A robust and versatile method for production and purification of large-scale RNA samples for structural biology

HAMPUS KARLSSON, LORENZO BARONTI, and KATJA PETZOLD
Department of Medical Biochemistry and Biophysics (MBB), Karolinska Institutet, SE-104 35 Stockholm, Sweden

ABSTRACT
Recent findings in genome-wide transcriptomics revealed that RNAs are involved in almost every biological process, across all domains of life. The characterization of native RNAs of unknown function and structure is particularly challenging due to their typical low abundance in the cell and the inherent sensitivity toward ubiquitous RNA degrading enzymes. Therefore, robust in vitro synthesis and extensive work-up methods are often needed to obtain samples amenable for biochemical, biophysical, and structural studies. Here, we present a protocol that combines the most recent advances in T7 in vitro transcription methodology with reverse phase ion pairing and ion exchange HPLC purification of RNAs for the production of yield-optimized large-scale samples. The method is easy to follow, robust and suitable for users with little or no experience within the field of biochemistry or chromatography. The complete execution of this method, for example, for production of isotopically labeled NMR samples, can be performed in less than a week.

Keywords: RNA; HPLC; in vitro transcription; sample production; structural biology

INTRODUCTION
The field of RNA biology is increasingly expanding, unveiling the diversity of this class of biomolecules among all organisms (Djebali et al. 2012; Ostankovitch and Pyle 2013). RNA is no longer considered only a mere intermediate to protein production and great interest has been recently focused toward the functional and regulatory roles carried out by noncoding RNAs (Gorski et al. 2017; Bartel 2018), as can also be seen by the ENCODE project (The ENCODE Project Consortium 2012). A detailed understanding of structure, dynamics, and biophysical properties of RNA molecules can thereby shed light on their functions and mechanism of action. A number of methods are available to obtain information in ever-greater structural detail, for example, X-ray crystallography (Shi 2014), small angle X-ray/neutron scattering (SAXS/SANS) (Chen and Pollack 2016), single particle electron cryo-microscopy (cryo-EM) (Vinothkumar and Henderson 2016), chemical mapping (e.g., SHAPE) (Weeks and Mauger 2011) and even dynamics of RNAs by nuclear magnetic resonance spectroscopy (NMR) (Dethoff et al. 2012; Kimsey et al. 2015), and single-molecule FRET (Lemer et al. 2018). While throughput and resolution of those techniques has improved over the years, an important bottleneck has persisted in the production of RNA samples of high enough purity, stability, and quantity. RNA samples can be produced by several methods, with the most commonly used being in vivo expression, in vitro enzymatic semi-synthesis, and solid-phase chemical synthesis and have been recently reviewed (Duss et al. 2012; Baronti et al. 2018; Zhang and Keane 2019). Standard protocols for the production of typical intermediate-sized samples (~20–150 nt) largely rely on in vitro transcription by T7 RNA polymerase and subsequent purification by preparative, denaturing polyacrylamide gel electrophoresis (PAGE) (Milligan et al. 1987; Milligan and Uhlenbeck 1989; Beckert and Masquida 2011). While these techniques are well established, they also have considerable shortcomings. During T7 in vitro transcription, both premature termination (−n) and noncoded nucleotide addition (+1) can occur. These phenomena increase 3′-end inhomogeneity of the transcripts. Purification with traditional preparative PAGE, is limited by the RNA elution step that is diffusion-based, time consuming and contamination prone. Continuous
elution versions of the technique, that address these short-comings have been described previously (Hagen and Young 1974; Cunningham et al. 1996) but are not extensively used. In addition, the handling of large quantities of acrylamide and phenol/chloroform solvents potentially increase the risk of chemical accidents.

A number of solutions have been suggested to resolve 3′-end inhomogeneity during the transcription reaction, including modification of the DNA template (Kao et al. 1999), mutations in the T7 RNA Polymerase (Guillerez et al. 2005), and usage of cis- or trans-acting ribozymes (Ferre-D’Amare and Doudna 1996; Shields et al. 1999).

These methods have been recently reviewed (Baronti et al. 2018). Liquid chromatography techniques have recently evolved as an alternative to the use of PAGE for purification of RNA (McGinnis et al. 2012; Martins et al. 2014). Among others, a number of non-denaturing fast performance liquid chromatography (FPLC) based methods were published (Luksavsky and Puglisi 2004; McKenna et al. 2007; Easton et al. 2010), using size exclusion or ion exchange matrices, which enable direct injection of large volumes of transcription reaction mixture, yielding high amounts of purified RNA in a short time. However, these methods can struggle in yielding a final product with homogenous 3′-ends. Another versatile technique is high performance liquid chromatography (HPLC), especially reverse phase ion-paring (RP-IP) and ion-exchange (IE) systems (McGinnis et al. 2012; Martins et al. 2014). RP-IP-HPLC systems have been shown to be successfully used for RNA purification, offering high-resolution for analytical setups and possibilities to obtain pure milligram (mg) amounts of nucleic acid material in more preparative system configurations of short, synthetic RNA molecules (Murray et al. 1994; McCarthy et al. 2009). Generally, the usability is not restricted to small, chemically synthesized RNAs (<50 nt) and RP-IP have been subsequently used for purification of in vitro transcribed RNA molecules in considerable amounts, up to 0.6 mg of material per injection (Weissman et al. 2013). Similarly, IE-HPLC has been used for purification of mg amounts of synthetic RNAs as well as for in vitro transcribed material, though in combination with trans-acting ribozymes and not requiring single-nucleotide resolution (Andersson et al. 1996; Shields et al. 1999). Methods utilizing ribozymes, RNase H, and RNA ligases in combination with denaturing IE-HPLC for production of segmentally labeled RNA constructs are also an alternative (Duss et al. 2010).

Despite the large body of literature on the topic, a comprehensive method that efficiently combines and optimizes all the relevant innovations does not currently exist. Here we present a fast, robust, and versatile method that builds upon three decades of improvements in the field of RNA sample preparation and compiles it into a complete protocol designed to be performed by inexperienced users (protocol in Supplemental Material). The method described is based on a sequence design step, in vitro transcription from synthetic DNA-strand, and the combined use of RP-IP and IE HPLC for purification. It has been used in our laboratory for the production of both isotopically labeled and unlabeled RNA samples for NMR, SPR, EMSA, SAXS characterization and experiments in human cell lines (Steiner et al. 2016; Schlagnitweit et al. 2018, 2019; Baronti et al. 2020). We here present the method applied on four diverse examples of RNA, a 22mer micro-RNA (Fig. 1D), a 29mer (Fig. 1E), a 46mer, and an 82mer (Fig. 6) RNA molecule, modeling a fragment of ribosomal RNA.

RESULTS

Sequence design

RNA sample production starts with the design of a DNA template to produce the RNA via T7 in vitro transcription (Fig. 1A–C). If the transcript is intended to model a smaller part of a larger RNA molecule (e.g., 29mer Fig. 1E), the function should be assayed to test if the designed RNA molecule has the same structural and/or functional properties as the original RNA molecule (Varani and Tinoco 1991). If that is not possible, secondary structure prediction software such as MC-Fold (Parisien and Major 2008) or the Vienna package (Lorenz et al. 2011) can be used to assess biophysical properties of the smaller RNA molecule to verify a representative fold and stability (Fig. 1B). In our laboratory we preferably use MC-Fold (https://major.irc.ca/MC-Fold/ [user friendly website]) or high-throughput version 2.32 (https://major.irc.ca/Web/mctools) for this purpose, accessible via a user-friendly website. With secondary structure prediction software it is also possible to estimate if possible deletion or addition products would fold differently and therefore result in an in-homogenous sample and negatively impact further experiments. If the transcript is deliberately intended to fold into an RNA “hairpin” an UUCG tetraloop with a closing G:C base pair (cUUCGg) can be introduced in the sequence to guarantee unimolecular folding and keep the hairpin stable (Fig. 1A,E; Nozinovic et al. 2010). Given the native T7 promoter sequence, the yield of the transcription reaction is increased if the DNA template is designed to encode for two guanines at the 5′-end of the RNA transcript (Kuzmine et al. 2003; Beckert and Masquida 2011). Furthermore, to reduce the noncoded addition of nucleotides at the 3′-end of the transcript, the use of C2′-methoxy groups (Fig. 1C), at the two last nucleotides at the 5′-end of the DNA template is recommended (Kao et al. 1999). The DNA template can then be purchased with these modifications (e.g., IDT DNA oligo, synthesized at 1 μmol scale with minimum guaranteed yield of 50 ODs). In addition, purchasing HPLC-purified DNA templates, especially recommended for constructs >25 nt, and T7 promoter DNA
reduces the generation of side-products during RNA transcription, due to a homogenous starting template.

DNA template annealing and in vitro transcription reaction

RNA is in vitro transcribed from the synthetic DNA template using T7 RNA polymerase (RNAP) (Milligan et al. 1987). The efficiency of the T7 RNAP strongly depends on the sequence to be transcribed. Thus, the concentrations of reagents and cofactors, most importantly Mg\(^{2+}\), Tris-buffer and DMSO, need to be optimized. Optimization reactions are typically carried out on a small 50 µL scale (Supplemental Table S1), prior to large-scale production to ensure high yield and minimize side-products (Milligan et al. 1987; Milligan and Uhlenbeck 1989; Chen and Zhang 2005). Transcription efficiency and presence of side-products are evaluated by loading the transcription reactions on denaturing (12%–20%) PAGE gels and stain with ethidium bromide. If no RNA is visible on the gel it might indicate that transcription did not work or that the RNA has been degraded. In the case of degradation it is recommended to check all reactants for contamination, for example by incubating them with a known RNA and if necessary, prepare new stock solutions. Annealing of a new DNA template and replacement of the gel staining solution is also advisable.

Recent developments suggest that the combined use of DMSO (Chen and Zhang 2005) as cosolvent and C2’-O–CH\(_3\) modifications indicated at the 5’-end of the DNA. (B) Example of an Mc-Fold (Parisien and Major 2008) based secondary prediction as sketched in A for an RNA where the secondary structures of a small RNA-hairpin are compared to the secondary structure of a small part of the larger system to verify similar distribution of energetically relevant states. (C) Schematic of a T7-transcription from single-stranded DNA template; highlighted are improvements to reduce undesired side-products and increase yield, e.g., using C2’-O–CH\(_3\) modifications indicated at the 5’-end of the DNA. (D) Sequence and predicted secondary structure of the 22mer; \(T_m\) experimentally determined. (E) Sequence and predicted secondary structure of the 29mer, \(T_m\) predicted using IDT oligoanalyzer tool (https://eu.idtdna.com/calc/analyzer).
the precipitate is allowed to form and then resuspended in a 500 mM EDTA solution to dissolve the particles and release the RNA from the precipitate. A moderate release of the RNA can already be achieved by just resuspending the pyrophosphate pellet with water (Fig. 2) but using an EDTA solution will completely dissolve the magnesium pyrophosphate precipitate. Dissolving the inorganic pyrophosphate pellets at a later stage with EDTA ensures retrieval of the entrapped RNA, while circumventing the need for the time-consuming production or purchase of iPPase. Avoiding usage of iPPase also makes the need to consider the optimization of reaction conditions for the iPPase unnecessary, one less parameter to optimize.

For cases, such as the 22mer, that cannot start the sequence with 5′GG, we found that an excess of NMP corresponding to the starting nucleotide in the RNA sequence (22mer: UMP 12–24 mM) and a reduction of the less abundant NTP (ATP 1 mM) was beneficial. Once an optimum transcription condition has been found (details for each sample in Materials and Methods), using a 20–50 µL reaction scale, it is linearly scaled up to a large-scale transcription, typically 10 mL or more, depending on the required final amount. Interestingly, we found that for the 22mer RNA used here, the large scale reaction resulted in higher yield if the 10 mL were split into 50 aliquots of 200 µL prior to incubation overnight at 37°C. It is rare that the yield of a scale up depends on the reaction volume, but this fact can be checked, if the large-scale yield is suboptimal. After transcription, the crude, concentrated transcription reaction mixture is purified using HPLC. Our optimized HPLC purification method involves two chromatographic stages, first a reverse phase-ion pairing step (RP-IP) followed by ion exchange (IE). Both stages are based on distribution of the material between several, consecutive HPLC injections.

Reverse phase ion pairing HPLC

To prepare the transcribed sample for the RP-IP purification step, possible pyrophosphate precipitate is spun down for 5' at 4900 rpm, the supernatant is collected and the pyrophosphate pellet is dissolved with a 500 mM EDTA solution, to release the remaining RNA, and pooled with the supernatant. The RNA sample needs to be heated and subsequently cooled to denature and precipitate the proteins, which are then filtered off with a 22 µm sterile single-use syringe filter. Removal of precipitated, denatured proteins before the HPLC injections reduces accumulation of these on the column and is crucial to avoid problems in the subsequent purification process.

| Condition no. | Solvent                  | Elution time (mM) |
|---------------|--------------------------|-------------------|
| 1             | H₂O                      | 5                 |
| 2             | Phosphate buffer         | 5                 |
| 3             | iPPase (in Tris buffer pH 8.0) | 40              |
| 4             | 5 mM EDTA                | 5                 |
| 5             | 25 mM EDTA               | 5                 |
| 6             | 50 mM EDTA               | 5                 |

FIGURE 2. 20% denaturing PAGE gel illustrating the release of the 29mer RNA sample from the inorganic pyrophosphate pellet. The individual pyrophosphate pellets have been resuspended with different solutions at both room temperature (RT) and at 37°C. The figure shows that EDTA solutions more concentrated than 25 mM efficiently release the RNA material already at room temperature.

FIGURE 3. Overview over the HPLC technique used for RNA purifications in this work. (A) In the RP-IP, hydrophobic cations form interactions with the negatively charged sugar phosphate backbone of the RNA—hence the name ion pairing—enabling it to interact with the hydrophobic stationary phase of a reverse phase HPLC column, where it will be eluted from with an organic solvent containing elution buffer. (B) In IE chromatography, the stationary phase carries cationic groups that bind the RNA, which RNA is then eluted from the column by increasing the amount of a competing anion. (C) A schematic chromatogram of a typical IE RNA separation/purification of a ~30-mer RNA using a 600 mM ClO₄⁻ buffer system as described here. Nonretained transcription reaction components elute with the column dead volume before the system has been exposed to any eluting buffer. The main RNA of interest is usually separated on a linear gradient in the middle of the chromatogram and elutes in one big peak together with transcription products of similar lengths. A subsequent period of 100% elution buffer helps eluting longer products and possible DNA-template material.
concentration step, when the sample is concentrated to a desired final volume. This volume depends on the loading-loop size of the HPLC and the maximum capacity of the column, here we load RNA material corresponding to ~0.6 mg of the main transcript on a 10 × 250 column using a 500 µL injection loop.

RP-IP-HPLC is based on ion pairing of the analyte with lipophilic cations to enable interaction with the stationary phase (Fig. 3A). The RP-IP step here, is based on a NH₄OAc/acetonitrile buffer system (Murray et al. 1994) using tetrabutylammonium hydrogen sulfate as ion pairing agent and serves as an initial purification step of the crude, concentrated transcription reaction mixture. In this step, a majority of deletion products of lower and higher molecular weight, such as cofactors, single NTPs and enzymes, are separated from the RNA of interest. Instead of using 20 mM NH₄OAc and 2 mM tetrabutylammonium hydrogen sulfate in the eluting buffer (buffer B), as suggested by Arnold and coworkers (Murray et al. 1994), we optimized the concentration to 4 mM NH₄OAc and 0.4 mM tetrabutylammonium hydrogen sulfate, to minimize precipitation/phase separation problems when using higher concentrations of NH₄OAc in acetonitrile. Although 4 mM NH₄OAc limits the buffering capacity somewhat, in our system it does not affect the separation or impact the stability of the purified RNA. An additional positive effect is that less material is wasted. For the system described here, the flow-rate for RP-IP HPLC was optimal at 3.0 mL/min for a column of 10 × 250 mm diameter/length and 19.6 cm³ volume. Best separation performance was achieved using 7% buffer B during loading (2.3 column volumes), to avoid unspecific binding. Thereafter, a gradient of 25% to 35% buffer B over ~4 column volumes was used to elute the target molecule, valid both for the 22mer and 29mer. The gradient depends on size (and with it charge) of sample and should be extended if the sample is larger in size. This was followed by a 100% buffer B wash to elute potential strongly bound fractions and clean the column. To prepare the column for the next injection, reequilibration was done for ~2 column volumes with 7% buffer B. During loading after one column volume, nonretained reaction components elute, already purifying the sample. Elution of longer RNA-molecules can be observed when binding to the stationary phase fails, due to, for example, overloading or an erroneous buffer system, therefore all fractions should always be collected and can be stored at −20°C.

![FIGURE 4. Shown here are RP-IP HPLC chromatograms and 20% denaturing PAGE profiles of two different RNA molecules. Data for the 29mer RNA are shown on the left (A,C,E) and for the 22mer RNA on the right (B,D,F). (A,B) Secondary structure and physical data for the RNA molecules. (C,D) Chromatograms shown represent one out of three RP-IP injections from a 10 mL transcription reaction of the RNA molecules injected on the 250 × 10 mm reverse phase column. (E,F) Fractions indicated by the red lines correspond to the red areas under the curves in the chromatograms of C,D and were chosen for the second IE purification step. Denaturing PAGE gels are stained with ethidium bromide; red arrows indicate position of the main transcript.]
The limiting factor for the RP-IP separations, in terms of maximum amount loaded, is the resolution of the RNA products during elution phase. Loading of larger amounts often leads to loss of separation of the discrete target peaks at ~30% buffer B, therefore containing a mix of longer products (red areas in Fig. 4), while an earlier part of the elution phase, ~25% buffer B, contains other, smaller contaminants, that can be separated. As an example, for 20–30 nt long transcripts, it is optimal to inject <0.6 mg crude, concentrated material of the main transcript for a 10 × 250 mm column. The RNA amount suitable for loading will depend on peak shape, distribution of side-products and is RNA sequence specific, and should therefore be optimized in a range of conditions. Optimization can be achieved by injecting increasing amounts, of less-valuable (e.g., nonisotope labeled) transcription reaction material, such as small-scale optimization reactions. During loading and equilibration phases of the RP-IP, 7% of the acetonitrile containing elution buffer was found optimal, to expose the hydrophobic stationary phase to a minimum of least 5% acetonitrile. The equilibration phase at the end of the gradient prepares the column for subsequent injection immediately after the gradient is finished. Fractions are usually collected in a region covering the 30%–35% Buffer B interval and these are analyzed on a denaturing PAGE-gel. Dilution effects need to be taken into consideration, when analyzing the fractions on the gel. One out of three RP-IP HPLC separations from the sample preparations of the 29mer and 22mer RNAs previously introduced in Fig. 1 are presented in Fig. 4.

**Ion exchange HPLC**

Fractions from the RP-IP stage where the main transcript is highly abundant, are pooled, concentrated and further purified in the second stage, the ion-exchange (IE) chromatography (Fig. 3B,C). RP-IP fractions contain considerable amounts of acetonitrile that can damage cellulose-based membranes in filter concentrators, therefore the flow through needs to be monitored by UV-absorbance or PAGE, for leakage of the RNA. The IE stage is based on a perchlorate buffer system as suggested by Andrus and Bloch (1998) and is performed at 75°C. This temperature is chosen to denature the structure of the RNA as it is close to the melting temperature (Tm) of many GC-rich oligonucleotides, at a high salt concentrations (Kamenetskii 1971). Additionally, a high concentration of chaotropic agents, such as formamide or urea, can be used to aid denaturation and push the equilibrium to a nonhelical, single-stranded RNA structure even further. Here, the ion-exchange buffer system does not include any chaotropic agent, to increase the lifetime of the HPLC-system and the column, as these agents can form crystals (e.g., urea), hazardous for the HPLC and the columns. For equilibration of the 22 × 250 DNAPac PA200 ion exchange column (Thermo Scientific) the following conditions were used: Flow rate of 4.0 mL/min and at a temperature of 75°C was chosen. For the first 0.4 column volumes 0% buffer B was used, followed by a gradient of 20%–30% buffer B between 0.6 column volumes, thereafter a plateau at 100% buffer B for 0.25 column volumes, and finally 0% buffer B for 0.4 column volumes. The gradient used for equilibration is a compromise between time needed and being able to expose the column to a sufficient amount of high salt containing elution buffer to activate it. According to our experience, insufficient exposure of the ion exchange column to elution buffer during the equilibration might lead to binding problems, when loading the RNA on the column. Therefore, a minimum of ~0.5 column volumes of elution buffer is recommended to avoid this. The same problem might occur if too much time passes between equilibration and injection of the RNA material (>1 h). For the RNA separations a flow rate of 4.5–5.5 mL/min and a gradient consisting of a shallow region in the beginning (0% buffer B for ~0.8 column volumes) followed by a gradient where the RNA elutes, starting at ~20% and usually with a steepness of ~0.2–0.8%/min for ~1.5 column volumes are used. The steepness and duration of the gradient for RNA elution/separation depends on the RNA sequence and structure. If no prior knowledge about the elution properties are available, starting with a steeper or longer gradient spanning a wider concentration range of elution buffer is advisable, for example, from 20% to 50% buffer B. Once the upper and lower bound of buffer B for elution has been optimized, the gradient can be made more flat for the subsequent injections. After elution of the main transcript a region with 100% elution buffer follows and at the end a period with 0% elution buffer to prepare the column for the next injection. Examples from representative injections from the IE step and the final concentrated material from all injections are presented in Fig. 5.

**Longer RNA molecules**

In order to display the performance of the method for the purification of longer RNA molecules, we present examples from a 46mer and a 82mer in (Fig. 6). These molecules were, as all other samples, transcribed on a 10 mL reaction scale and 1/3 (corresponding to ~3.3 mL of transcription reaction) of the material was loaded for HPLC purification (Fig. 6).

The yield for the 46mer in (Fig. 6), representing one injection of 1/3 of the material from a 10 mL reaction, (based on UV absorbance @ 260 nm) corresponds to ~0.15 mg of material, equivalent of a 250 µL NMR sample of a concentration of 42.3 µM of >74% purity. In the three rightmost lanes (Fig. 6G), the less pure fractions were estimated to correspond to 0.24 mg and an NMR sample concentration of 66.8 µM of >28% purity. The estimated yield for all fractions pooled, based on a percentage yield in
respect to the most abundant nucleotide in the transcript, is 0.27% (see Materials and Methods). The yield calculation equivalent for the 82mer sample, again only one 1/3 representative injection was performed, corresponds to 0.26 mg of 100% pure material, NMR sample concentration 41.7 µM and for the less pure (>33%), 2.02 mg material and a NMR sample concentration of 321.2 µM. The estimated yield, off all fractions pooled together, is 2.7% (see Materials and Methods).

DISCUSSION

In this work we present an RNA-sample production route using in vitro transcription from 5′-methylated DNA-templates, followed by a HPLC purification procedure consisting of a RP-IP purification step and an IE purification step. We present an overview of a simple method, suitable for beginners to purify RNA transcripts. The amount of RNA material that is possible to obtain with the procedure is in the milligram range, however, the yield depends on the RNA sequence and on what purities are expected. We have indicated in Figures 5 and 6 fractions that are considered highly pure (fractions indicated by green area/lines) and less pure but enriched in target transcript (indicated by red area/line). As expected, we see a clear dependence on the construct sequence, as for example the 22mer, starting with an unfavorable uridine, has a large amount of side products in the crude IVT (Fig. 4F 2nd lane), automatically leading to a lower purity even in the highly pure fraction, while even longer sequences that

FIGURE 5. Shown here are IE HPLC chromatograms and 20% denaturing PAGE profiles of the two different RNA molecules. (A) Loading of 1/5 of the material (final amount 1.8 mg), concentrated from the first RP-IP purification step, of the 29mer RNA construct onto the 250 × 22 mm ion exchange column, one representative out of five injections. (B) Loading of ~40% of the 22mer RNA material (final amount 0.67 mg) concentrated from the first RP-IP purification, one out of three injections. (C,D) Fractions indicated by the green line correspond to the green area under the curve of the chromatogram in A and B containing appropriate fractions (no or little second band detectable) and were pooled from all injections to a final sample. (E) Overloaded PAGE to check for impurities, indicated by the volume loaded of final concentrated sample for the 29mer. The three samples to left (442 µM) represent fractions considered pure, and the three to the right (612 µM) are considered less, but sufficiently pure. All material was pooled together into a final sample with an estimated yield of 5.6% (1.8 mg of material) corresponding to a NMR sample concentration of 0.84 mM in 250 µL. (F) Final purity of the 22mer RNA. Right lane shows the pooled, less pure fractions from the three ion exchange purifications, while the middle lane shows the more pure fractions; left is a 21 nt reference. The final yield for the 22mer RNA sample was 2.0% (0.67 mg of material) and a NMR-sample concentration of 0.4 mM. Differences between the 5.6% of the 29mer to the 2% of the 22mer highlights the increase in yield if the advantageous starting sequence (5′GG) is used.
have a more favorable sequence for IVT transcription, such as the 46mer and even 82mer can be produced with very high purity using this method (Fig. 6). This indicates, as anticipated, that a purification is just an enrichment of one product (the target) over others. In order to estimate the final and correct purity, one has to consider the detection limits of the impurities. We therefore have used a dilution series of the pure product, where one can, at the highest concentration, always see side products and hence estimate the purity (29mer: Fig. 5E, 22mer: Fig. 5F, 46mer: Fig. 6G, 82mer: Fig. 6H, respectively); however, the concentration of the main transcript is often outside of the dynamical range (Supplemental Fig. S5) for exact concentration estimation, therefore values determined this way represent the lower end of the estimated purity and yield. For the concentrated samples of the IE fractions (Fig. 6G, H), quantification using imageJ was performed to estimate the abundance of the main transcript and deletion products (see details in Materials and Methods as well as in Supplemental Figs. S1–S5; Supplemental Table S3). To achieve even higher purity, one can repeat the IE step of the highly pure fraction and to increase yield, the less pure fractions can be pooled and repurified; however, this will cost time and reduce yield material.

Comparison with other methods

For comparison, our purification method is direct injection of large amounts of crude RNA from T7 based, single-stranded in vitro transcription into an RP/IE HPLC system. Literature is somewhat sparse with examples that are immediately translatable to the work presented here (Arghavani and Romano 1995; Andersson et al. 1996; McCarthy et al. 2009). Additionally, HPLC purification of solid phase synthesized RNAs has been performed (Arghavani and Romano 1995; Andersson et al. 1996; McCarthy et al. 2009). Andersson et al. (1996) addresses the impact of secondary structures, which can be estimated by using UV-melting to obtain the melting point ($T_m$) in dependence of cation ions used to stabilize the RNA secondary structure. In our approach we heat our IE column to 75°C with the intention to denature the RNA secondary structure and therefore obtain a homogenous population of single-stranded RNA. 75°C is, however, close to or slightly above the $T_m$ of most RNAs in the size range discussed here, especially in the salty IE buffer system. Therefore, in our systems one can expect multiple conformers of secondary structures, exchanging fast with the melted, linear form, reducing the separation capacity; however, no time is invested to assess ionic requirements and reduced resolution is compensated by repeated smaller injections. We have tested usage of formamide as a chaotropic agent to reduce the problem of RNA secondary structures (Kondo et al. 2014), but in our experience, the gain in separation efficiency is small. RP-IP purifications of solid phase synthesized RNAs (Arghavani and Romano 1995; McCarthy et al. 2009) can be compared, since they address a similar RNA scale, ~1 mg range, and show the impact on the chromatogram of loading larger amounts of material. However, solid-phase synthesized RNA can be expected to contain less deletion products than IVT produced RNA, with its enrichment of ±1nt products, therefore simplifying the purification of solid-phase synthesized RNA.

Alternatively, to improve the purification method, one can address the shortcomings of IVT by starting with less by-products, for example, by using a methylated template, as suggested in Kao et al. (1999) or tandem templates with an optimized starting sequence (Feyrer et al. 2020). Usage of ribozymes or enzymatic methods is also an alternative; these options can improve homogeneity and counteract the inherent shortcomings of the chromatographic methods that exist. Shields et al. (1999) utilizes the trans-cleavage by a hammerhead ribozyme to aid in separation and improve 3’-end inhomogeneity. Yet the hammerhead ribozyme does have some sequence restrictions, such as the 5’ relative to the cleavage site, and a slight 3’-end extension is required, hence this method is not applicable for systems such as miRNAs where the sequence cannot at all be modified. Other trans acting ribozymes exist, for instance antigenomic hepatitis delta ribozyme that has no sequence requirements 5’ relative to the cleavage position, however, it is again necessary to extend a transcript to be cleaved with a recognition sequence for the ribozyme (Wichlacz 2004), similar for the Varkud satellite (VS) ribozyme (Ferre-D’Amare and Doudna 1996). So although it would be possible to use these ribozymes, which can be nonisotopically labeled (important for NMR) in trans with a $^{15}$N/$^{13}$C labeled transcript of interest, optimization of ribozyme cleavage/sequence, an additional separate purification step prior to ribozyme cleavage and the fact that some labeled nucleotides will be cleaved away are factors to be considered when compared to our simpler approach. There are also several examples of when cis-acting ribozymes are used, both cis-acting hammerhead ribozymes (Price et al. 1995) and ribozymes, such as the hepatitis delta ribozyme (HDV), can be used (Scherer 2002) to reduce 3’-end inhomogeneity and even if optimized cleavage reactions (Kanwal et al. 2019) and shorter versions of these ribozymes have been discussed more recently (Akoopie and Müller 2018). However, these ribozymes are still in the ~60–80 nt size range and to use that in cis and cleave from a $^{15}$N/$^{13}$C labeled transcript is not a very attractive option, as the yield based on input nucleotides will be automatically cut by 50% for RNAs sized below 100 nt. Combined usage of ribozymes, RNases and ligases to obtain segmentally labeled, homogenous length transcripts are also alternatives (Duss et al. 2010). Although all of these ribozyme/ enzymatic based approaches could be alternatives, they can be complementary to our method, and the possible
FIGURE 6. Example of the IE purification of the 46 and 82mer. (A) Physical data for the 46mer. (B) Physical data for the 82mer. (C) Chromatogram from injection of the 46mer, material corresponding to 3.3 mL transcription reaction mixture, prepurified by RP-IP HPLC according to the protocol, flow rate 7 mL/min. (D) Chromatogram from injection of the 82mer, material corresponding to 3.3 mL transcription reaction mixture, prepurified by RP-IP HPLC according to the protocol, flow rate 7 mL/min. (E) 20% PAGE gel with fractions from the labeled areas in the chromatogram (C), gel stained with ethidium bromide. Fractions below the red line were collected from the red area in the chromatogram and were considered less pure, those below the green line were collected from the green area and were considered purer. (F) 10% PAGE gel with fractions from the labeled areas in the chromatogram (D), gel stained with ethidium bromide. Fractions below the red line were collected from the red area in the chromatogram and were considered less pure, those below the green line were collected from the green area and were considered purer. Leftmost lane represents the transcription reaction. (G) Concentrate of the IE fractions of the 46mer. The three leftmost lanes represent the concentrate of the fractions considered purer (green in panel E) and loaded as 1×, 10×, 100× dilutions of the concentrate. The three rightmost lanes are the same dilutions of the concentrate from fractions considered less pure (red in panel E). The value in percent above each lane is an ImageJ based estimate of main transcript abundance for that sample (see discussion and Supplemental Fig. S1–S5; Supplemental Table S3 in Supplemental Information). (H) Concentrate of the IE fractions from the 82mer. Leftmost lane represents the crude transcription reaction. The three subsequent lanes represent 1×, 10×, 100× dilutions of the purer material (green, panel F) and the three rightmost lanes of the same dilution of the less pure material (red in panel F). The value in percent above each lane is an ImageJ based estimate of main transcript abundance for that sample (see discussion and Supplemental Fig. S1–S5; Supplemental Table S3 in Supplemental Information).
additional sequence requirements and optimization steps are complicating factors that we try to avoid and compensate for by accepting lesser separation and performing more injections.

An alternative purification method is based on FPLC. This category of protocols constitutes a competitive alternative to our method as existing literature shows efficient purification of high amounts of RNA in short time spans. Size exclusion FPLC methods have been used to produce highly concentrated NMR samples, although with inhomogeneous ends (Lukavsky and Puglisi 2004; Kim et al. 2006; McKenna et al. 2007), similar to HPLC methods. Anion exchange versions of the FPLC method is also able to purify large amounts of RNA (>10 mL) in a couple of hours (Easton et al. 2010; Koubek et al. 2013). However, these methods also struggle with the same homogeneity issue of the purified material, even though the speed and capacities are high. The advantage of HPLC over FPLC is that one can heat HPLC columns to a larger extent and use stronger salt conditions on them, decreasing some of the homogeneity issues further than possible with FPLC. FPLC-based methods are superior, if the purification is based on a tag or a specific sequence the column can be primed with (Pereira et al. 2016); however, this reduces the variability to any type of RNA construct and is usually used if one specific RNA is needed in larger amounts (>grams).

Finally, we focus on a classic purification technique, preparative PAGE separation, followed by elution from the gel (Ogden and Adams 1987; Petrov et al. 2013). The elution from the gel matrix can either be done by passive diffusion of the nucleic acid material out of the gel, for example, the crush and soak method, or by electroelution (Chen and Ruffner 1996; Fadouloglou 2013). Preparative PAGE is well established and widely used due to its robustness and reliability; continuous elution versions of the techniques are also interesting (Hagen and Young 1974; Cunningham et al. 1996). For example, the method used by Cunningham et al. (1996) using the Model 491 Prep cell is able to purify 18.5 mg of RNA from 100 mL of transcription reaction for RNA in the ∼30mer size range though with ±n bands present. A limitation here is the length of the column for proper separation and it is surprising that this method has not found more application. Unfortunately, these methods can be prone to contamination, are somewhat work extensive (handling of large PAGE gels) and numerous instrumentations have been discontinued. Although this technique is a good alternative to HPLC/FPLC based purification, we emphasize the simplicity and straightforwardness as strong points of our technique.

Staining dye and impact on purity estimate

To get an accurate estimate of the purity of a nucleic acid sample, or abundance of a particular length nucleic acid molecule among others, can be complicated by a number of factors. Two important factors are the separation technique with a guaranteed single nucleotide resolution and a safe detection technique that would give a signal that is directly proportional to the amount of an RNA of a particular length. Under certain circumstances gel electrophoresis in combination with radioactive labeling or some LC-MS techniques could provide this, but these techniques are not always available to the everyday researcher and require specific/expensive equipment/rooms. Therefore, gel electrophoresis with a staining dye is the easiest option and often used. One of the historically relied upon stains is Ethidium bromide (EtBr). In recent years, many supposedly safer alternatives have been developed, for example, GelRed, GelGreen, tryphan blue and compound in the SYBR product series, SYBRGreen, SYBRSafe, and SYBRGold (Tuma et al. 1999). We have chosen to use SYBRGold, in our hands the closest in sensitivity of EtBr, in parallel to EtBr, to get a better idea of the purity of our RNA material. SYBRGold is supposed to be more sensitive than EtBr and safer (Tuma et al. 1999; Kirsanov et al. 2010). As sensitivity has been an ongoing issue for denaturing 20% PAGE gels, we present in Supplemental Figures S6, S7 the same samples loaded on PAGE gels stained with EtBr and SYBRGold. From this work we have concluded that SYBRGold probably is more sensitive for lower concentration bands than EtBr for this type of gel and therefore excellent in detecting impurities and is most often able to detect at least as many products as EtBr. However, it is clearly visible in our data (Supplemental Figs. S5–S7) that band intensity between products differ less for SYBRGold than for EtBr, indicating that the dynamic range of SYBRGold is smaller than for EtBr. This can lead to wrong estimates in purity and yield and need to be considered in the application. We further find that staining of SYBRGold is less homogenous and more prone to uneven staining for 10%–20% denaturing PAGE.

Simplification by excluding RP-IP step

As we have stressed the simplicity of this method, we also want to point out that for many samples it is possible to even skip the RP-IP step entirely, saving a whole day of time and reducing sample loss with the same result as when using RP-IP. An example for this on a 39mer is shown in Supplemental Figure S8 (Supplemental Information). Although this might indicate that the RP-IP step is redundant, in many cases we still find it useful, primarily to reduce the load on the IE column and in some cases to aid purification. In some instances, only utilizing the less efficient RP-IP purification can be an option, if one wants to avoid the high salt concentrations associated with the IE.
Conclusion

In this work we describe an RNA sample design, production, and purification method based on the combined use of efficient T7 in vitro transcription with two different HPLC techniques, RP-IP and IE HPLC for purification. The method is robust and versatile and suitable for RNA sample production for structural biology techniques, for example, NMR, X-Ray crystallography, or SAXS experiments, that demand material in the milligram range. The method can be used for production of samples in a wider size range, in the ~20–80 nt range as exemplified here. It is also well suited for production of RNA material for other types of experiments, for example, drug screening, CRISPR RNAs, siRNA for gene modification of expression regulation, or cell transfection experiments where purified RNA material is needed.

MATERIALS AND METHODS

Inorganic pyrophosphate pellet resuspending experiment

A 12 mL transcription reaction of the 29mer was performed. After transcription the reaction mixture was extensively vortexed, to create a homogenous suspension. The suspension was distributed between 12 different falcon tubes, 1 mL per tube. These falcon tubes were centrifuged 3 min at 9400 rpm to form pellets. The supernatants from these tubes were removed. To each pair (six in total) of these tubes 500 µL of the following solutions were added: ddH2O, Phosphate buffer (15 mM sodium phosphate 25 mM NaCl, 0.1 mM EDTA, pH 6.5), iPPase (in Tris buffer pH 8.0, 10 mM MgCl2), 5 mM EDTA, 25 mM EDTA, and 50 mM EDTA. The pellets were resuspended with these solutions and then incubated, in each pair of tubes. One tube was left at room temperature (RT) and one in a water bath at 37°C. All tubes except from the iPPase containing ones were incubated for 5 min, whereas those with iPPase were incubated for 40 min. After incubation the remaining precipitate was spun down again (3 min at 9400 rpm), 0.5 µL of this supernatant was then loaded on a 20% PAGE gel and stained with ethidium bromide.

DNA annealing and in vitro transcription

Synthetic DNA templates (Integrated DNA Technologies) with C2-methoxy groups at the two last nucleotides at the 5’-end and the T7 promoter sequence at the 3’-end were annealed with equimolar amounts of DNA complementary to the T7 promoter by heating to 95°C for 5 min, followed by rapid cooling, using DNA concentration of 25 µM in the presence of 3 mM Mg2+.

The DNA sequences used for the 29mer RNA molecule were 5’-GGACACATGGCCAGGCTTCTTGGGCCATATGCGAGCGTTTATATGCGAGCGTTGCTATATGCGAGCGTTGCTATATGCGAGCGTTGCTATATGCGAGCGTTGCTATATAA-3’ and for the 22mer RNA molecule were 5’-ACAACC AGCTAAGACACTGCCATATAGTGAGTCGTTATTAA-3’. Optimizations of in vitro transcription reaction conditions for the 22mer RNA were done on a 50 µL reaction scale varying the concentrations of Tris-HCl (50–150 mM), MgCl2 (30–60 mM), and DMSO (0%–30% v/v), with NTPs kept at 3 mM concentration. Addition of NMP corresponding to the starting nucleotide in the RNA sequence at 3 mM concentration and influence of E. coli inorganic pyrophosphatase (produced in-house) and RNase inhibitor (RNaseOUT, ThermoFisher), used according to manufacturer’s specifications were tested. Similarly, the 22mer RNA in vitro transcription reaction conditions were optimized in 50 µL reaction scale varying the concentrations of Tris-HCl (20–100 mM), MgCl2 (10–40 mM), and DMSO (5%–20% v/v). Reactions were preheated for 5 min at 37°C prior to addition of T7 polymerase (produced in-house) and RNase-inhibitors and inorganic pyrophosphatase, then incubated at 37°C for 2 h. Transcription reaction reagents used that were not produced in-house or purchased from ThermoFisher were purchased form Sigma-Aldrich or Fischer Scientific. All reagents for transcription were prepared with ultrapure type 1 water (18.2 MΩ·cm at 25°C), sterile filtered and stored at ~20°C. Efficiency of transcription was evaluated by loading 0.5–1 µL of the reaction mixture, mixed with 9 µL loading dye (bromophenol blue and EDTA in formamide), on 22 x 20 cm denaturing PAGE gel (20% polyacrylamide, 8 M Urea) and detected with ethidium bromide on a WVR Genoview mini 5 UV-board. Large scale reaction for the 29mer RNA was done in a 10 mL reaction scale by linear scale up of the optimum reaction condition (100 mM Tris-HCl, 40 mM MgCl2, 30 mM DTT, 2 mM Spermidine, 10% [v/v] DMSO, 3 mM each of NTP and GMP, 0.4 µM DNA, and 200 µL of T7 polymerase) and was incubated overnight at 37°C. Similarly, the large scale reaction for the 22mer RNA was done on a 10 mL reaction scale by linear scale up of the optimum reaction condition (80 mM Tris-HCl, 20 mM MgCl2, 60 mM DTT, 2 mM Spermidine, 10% [v/v] DMSO, 1 mM ATP, 3 mM GTP, 3 mM CTP, 3 mM UTP 24 mM UMP, 0.4 µM DNA, and 200 µL of T7 polymerase) and by distributing the reaction between 50 different Eppendorf tubes (200 µL each). In the example shown here, 13C and 15N fully labeled ATP and UTP (Sigma-Aldrich) were used.

Reverse phase ion pairing HPLC purification

HPLC purifications were done on a Dionex Ultimate 3000 UHPLC system (Thermo Scientific) with a LPG-3400RS pump, a TCC-3000RS column thermostat, a WVD-3100 detector, an AFC-3000 HPLC fraction collector, a two-way switching valve, a reverse-phase column operated at room temperature and an ion-exchange column in the column oven at 75°C. Prior to RP-IP, the 29mer RNA 10 mL large scale transcription reaction was centrifuged at 2000 g, 5 min, supernatant was collected and distributed between 10 Eppendorf tubes, 1 mL each. The remaining pellet was dissolved by vortexing in 5 mL of a 100 mM EDTA solution, pH ~6.5 and distributed between five Eppendorf tubes. The Eppendorf tubes were then heated to a temperature above 70°C, then cooled on ice to room temperature and precipitate spun down (2000 g, 4 min). The mixture was filtered using a 0.2 µm syringe filter (Pall Laboratory), and concentrated with centrifugal filter unit 3 kDa MWCO (Amicon ultra) to a final volume <1 mL and distributed between three different injections in fractions of 310 µL, corresponding to ~0.6 mg/injection on the RP-IP column (Hypersil Gold 250 x 10 reverse-phase semi preparative column, Thermo Scientific). The 22mer RNA 10 mL large-scale transcription reaction was centrifuged at 2000 g, 5 min.
The pellet was dissolved in 10 mL of a 100 mM EDTA solution, pH ~6.5 and added to the reaction supernatant. The resulting 20 mL mixture was concentrated with centrifugal filter unit 3 kDa MWCO (Amicon ultra) to a final volume <1.5 mL and collected in one Eppendorf tube. Again, the concentrated solution was heated to a temperature above 70°C, then cooled on ice to room temperature and precipitate spun down (2000g, 4 min). The resulting supernatant was distributed between three different injections on the RP-IP column (Hypersil Gold 250 × 10 reverse-phase semi preparative column, Thermo Scientific). Unless otherwise stated, separation was performed under the same conditions for both RNA samples and achieved at room temperature with a buffer A: 100 mM NH₄OAc, 2 mM tetrabutyl ammonium hydrogen sulfate, pH 6.5, degassed under vacuum with stirring for ≥30 min. The eluting buffer B contained 4 mM NH₄OAc, 0.4 mM tetrabutyl ammonium hydrogen sulfate (prepared from 100× stock solution), pH 6.5. Fresh acetonitrile was added to a final concentration of 80% (v/v) and degassed in an ultrasound bath for ≥30 min.

Equilibration of the reverse-phase column was achieved at a flow rate of 2.5 mL/min, with the following sequence: gradient of 20%/min for 5 min, 5–15 min (100% buffer B), gradient 20%/min for 5 min 20–45 min (7% buffer B). Purification of the 29mer RNA was done using a flow rate of 3.0 mL/min and 7% buffer B for 0–15 min, gradient of 12.5%/min for 2 min, gradient of 0.33%/min from 17–47 min, reaching 35% buffer B, gradient of 13%/min for 5 min, 52–62 min (100% buffer B), gradient of 20%/min for 5 min, 67–80 min (7% buffer B). An amount of 1 mL fractions were collected between 32.5–52.5 min. For the 22mer RNA 1 mL fractions were collected between ~36–41 min. The RP-IP purification of the longer RNAs (the 82mer and 29mer) were pooled together and fractions of satisfying purity were pooled together and concentrated down to a volume 0.1 mL using acetic acid and degassed under vacuum with a stirring magnet for ≥30 min before usage. Equilibration of the ion-exchange column with this buffer system was done at 75°C, at a flow rate of 4.0 mL/min. With the following sequence: 0% buffer B for 0–10 min, gradient of 20%/min for 1 min, gradient of 0.71%/min from 11–25 min, reaching 30% buffer B, gradient of 35%/min for 2 min, 27–33 min (100% buffer B), gradient of 50%/min for 2 min, 35–45 min (0% buffer B).

IE-HPLC of the 29mer was done at 75°C and a flow rate of 4.5 mL/min using for the 1st injection: 0% buffer B for 0–15 min, gradient of 17%/min for 1 min, gradient of 0.74%/min from 16–43 min, reaching 37% buffer B, gradient 63%/min for 1 min 44–57 min (100% buffer B), gradient 100%/min for 1 min, 58–73 min (0% buffer B). For the second and third injection; 0% buffer B for 0–10 min, gradient of 17%/min for 1 min, gradient of 0.74%/min from 11–45 min, reaching 42% buffer B, gradient of 58%/min for 1 min 46–59 min (100% buffer B), gradient 100%/min for 1 min, 60–75 min (0% buffer B). For the fourth and fifth injection the gradient between 11 and 45 min was adjusted to 22%–42% (0.59%/min) and 22%–36% (0.41%/min), respectively. The 22mer sample buffer system contained additional 10% (v/v) acetonitrile as cosolvent for both buffers A and B. Equilibration was performed as for the 29mer whereas the three injections were structured as follows: 0% buffer B for 0–15 min, gradient of 20%/min for 1 min, gradient of 0.2%/min from 16–46 min, reaching 26% buffer B, gradient 74%/min for 1 min, 47–62 min (100% buffer B), gradient 100%/min for 1 min, 63–80 min (0% buffer B). For the 46mer, the temperature was 75°C and the flow rate was 7.0 mL/min. The gradient was: 0% buffer B for 0–10 min, gradient of 20%/min for 1 min, gradient of 0.51%/min for 39 min, gradient of 20%/min for 3 min, 53–63 min (100% buffer B), gradient of 33.3%/min for 3 min, 66–80 min (0% buffer B). For the 82mer the temperature was 75°C and the flow rate was 7.0 mL/min. The gradient was: 0% buffer B for 0–10 min, gradient of 25%/min for 1 min, gradient of 0.64%/min for 39 min, gradient of 16.6%/min for 3 min, 53–63 min (100% buffer B), gradient of 33.3%/min for 3 min, 66–80 min (0% buffer B). Fractions from each individual injection were analyzed on PAGE-gel (20% polyacrylamide, 8 M Urea) detected with ethidium bromide and for the 46mer and 82mer also with SYBRGold on a WWR Genoview mini 5 UV-board and fractions of satisfying purity were pooled together concentrated down and buffer exchanged into a final sample, fractions considered more and less pure were pooled and concentrated separately.

Yield calculations

The maximum possible molar yield for the RNA material was calculated as molar amount of NTP (for the most abundant nucleotide in the transcript sequence) divided by the number of occurrences of that nucleotide in the sequence. Percentage yields were then calculated as:

\[
\text{Percentage yield} = \frac{\text{Obtained molar amount RNA}}{\text{Maximum possible molar amount RNA}} \times 100.
\]

For the 22mer and the 29mer, these calculations were done based on A₂₆₀ values for the concentrated final material from all 10 mL of transcription reaction. For the 46mer and 82mer injections corresponding to 1/3 (~3.3 mL) of the total amount of material were done. A₂₆₀ values from the material from these injections were then used to estimate total amount of material and yield as if all 10 mL of transcription reaction would have been purified by subsequent injections.
Large-scale RNA sample preparation

ACKNOWLEDGMENTS

The authors are grateful to the Swedish Foundation for Strategic Research (project no. ICA14-0023), the Swedish Research Council (2014-04303), the Ragnar Söderberg Foundation (M91/14), and the Karolinska Institute and the department of Medical Biochemistry and Biophysics (KID 2-3707/2013 and the support from the Karolinska Institute and the department of Medical Research (project no. ICA14-0023), the Swedish Research Council (Dnr 2015-11949), and the Swedish Foundation for Strategic Research (project no. ICA14-0023). The ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. Nature 495: 56–74. doi:10.1038/nature11815

Funding. The Swedish Research Council (Dnr 2015-11949), the Swedish Foundation for Strategic Research (project no. ICA14-0023), and the Swedish Research Council (Dnr 2015-11949). The ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. Nature 495: 56–74. doi:10.1038/nature11815

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

REFERENCES

Akoopie A, Müller UF. 2018. Cotranscriptional 3′-end processing of T7 RNA polymerase transcripts by a smaller HDV ribozyme. J Mol Evol 86: 425–430. doi:10.1007/s00239-018-9861-9
Andersson AC, Scaringe SA, Earp BE, Frederick CA. 1996. HPLC purification for crystallography and NMR. Anal Biochem 231: 201–209. doi:10.1006/abio.1995.1521
Baronti L, Karlsson H, Marušić M, Petzold K. 2018. A guide to largescale RNA sample preparation. Anal Bioanal Chem 410: 3239–3252. doi:10.1007/s00216-018-9493-8
Baronti L, Guzzetti I, Ebahrimi P, Sandoz SF, Steiner E, Schlagintweit J, Fromm B, Silva L, Fontana C, Chen A, et al. 2020. Base-pair conformational switch modulates miR-34a targeting of Sirt1 mRNA. Nature. doi:10.1038/s41586-020-2336-3
Bartel DP. 2018. Metazoan microRNAs. Cell 173: 20–51. doi:10.1016/j.cell.2018.03.006
Beckert B, Masquida B. 2011. Synthesis of RNA by in vitro transcription. Methods Mol Biol 703: 9–41. doi:10.1007/978-1-59745-248-9_3
Chen Y, Pollack L. 2016. SAXS studies of RNA: structures, dynamics, and interactions with partners: SAXS studies of RNA. WIREs RNA 7: 512–526. doi:10.1002/wrna.1349
Chen Z, Ruffner DE. 1996. Modified crush-and-soak method for recovering oligodeoxynucleotides from polycrylamide gel. BioTechniques 21: 820–822. doi:10.2144/96215bm14
Chen Z, Zhang Y. 2005. Dimethyl sulfoxide targets phage RNA polymerases to promote transcription. Biochem Biophys Res Commun 333: 664–670. doi:10.1016/j.bbrc.2005.05.166
Cunningham L, Kättikamron K, Lu Y. 1996. Preparative-scale purification of RNA using an efficient method which combines gel electrophoresis and column chromatography. Nucleic Acids Res 24: 3647–3648. doi:10.1093/nar/24.18.3647
Dethoff EA, Petzold K, Chugh J, Casiano-Negroni A, Al-Hashimi HM. 2012. Visualizing transient low-populated structures of RNA. Nature 491: 724–728. doi:10.1038/nature11498
Djebali S, Davis CA, Merkel A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, et al. 2012. Landscape of transcription in human cells. Nature 489: 101–108. doi:10.1038/nature11233
Duss O, Mars C, Von Schroetter C, Alain FH-T. 2010. A fast, efficient and sequence-independent method for flexible multiple segmental isotope labeling of RNA using ribozyme and RNase H cleavage. Nucleic Acids Res 38: e188. doi:10.1093/nar/gkq756
Duss O, Lukavsky PJ, Alain FH-T. 2012. Isotope labeling and segmental labeling of larger RNAs for NMR structural studies. In Isotope labeling in biomolecular NMR (ed. Atreya HS), pp. 121–144. Springer, Dordrecht/New York.
Easton LE, Shibata Y, Lukavsky PJ. 2010. Rapid, nondenaturing RNA purification using weak anion-exchange fast performance liquid chromatography. RNA 16: 647–653. doi:10.1261/ma.1862210
The ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. Nature 495: 57–74. doi:10.1038/nature11247
Fadouloglou VE. 2013. Electroelution of nucleic acids from polycrylamide gels: a custom-made, agarose-based electroeluter. Anal Biochem 437: 49–51. doi:10.1016/j.ab.2013.02.021
Ferre-D’Amare AR, Doudna JA. 1996. Use of cis- and trans-ribozymes to remove 5′ and 3′ heterogeneties from milligramms of in vitro transcribed RNA. Nucleic Acids Res 24: 977–978. doi:10.1093/nar/24.5.977
Feyrer H, Munteanu R, Baronti L, Petzold K. 2020. One-pot production of RNA in high yield and purity through cleaving tandem transcripts. Molecules 25: 1142. doi:10.3390/molecules25051142
Gorski SA, Vogel J, Doudna JA. 2017. RNA-based recognition and targeting: sowing the seeds of specificity. Nat Rev Mol Cell Biol 18: 215–228. doi:10.1038/nrm.2016.174
Guillerez J, Lopez P, Proux F, Launay H, Dreyfus M. 2005. A mutation of T7 RNA polymerase that facilitates promoter clearance. Proc Natl Acad Sci 102: 5958–5963. doi:10.1073/pnas.0407141102
Hagen SH, Young ET. 1974. Preparative polyacrylamide gel electrophoresis of ribonucleic acid. Identification of multiple molecular species of bacteriophage T7 lysozyme messenger ribonucleic acid. Biochemistry 13: 3394–3400. doi:10.1021/bi00713a033
Heiling C, Keyhani S, Sochor F, Fürtig B, Hengesbach M, Schwalbe H. 2015. Rapid NMR screening of RNA secondary structure and binding. J Biomol NMR 63: 67–76. doi:10.1007/s10858-015-9967-y
Vinothkumar KR, Henderson R. 2016. Single particle electron cryomicroscopy: trends, issues and future perspective. Q Rev Biophys 49: e13. doi:10.1017/S0033583516000068

Weeks KM, Mauger DM. 2011. Exploring RNA structural codes with SHAPE chemistry. Acc Chem Res 44: 1280–1291. doi:10.1021/ar200051h

Weissman D, Pardi N, Muramatsu H, Karikó K. 2013. HPLC purification of in vitro transcribed long RNA. In Synthetic messenger RNA and cell metabolism modulation (ed. Rabinovich PM), pp. 43–54. Humana Press, Totowa, NJ.

Wichlacz A. 2004. Generating in vitro transcripts with homogenous 3’ ends using trans-acting antigenomic delta ribozyme. Nucleic Acids Res 32: e39. doi:10.1093/nar/gnh037

Zhang H, Keane SC. 2019. Advances that facilitate the study of large RNA structure and dynamics by nuclear magnetic resonance spectroscopy. Wiley Interdiscip Rev RNA 10: e1541. doi:10.1002/wrna.1541