two envelopes reveals a distinct lobe of density corresponding to the tRNA insert and hence, the vicinity of domain II within the intron component. The volume of this difference density is consistent with the known structure of tRNA when manually docked (Figure 7B). The general concordance of the shapes determined and the agreement of the difference volume with available structures of tRNA also suggests that the insertion does not fundamentally alter the native gross properties of the +A RNP structure.

Reconciliation of the low resolution molecular envelope of the +A group II intron with available structural models

Our results provide the opportunity for direct comparison of the +A particle with the spatial features of the ΔA precursor particle previously derived by electron microscopy (25). ΔA is a 250 × 200 × 140 Å particle with an apparent porosity we previously hypothesized to account for its unusually large volume-to-mass ratio (24,25) (Figure 8A). The bulk of the volume of the particle is defined by a main lobe of elongated volume referred to as the 'forehead' feature. Similarities with the structure determined for the +A particle are readily apparent and manual docking of the +A SAXS reconstruction into the main lobe of the ΔA EM reconstruction further underscores these similarities (Figure 8A). The spatial properties of the +A shape correlate well with the main 'forehead'
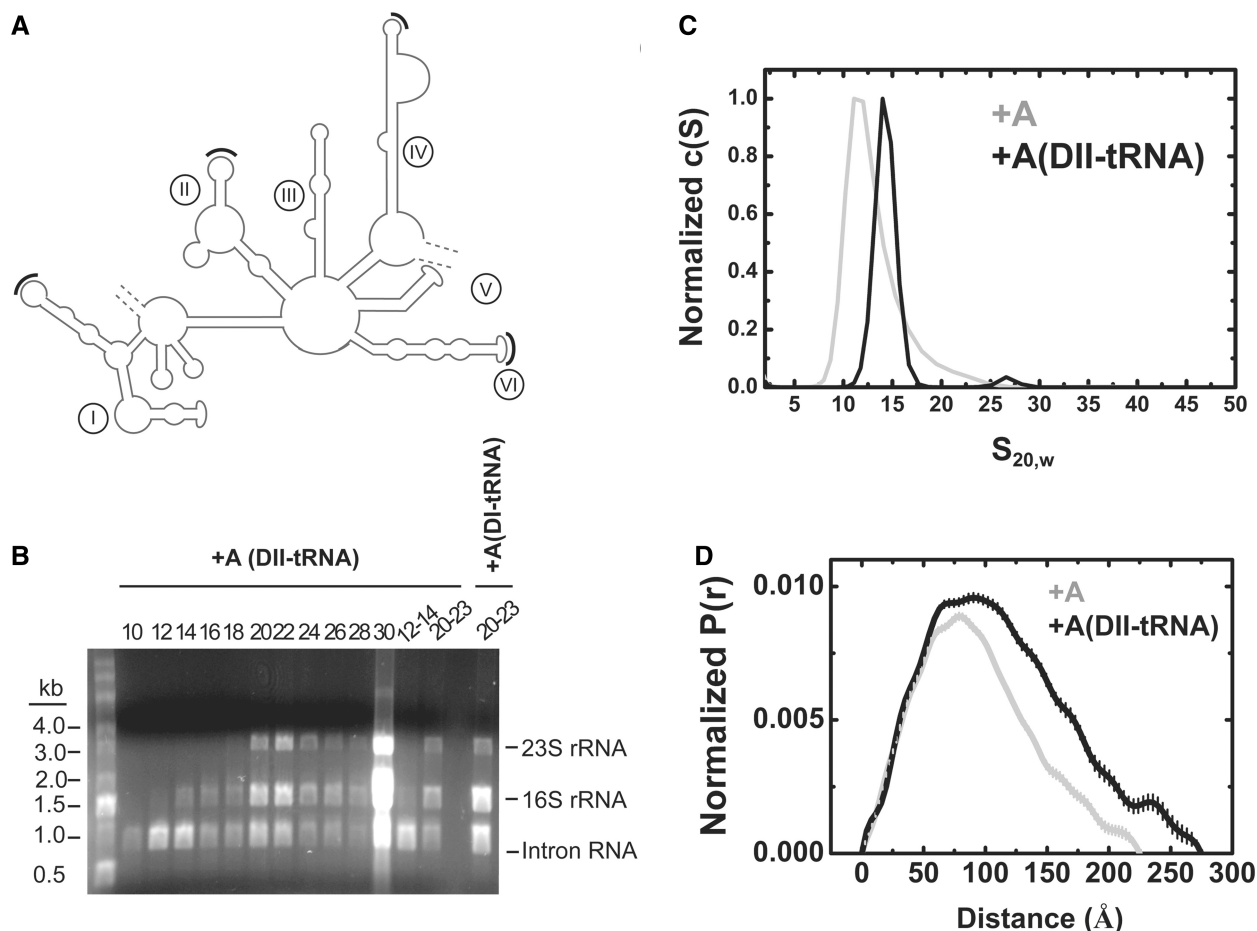
Figure 6. Properties of a tRNA-containing group II intron. (**A**) Secondary structure map of the Ll.LtrB intron with black arcs showing sites of tRNA insertion in domains I, II, IV and VI. In all four cases, a BglII site AGATCT was created in the terminal loop from the following sequences AGACAA (I), AGCACU (II), AAACCU (IV) and AGUAAU (VI) for integration of the tRNA sequences. (**B**) RNA analysis of purified +A-DII (tRNA) intron. The RNA content of alternating sucrose gradient fractions are shown separated on a 1.2% agarose gel containing 2.2 M formaldehyde as in Figure 2B. Pooled fractions 12–14 contain pure intron, whereas pooled fractions 20–23 contain rRNAs as well. In contrast the +A-DI (tRNA) construct on the right yielded no pure intron (blank lane, fractions 12–14) but the pooled fractions 20–23 show intron with rRNA. (**C**) Sedimentation velocity analysis. c(S) distributions are shown that compare the properties of the +A RNP (red) to that of the DII(tRNA) insertion construct (blue). (**D**) Pr analysis. Pr analysis derived from SAXS analysis, comparing the spatial properties of the +A (red) and +A-DII(tRNA) (blue) constructs.

lobe of the ΔA particle, possibly defining a core body preserved in the transition from precursor to spliced state. Taken together, the differences observed between the ΔA and +A shapes indicates an asymmetric and dramatic conformational change that compacts the intron particle from an extended state to a more compact form. These observations are further corroborated by comparison of the calculated Pr function for the cryo-EM ΔA shape with that determined by SAXS for the +A particle (Figure 8B and Table 2).

DISCUSSION

In this study, we have defined biochemical, biophysical and structural properties of the native (+A) group II intron from *L. lactis*, poised for attack of a DNA target. We reconciled these properties with those already determined for the intron in its precursor state (24,25). SAXS is a long-established technique that has been valuable in understanding the properties of native RNPs, including prokaryotic ribosomes (34,35) and

other nucleoprotein assemblies like the nucleosome (37,47,48). The combination of scattering methods with the biochemical and biophysical experiments presented in this report has led to an understanding of the properties of a group II intron RNP as it exists in its native state in solution The experiments have also allowed us to propose a coarse-grain quaternary model for the RNP complex to guide future studies (Figure 7C).

Although purified group II introns RNPs formed in native conditions are highly active, our experimental approach only yields ~1.5 μg RNP/l of culture. These quantities have proved sufficient for study by methods such as electron microscopy (25), mass spectrometry, analytical centrifugation, light scattering and SAXS. Our work with these particles confirms the composition, stoichiometry, and high-affinity interaction between the two components of the +A particle previously determined by *in vitro* analyses (13). This work is a critical first step towards the long term goal of determination of three-dimensional structures of these RNPs and will guide forthcoming cryo-EM reconstructions on the +A particle.
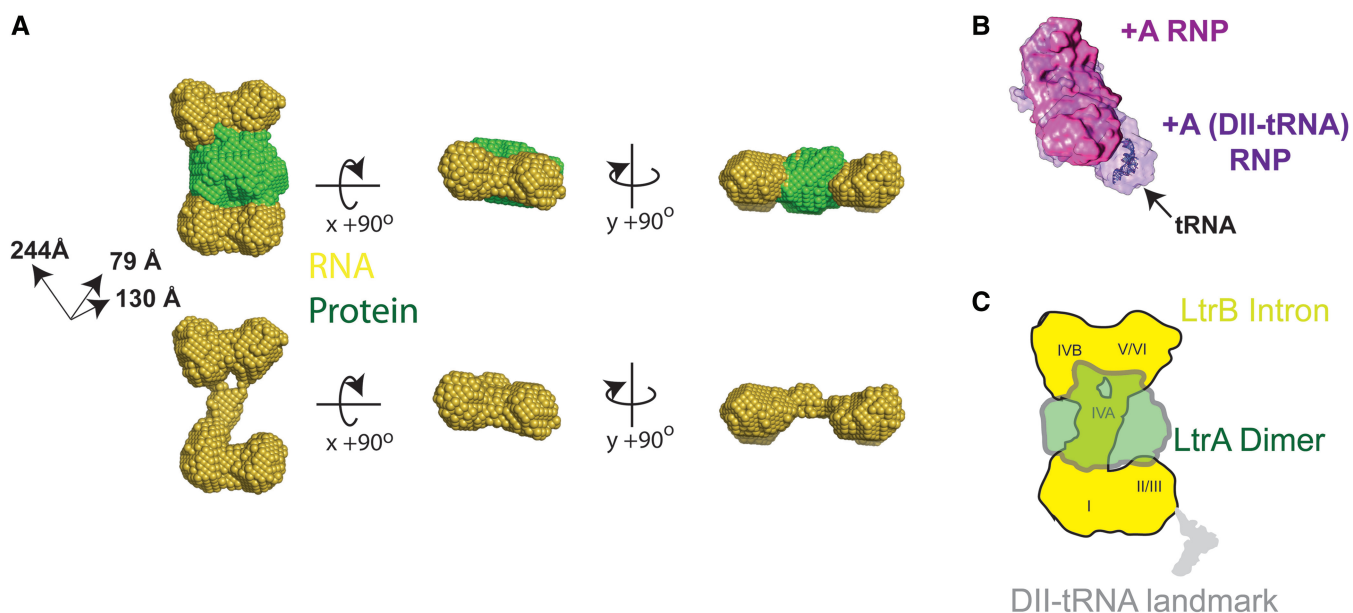
Figure 7. Shape reconstruction of the +A RNP particle. (**A**) The final averaged and filtered two phase shape was calculated by MONSA, shown in orthogonal views. The upper series shows the reconstruction with both RNA (yellow spheres) and protein (green spheres) phases (upper series) and with the RNA phase only (lower series). The final overall particle has dimensions of ∼244 × ∼140 × ∼80 Å. The final NSD figure for the averaged and filtered envelope is 0.77 ± 0.16. (**B**) Manual superposition of SAXS reconstructions for +A (magenta) and +A-DII(tRNA) (purple), revealing the different spatial feature assigned to the tRNA insertion. The available crystallographic structure of phenylalanine-tRNA [2TRA (46), blue] is manually docked into this lobe of difference density. The +A-DII(tRNA) variant is a prolate ellipsoid with comparable dimensions of 177 × 181 × 260 Å. Galleries of individual reconstructions for both particles are provided in Supplementary Figure S5. (**C**) Speculative model for the quaternary arrangement of the +A RNP. Roman numerals denote the speculative placement of the structural domains of the RNA component (yellow) based on the tRNA insertion data, connectivity and percentage volumes estimated for each domain based on sequence and the requirement for LtrA dimer binding to domain IV.

In contrast to previous studies of the particle in its precursor state (ΔA), a relatively more compact group II intron RNP is observed here, with an atomic volume that reconciles well with its composition and stoichiometry. The contrast that is drawn with the experimental structures of its precursor state (ΔA) provides important insight into the dramatic nature of the structural changes that occur upon formation of the native active state. Lacking in the +A shape reconstruction are the porous features that mark the experimental shape of the ΔA particle as well as similar cavities observed in cryo-EM studies of the spliceosome (49,50). This is to be expected, as the shape reconstruction algorithms utilized in this study generally provide descriptions of excluded volume in solution and generally do not attempt to model internal structure. However, the detection of flexibility and large volume-to-mass ratios in our analyses indicate that the +A RNP may also be somewhat distended in solution. We have previously reported on the unusual volume-to-mass determined by biophysical analysis of the ΔA particle, and suggested that the breadth and time-dependent properties of gel filtration and sedimentation velocity profiles might indicate a flexible species where multiple conformations are sampled in solution. Taken together with the results presented here, our studies indicate that the transition from the precursor to native state is accompanied by a significant size reduction and gain-of-order and compactness in the assembly. The driving force behind this structural transition, in the absence of ATP hydrolysis or other co-factors, is not yet understood.

Volumetric differences between the +A and ΔA particles may be due in part to the modest difference in composition (902 nt and 984 nt, respectively), although this discrepancy cannot entirely account for the large changes observed. It might be inferred from manual superposition of the +A and ΔA projections that these conformational changes correspond to RNA domain motions within the intron, as the lobes of volume which differ have elongated properties consistent with extended RNA domains. Indeed, previous proposals have suggested that domains II and VI mediate a conformational change that underlies self-splicing in both group IIA and IIB introns via the conserved η–η′ interaction (42,43,51). As domain II is known to enhance catalysis and domain VI contributes the catalytic adenosine residue, a structural model that couples this conformational change with the formation of a competent active site is plausible. In future studies, the methods employed here might provide the means to link the contributions of putative conserved tertiary interactions to overall quaternary structure of the *L. lactis* intron in different catalytic states, both with natively formed particles and materials prepared *in vitro*.

The shape of the +A particle determined by this study bears mechanistic and structural significance. First, our data indicate that the RNA component of the +A particle is distended by virtue of its flat ellipsoidal arrangement, but still more compact than the ΔA particle. It has been proposed that the binding of the LtrA protein to the LtrB RNA imparts a spatial organization that is required for RNA splicing (33). In the context of the model
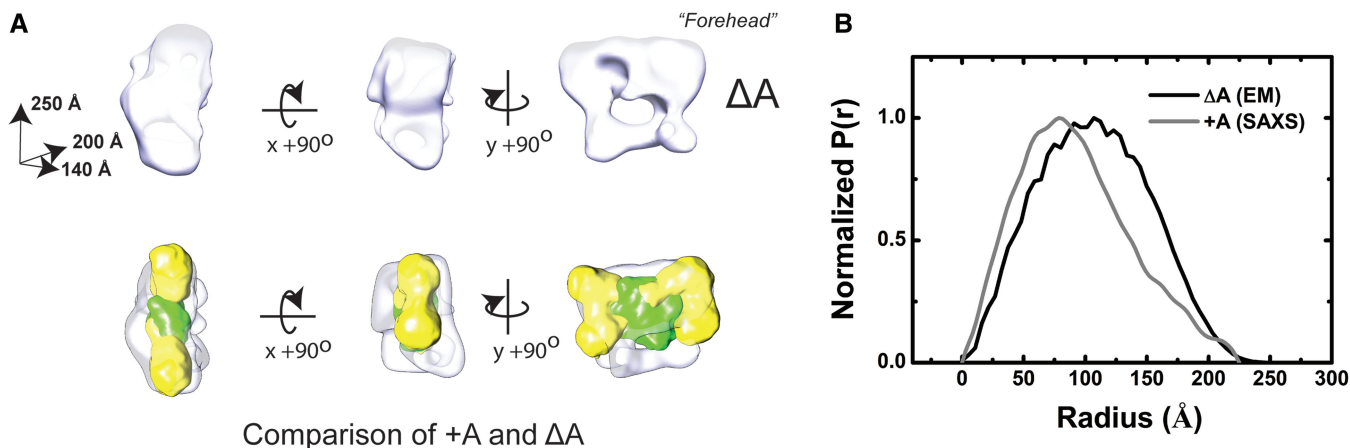
Figure 8. Comparison of the +A and ΔA experimental models. (**A**) The ΔA cryo-EM reconstruction. The model [white, (25)] is shown in orthogonal views. (**B**) Manual docking of the SAXS-derived +A particle (yellow phase for RNA and green phase for protein) into the ΔA shape, revealing spatial differences in RNP structure that coincide with the transition from the precursor to final states. (**C**) Shape distribution function analysis, comparing the interatomic vector distribution of the +A particle (SAXS, magenta) and ΔA shapes (cryo-EM, gray).

Table 2. Comparison of properties of +A and ΔA experimental shapes with their experimental and calculated properties

Model	Experimental model					Sedimentation velocity[b]			SEC[c]	Calculated			
	R_s(Å)	R_g(Å)	D_{max}(Å)	Volume(Å^3)	$S_{T,b}(f/f_o)$	$S_{20,w}$ (f/f_o)	MW (kD)		R_s (Å)	Volume(Å^3)	MW (kD)	v_{bar} (cm^3/g) [g]	
+A (SAXS)[a]	89.5	77.5	275	940 300	15.6	13.3 (2.1)	453		77	450 339	432	0.63	
ΔA (EM) (24)	101.0	83.9	267	2 974 000	14.4	19.7 (1.6)	473		116	475 124	458	0.62	

[a]This work. The calculation of hydrodynamic properties from shape reconstructions are described in Supplementary Methods.
[b]Measurement performed at 4°C.
[c]Measurement performed at room temperature.

provided in this study, it can be envisioned that the high-affinity binding of the LtrA protein to the extended and exposed RNA region assigned as domain IV imparts significant structural ordering of the RNA component, priming the assembly for catalysis by bringing the distal globular RNA components into the proper spatial organization. The LtrA protein has also been shown to have secondary interactions with domains I, V and VI (18). Although SAXS lacks the spatial resolution to discern finer structural details such as secondary structure, the observed positioning of the protein phase in this reconstruction could readily accommodate additional contacts to the larger RNA lobes observed and hence would be compatible with these reported secondary interactions.

Our results also provide an opportunity to reconcile other proposed structural models with experimental solution properties. For example, previous studies of the intron-encoded LtrA protein by threading existing structures of HIV reverse transcriptase with LtrA sequence yielded a three-dimensional model that suggests RNA-binding surface (16). The spatial properties of the LtrA homology model dimer correlates well with the experimental SAXS envelope (Supplementary Figure S7). Manual docking of this structural model into the protein phase of the SAXS-derived envelope indicates that these models are generally compatible with the

proposed interface, with a predicted binding surface near the center of mass.

Crystallographic structures of the related group IIC intron RNAs (22,23) are expected to be representative of the overall conserved secondary structure and gross spatial properties of the first five of the six structural domains of the intron RNA studied here. However, the composition of LtrB intron from *L. lactis* differs in mass by almost 3-fold, and the different subgroups of the group II introns vary significantly with respect to several predicted structural features. Hence, it is not expected that the available bacterial crystal structures can be directly reconciled with our experimental data, nor entirely account for its spatial properties. However, we note that the spatial arrangement of these structures do reveal a relatively planar organization of the structural elements of the intron.

An *ab initio* model derived from experimentally determined crosslinks accounting for 735 of the 902 nucleotides in the ΔORF RNP construct used in this study has also been reported (20). Although the missing nucleotides comprise a significant fraction of the overall inventory, the spatial extent and depth provided by this model can still be considered in light of our experimental results. Visual comparisons reveal that the solution properties of the intron RNA phase are substantially

larger than those implied by the *ab initio* model (Supplementary Figure S7). Possible explanations for the apparent disparity between these structural models is the assumption that long range tertiary interactions previously identified in other self-splicing group II introns will be preserved (52,53) or that if they are maintained, that the interactions are transient.

The distinctive shape of the +A particle also bears significance with regards to macromolecular interactions. A consequence of anisotropic shapes is increased surface-to-volume (S/V) ratios. The axial ratios determined in this study for the +A particle indicate a particle volume much larger than would be predicted for a compact sphere of identical composition and as a consequence, a larger accessible surface. One possibility is that this elongated particle provides a lengthy interface for interaction of the intron with the DNA target over 14 base pairs (54). Another possibility relates to a large number of interacting host factors, including tRNA synthetases (11), DEAD box proteins (55,56), RNAses (57) and even the 30S ribosome (32). The structural significance of the oblate properties of the +A particle may lie in providing accessible interaction surfaces that are not mutually exclusive of other interactions and activities. The experimental approach outlined in this study provides a tractable strategy to probe the structural properties of these even-larger macromolecular assemblies alongside their biochemical properties, including hydrodynamic parameters, composition and stoichiometry. More broadly, such an integrative approach is applicable to the study of other model systems from which native assemblies can be obtained. Recent advances in synchrotron radiation applied to SAXS, detectors and microfluidics further this promise.

SUPPLEMENTARY DATA

Supplemental Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr Robert Singer for the kind gift of the pSL12xMS2 plasmid. We thank Richard Gillilan (CHESS) for stimulating discussion and assistance with data collection. We also thank Liang Guo (APS) and Lin Yang and Marc Allaire (NSLS) for assistance with data collection.

FUNDING

REFERENCES

1. Belfort,M., Derbyshire,V., Cousineau,B. and Lambowitz,A. (2002) In: Craig,N., Craigie,R., Gellert,M. and Lambowitz,A. (eds), *Mobile DNA II.* ASM Press, Washington D.C., pp. 761–783.
2. Dai,L., Toor,N., Olson,R., Keeping,A. and Zimmerly,S. (2003) Database for mobile group II introns. *Nucleic Acids Res.*, **31**, 424–426.
3. Lehmann,K. and Schmidt,U. (2003) Group II introns: structure and catalytic versatility of large natural ribozymes. *Crit. Rev. Biochem. Mol.*, **38**, 249–303.
4. Chalamcharla,V.R., Curcio,M.J. and Belfort,M. (2010) Nuclear expression of a group II intron is consistent with spliceosomal intron ancestry. *Genes Dev.*, **24**, 827–836.
5. Cech,T.R. (1986) The generality of self-splicing RNA: relationship to nuclear mRNA splicing. *Cell*, **44**, 207–210.
6. Cavalier-Smith,T. (1991) Intron phylogeny: a new hypothesis. *Trends Genet.*, **7**, 145–148.
7. Sharp,P.A. (1991) Five easy pieces. *Science*, **254**, 663.
8. Shukla,G.C. and Padgett,R.A. (2002) A catalytically active group II intron domain 5 can function in the U12-dependent spliceosome. *Mol. Cell*, **9**, 1145–1150.
9. Fica,S.M., Tuttle,N., Novak,T., Li,N.S., Lu,J., Koodathingal,P., Dai,Q., Staley,J.P. and Piccirilli,J.A. (2013) RNA catalyses nuclear pre-mRNA splicing. *Nature*, **503**, 229–234.
10. Beauregard,A., Curcio,M.J. and Belfort,M. (2008) The take and give between retrotransposable elements and their hosts. *Annu. Rev. Genet.*, **42**, 587–617.
11. Lambowitz,A.M. and Perlman,P.S. (1990) Involvement of aminoacyl-tRNA synthetases and other proteins in group I and group II intron splicing. *Trends Biochem. Sci.*, **15**, 440–444.
12. Lambowitz,A.M. and Belfort,M. (1993) Introns as mobile genetic elements. *Annu.Rev.Biochem.*, **62**, 587–622.
13. Rambo,R.P. and Doudna,J.A. (2004) Assembly of an active group II intron-maturase complex by protein dimerization. *Biochemistry*, **43**, 6486–6497.
14. Moran,J.V., Zimmerly,S., Eskes,R., Kennell,J.C., Lambowitz,A.M., Butow,R.A. and Perlman,P.S. (1995) Mobile group II introns of yeast mitochondrial DNA are novel site-specific retroelements. *Mol. Cell Biol.*, **15**, 2828–2838.
15. Zimmerly,S., Guo,H., Perlman,P.S. and Lambowitz,A.M. (1995) Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell*, **82**, 545–554.
16. Blocker,F.J., Mohr,G., Conlan,L.H., Qi,L., Belfort,M. and Lambowitz,A.M. (2005) Domain structure and three-dimensional model of a group II intron-encoded reverse transcriptase. *RNA*, **11**, 14–28.
17. Saldanha,R., Chen,B., Wank,H., Matsuura,M., Edwards,J. and Lambowitz,A.M. (1999) RNA and protein catalysis in group II intron splicing and mobility reactions using purified components. *Biochemistry*, **38**, 9069–9083.
18. Wank,H., SanFilippo,J., Singh,R.N., Matsuura,M. and Lambowitz,A.M. (1999) A reverse-transcriptase/maturase promotes splicing by binding at its own coding segment in a group II intron RNA. *Mol. Cell*, **4**, 239–250.
19. Huang,H.R., Chao,M.Y., Armstrong,B., Wang,Y., Lambowitz,A.M. and Perlman,P.S. (2003) The DIVa maturase binding site in the yeast group II intron aI2 is essential for intron homing but not for in vivo splicing. *Mol. Cell Biol.*, **23**, 8809–8819.
20. Dai,L., Chai,D., Gu,S.Q., Gabel,J., Noskov,S.Y., Blocker,F.J., Lambowitz,A.M. and Zimmerly,S. (2008) A three-dimensional model of a group II intron RNA and its interaction with the intron-encoded reverse transcriptase. *Mol. Cell*, **30**, 472–485.

21. Zhou,Z. and Reed,R. (2003) Purification of functional RNA-protein complexes using MS2-MBP. *Curr. Protoc. Mol. Biol.*, **Chapter 27**, Unit 27.3.

22. Chan,R.T., Robart,A.R., Rajashankar,K.R., Pyle,A.M. and Toor,N. (2012) Crystal structure of a group II intron in the pre-catalytic state. *Nat. Struct. Mol. Biol.*, **19**, 555–557.

23. Toor,N., Keating,K.S., Taylor,S.D. and Pyle,A.M. (2008) Crystal structure of a self-spliced group II intron. *Science*, **320**, 77–82.

24. Slagter-Jäger,J.G., Allen,G.S., Smith,D., Hahn,I.A., Frank,J. and Belfort,M. (2006) Visualization of a group II intron in the 23S rRNA of a stable ribosome. *Proc. Natl Acad. Sci. USA*, **103**, 9838–9843.

25. Huang,T., Shaikh,T.R., Gupta,K., Contreras-Martin,L.M., Grassucci,R.A., Van Duyne,G.D., Frank,J. and Belfort,M. (2011) The group II intron ribonucleoprotein precursor is a large, loosely packed structure. *Nucleic Acids Res.*, **39**, 2845–2854.

26. Siatecka,M., Reyes,J.L. and Konarska,M.M. (1999) Functional interactions of Prp8 with both splice sites at the spliceosomal catalytic center. *Genes Dev.*, **13**, 1983–1993.

27. Matsuura,M., Saldanha,R., Ma,H., Wank,H., Yang,J., Mohr,G., Cavanagh,S., Dunny,G.M., Belfort,M. and Lambowitz,A.M. (1997) A bacterial group II intron encoding reverse transcriptase, maturase, and DNA endonuclease activities: biochemical demonstration of maturase activity and insertion of new genetic information within the intron. *Genes Dev.*, **11**, 2910–2924.

28. Gupta,K., Diamond,T., Hwang,Y., Bushman,F. and Van Duyne,G.D. (2010) Structural properties of HIV integrase. Lens epithelium-derived growth factor oligomers. *J. Biol. Chem.*, **285**, 20303–20315.

29. Schuck,P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.*, **78**, 1606–1619.

30. Laue,T.M., Shah,B.D., Ridgeway,T.M. and Pelletier,S.L. (1992) *Analytical Ultracentrifugation in Biochemistry and Polymer Science*. Royal Society of Chemistry, Cambridge, U.K.

31. Bertrand,E., Chartrand,P., Schaefer,M., Shenoy,S.M., Singer,R.H. and Long,R.M. (1998) Localization of ASH1 mRNA particles in living yeast. *Mol. Cell*, **2**, 437–445.

32. Contreras-Martin,L.M., Huang,T., Piazza,C.L., Smith,D., Qu,G., Gelderman,G., Potratz,J., Russell,R. and Belfort,M. (2013) Group II intron-ribosome association protects intron RNA from degradation. *RNA*, **19**, 1497–1509.

33. Matsuura,M., Noah,J.W. and Lambowitz,A.M. (2001) Mechanism of maturase-promoted group II intron splicing. *EMBO J.*, **20**, 7259–7270.

34. Hill,W.E., Thompson,J.D. and Anderegg,J.W. (1969) X-ray scattering study of ribosomes from Escherichia coli. *J. Mol. Biol.*, **44**, 89–102.

35. Kearney,K.R. and Moore,P.B. (1983) X-ray solution-scattering studies of active and inactive Escherichia coli ribosomal subunits. *J. Mol. Biol.*, **170**, 381–402.

36. Rambo,R.P. and Tainer,J.A. (2010) Improving small-angle X-ray scattering data for structural analyses of the RNA world. *RNA*, **16**, 638–646.

37. Damaschun,H., Damaschun,G., Pospelov,V.A. and Vorob'ev,V.I. (1980) X-ray small-angle scattering study of mononucleosomes and of the close packing of nucleosomes in polynucleosomes. *Mol. Biol. Rep.*, **6**, 185–191.

38. Feigin,L.A. and Svergun,D. (1987) *Structure Analysis by Small-Angle X-ray and Neutron Scattering*. Plenum Press, New York.

39. Hyeon,C., Dima,R.I. and Thirumalai,D. (2006) Size, shape, and flexibility of RNA structures. *J. Chem. Phys.*, **125**, 194905.

40. Stuhrman,H.B. (1974) Neutron small-angle scattering of biological macromolecules in solution. *J. Appl. Cryst.*, **7**, 173–178.

41. Hjelm,R.P., Kneale,G.G., Sauau,P., Baldwin,J.P., Bradbury,E.M. and Ibel,K. (1977) Small angle neutron scattering studies of chromatin subunits in solution. *Cell*, **10**, 139–151.

42. Costa,M., Deme,E., Jacquier,A. and Michel,F. (1997) Multiple tertiary interactions involving domain II of group II self-splicing introns. *J. Mol. Biol.*, **267**, 520–536.

43. Pyle,A.M. (2010) The tertiary structure of group II introns: implications for biological function and evolution. *Crit. Rev. Biochem. Mol.*, **45**, 215–232.

44. Svergun,D.I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing (Erratum vol 76, pg 2879, 1999). *Biophys. J.*, **77**, 2896.

45. Petoukhov,M.V., Franke,D., Shkumatov,A.V., Tria,G., Kikhney,A.G., Gajda,M., Gorba,C., Mertens,H.D.T., Konarev,P.V. and Svergun,D.I. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J.Appl.Cryst.*, **45**, 342–350.

46. Westhof,E., Dumas,P. and Moras,D. (1988) Restrained refinement of two crystalline forms of yeast aspartic acid and phenylalanine transfer RNA crystals. *Acta Crystallogr. A*, **44(Pt 2)**, 112–123.

47. Greulich,K.O., Ausio,J. and Eisenberg,H. (1985) Nucleosome core particle structure and structural changes in solution. *J. Mol. Biol.*, **186**, 167–173.

48. Inoko,Y., Yamamoto,M., Fujiwara,S. and Ueki,T. (1992) X-ray scattering study of the shape of the DNA region in nucleosome core particle with synchrotron radiation. *J.Biochem.*, **111**, 310–316.

49. Azubel,M., Wolf,S.G., Sperling,J. and Sperling,R. (2004) Three-dimensional structure of the native spliceosome by cryo-electron microscopy. *Mol. Cell*, **15**, 833–839.

50. Boehringer,D., Makarov,E.M., Sander,B., Makarova,O.V., Kastner,B., Luhrmann,R. and Stark,H. (2004) Three-dimensional structure of a pre-catalytic human spliceosomal complex B. *Nat. Struct. Mol. Biol.*, **11**, 463–468.

51. Chanfreau,G. and Jacquier,A. (1996) An RNA conformational change between the two chemical steps of group II self-splicing. *EMBO J.*, **15**, 3466–3476.

52. Swisher,J., Duarte,C.M., Su,L.J. and Pyle,A.M. (2001) Visualizing the solvent-inaccessible core of a group II intron ribozyme. *EMBO J.*, **20**, 2051–2061.

53. Michel,F. and Ferat,J.-L. (1995) Structure and activities of group II introns. *Ann. Rev. Biochem.*, **64**, 435–461.

54. Mohr,G., Smith,D., Belfort,M. and Lambowitz,A.M. (2000) Rules for DNA target site recognition by a Lactococcal group II intron enable retargeting of the intron to specific DNA sequences. *Genes Dev.*, **14**, 559–573.

55. Huang,H.R., Rowe,C.E., Mohr,S., Jiang,Y., Lambowitz,A.M. and Perlman,P.S. (2005) The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box protein with RNA chaperone function. *Proc. Natl Acad. Sci. USA*, **102**, 163–168.

56. Halls,C., Mohr,S., Del Campo,M., Yang,Q., Jankowsky,E. and Lambowitz,A.M. (2007) Involvement of DEAD-box proteins in group I and group II intron splicing. Biochemical characterization of Mss116p, ATP hydrolysis-dependent and -independent mechanisms, and general RNA chaperone activity. *J. Mol. Biol.*, **365**, 835–855.

57. Coros,C.J., Piazza,C.L., Chalamcharla,V.R. and Belfort,M. (2008) A mutant screen reveals RNase E as a silencer of group II intron retromobility in Escherichia coli. *RNA*, **14**, 2634–2644.