Published online 11 November 2010

Nucleic Acids Research, 2011, Vol. 39, No. 3 ***e18***
doi:10.1093/nar/gkq1084

The ARiBo tag: a reliable tool for affinity purification of RNAs under native conditions

Geneviève Di Tomasso, Philipe Lampron, Pierre Dagenais, James G. Omichinski and Pascale Legault*

Département de Biochimie, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, QC, H3C 3J7 Canada

Received August 5, 2010; Revised October 14, 2010; Accepted October 15, 2010

ABSTRACT

Although RNA-based biological processes and therapeutics have gained increasing interest, purification of *in vitro* transcribed RNA generally relies on gel-based methods that are time-consuming, tedious and denature the RNA. Here, we present a reliable procedure for affinity batch purification of RNA, which exploits the high-affinity interaction between the *boxB* RNA and the N-peptide from bacteriophage λ. The RNA of interest is synthesized with an ARiBo tag, which consists of an activatable ribozyme (the *glmS* ribozyme) and the λ*BoxB* RNA. This ARiBo-fusion RNA is initially captured on Glutathione-Sepharose resin via a GST/λN-fusion protein, and the RNA of interest is subsequently eluted by ribozyme self-cleavage using glucosamine-6-phosphate. Several GST/λN-fusion proteins and ARiBo tags were tested to optimize RNA yield and purity. The optimized procedure enables one to quickly obtain (3 h) highly pure RNA (>99%) under native conditions and with yields comparable to standard denaturing gel-based protocols. It is widely applicable to a variety of RNAs, including riboswitches, ribozymes and microRNAs. In addition, it can be easily adapted to a wide range of applications that require RNA purification and/or immobilization, including isolation of RNA-associated complexes from living cells and high-throughput applications.

INTRODUCTION

Several recent discoveries have emphasized the importance of RNA-based processes in biology and brought RNA molecules at the forefront of basic and applied biomedical research. As a result, there has been an increased demand to quickly generate large amount of RNAs that are chemically pure and folded in their native conformation. The traditional approach to purify RNA produced from *in vitro* transcription relies on denaturing polyacrylamide gel electrophoresis (1–4). This methodology involves denaturation of the RNA molecule (5), often results in contamination of the purified RNA with acrylamide oligomers (6) and is generally very time consuming and tedious. Therefore, alternative purification methods have recently been developed to purify RNA in a time-efficient manner and under non-denaturing conditions. Among these approaches, size-exclusion and ion-exchange chromatography present several advantages (6–11), however, affinity purification is very rapid and can be easily adapted to any molecular weight RNA and to high-throughput applications (10,12–17).

At this time, only a few procedures for affinity purification of *in vitro* transcribed RNA have been reported (12–17). They all incorporate four main steps: (i) transcription of a hybrid RNA that contains both the RNA of interest and a 3′-affinity tag; (ii) immobilization of the transcribed RNA on an affinity matrix; (iii) removal of impurities from the affinity matrix; and (iv) elution of the target RNA by cleavage of the affinity tag. The RNA immobilization and cleavage steps are the most critical aspects of the procedure. Several strategies have been employed for RNA immobilization, including RNA/DNA hybridization and high-affinity RNA/protein interactions (12–17). RNA cleavage has been achieved in trans using a DNAzyme, however it requires additional purification steps to remove the co-eluting enzyme (12). The use of RNA tags containing activatable ribozymes has been shown to substantially simplify the procedure, since no additional purification step, other than a buffer exchange, is required after the RNA elution (13–17). Recently, an elegant protocol was reported that exploits a His-tagged MS2 coat protein attached to a Ni-NTA resin for RNA immobilization and the efficient *glmS* ribozyme activated by

*To whom correspondence should be addressed. Tel: 514 343 7326; Fax: 514 343 2210; Email: pascale.legault@umontreal.ca

glucosamine-6-phosphate (GlcN6P) for RNA elution (14). This procedure allows for non-denaturing purification of μg to mg quantities of RNA in only a few hours. However, like other procedures reported so far for affinity purification of transcribed RNA, it was not developed to maximize RNA purity and yield. Thus, it is still not clear if this or other known affinity purification procedures can consistently produce RNA with the yields and purity required for the most demanding applications, such as accurate biochemical, biophysical and structural characterizations.

In this work, we establish a quick and reliable affinity batch purification method that yields highly pure native RNA with yields comparable to standard denaturating gel methods. Our procedure is based on a novel affinity tag, termed ARiBo tag, which contains an Activatable Ribozyme (the *glmS* ribozyme) and the *BoxB* RNA from bacteriophage λ. The RNA is first transcribed as an ARiBo-fusion RNA, captured on Glutathione-Sepharose (GSH-Sepharose) resin via a Glutathione-S-Transferase (GST) fusion with the λN peptide and eluted by self-cleavage of the *glmS* ribozyme upon activation by glucosamine-6-phosphate (GlcN6P). A simple gel-based method is described that provides a quantitative evaluation of the performance of our method.

MATERIALS AND METHODS

Cloning of protein expression vectors

All vectors used for protein expression are illustrated in Supplementary Figure S1. The pGEX2T-λN plasmid was described previously (18). For pGEX2T$_{Td}$-L-λN, a DNA fragment coding for G_8-λ$N_{1–22}$ was inserted at the BamH1 and EcoR1 sites of pGEX2T (GE Healthcare), and the thrombin cleavage-site was destroyed (LVPRGS to LVPGGS) by mutagenesis to prevent cleavage of the fusion protein by contaminant proteases. All mutageneses reported here were conducted according to the Stratagene QuikChangeII procedure. For pET42a-λN-L-GST, a DNA fragment coding for λ$N_{1–22}$-G_8 was inserted into the NdeI site of the pET42a vector (Novagen), and a stop codon was created by mutagenesis at the end of the GST-coding sequence. Mutagenesis of pET42a-λN-L-GST was carried out to introduce the M1G/D2N/Q4K mutations (19) within the coding sequence of the λN peptide, which yielded pET42a-λN^+-L-GST. Mutagenesis of this latter vector was carried out to change the G_8-linker sequence to $(GA)_{10}$ and thereby obtain pET42a-λN^+-L^+-GST. The pET42a-2λN^+-GST plasmid was obtained as a by-product of a multi-step cloning procedure. First, NheI and AatII restriction sites were introduced by mutagenesis just upstream of the GST-coding sequence of pET42a-λN^+-L^+-GST. A PCR fragment generated from pET42a-λN^+-L-GST and coding for λN^+-G_8 was inserted within the new NheI and AatII sites. All sequences were verified by DNA sequencing.

Expression and purification of GST-fusion proteins

The pGEX2T-based plasmids used for expression of GST-λN and GST-L-λN were transformed into BL21 cells, whereas the pET42a-based plasmids used for expression of λN-L-GST, λN^+-L-GST, λN^+-L^+-GST and 2λN^+-GST were transformed into BL21(DE3) cells (Figure 3). All cells were grown at 37°C in Luria–Bertani media, and protein expression was induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) for 4 h at 30°C. The cells were harvested by centrifugation and resuspended in homogenization buffer [20 mM Tris pH = 7.4, 1 M NaCl, 1 mM DTT, 0.2 mM EDTA and 0.15% w/v protease inhibitor cocktail (Sigma-Aldrich)]. The cells were lysed by French press, sonicated 10 s and centrifuged at 138 000*g* for 1 h (4°C). The supernatant was incubated for 1 h at 4°C with GSH-Sepharose 4B resin (GE Healthcare). The resin was washed two times with wash buffer (homogenization buffer with 2 M urea) and then two times with PBS buffer (10 mM Na_2HPO_4, 2 mM KH_2PO_4, 2.7 mM KCl, 140 mM NaCl and pH = 7.4). The GST-fusion proteins were eluted from the bound resin by two incubations of 15 min at room temperature with 20 mM reduced L-glutathione pH = 8.0 (Sigma-Aldrich). The supernatant was dialyzed overnight at 4°C in FPLC-A buffer (20 mM sodium phosphate pH = 7.4, 1 mM EDTA and 1 mM DTT) and then applied to an SP Sepharose high-performance column (GE Healthcare; 100 ml bed volume) equilibrated with FPLC-A buffer. The proteins were eluted from the column using a gradient (from 0 to 100% over 525 ml) of FPLC-B buffer (FPLC-A with 2 M NaCl). The pooled fractions containing the protein of interest were dialyzed into water and then lyophilized. The high purity (≥95%) and correct mass (with error ≤ ±1.4 Da) of all purified proteins were verified by SDS–PAGE and mass spectrometry, respectively.

Cloning of the pARiBo-based plasmids

All transcription vectors are depicted in Supplementary Figure S2. To generate the pARiBo1 plasmid, a DNA fragment was generated which contains the T7 promoter (5′-TAATACGACTCACTATA-3′) and codes for a 5′-GGCGAA-3′ sequence followed by the ARiBo1 RNA. This fragment was inserted between the HindIII and EcoR1 sites of the pTZ19R-derived pTR-4 vector (20). The sequence of the ARiBo1 RNA was designed to incorporate an ApaI restriction site in the P1 helix (Figure 2). For the pRSA$_{U65C}$-ARiBo1 plasmid, a DNA fragment containing the T7 promoter and coding for the RSA$_{U65C}$ RNA was first generated by PCR amplification of the pRSA$_{U65C}$-VS plasmid (21) and then inserted between the HindIII and ApaI sites of the pARiBo1 plasmid. The pRSA$_{U65C}$-ARiBo2 and pRSA$_{U65C}$-ARiBo3 plasmids were obtained by mutagenesis of the pRSA$_{U65C}$-ARiBo1 plasmid using the QuikChangeII site-directed mutagenesis procedure. The sequences of the ARiBo2 and ARiBo3 RNAs were designed to incorporate a KpnI restriction site in the P1 helix (Figure 2). All sequences were verified by DNA sequencing.

In vitro transcription of ARiBo-fused RNAs

Large-scale preparations (~2–3 mg) of plasmid DNA template were typically obtained by growing 0.5 l of plasmid-transformed DH5α cells (Invitrogen), purifying the plasmid using the QIAGEN Plasmid Mega Kit and linearizing it overnight with EcoRI (New England Biolabs). The ARiBo-fused RNAs were transcribed at 37°C for 3 h using the following reaction conditions: 40 mM HEPES pH = 8.0, 50 mM DTT, 0.1% Triton X-100, 1 mM spermidine, 4 mM ATP, 4 mM CTP, 4 mM UTP, 8 mM GTP, 25 mM $MgCl_2$, 60 μg/ml His-tagged T7 RNA polymerase, 3 U/ml RNAsin Ribonulease Inhibitor (Promega) and 80 μg/ml of linearized plasmid DNA template. Transcription reactions were stopped by adding 50 mM EDTA and stored at −20°C.

Small-scale affinity batch purification of ARiBo-fused RNAs

For typical small-scale purifications, 35 nmol of GST/λN-fusion protein was first added to a small transcription volume (~140 μl for RSA_{U65C}) that corresponds to 7 nmol of ARiBo-fused RNA, and the total volume was adjusted to 800 μl with equilibration buffer (50 mM HEPES pH = 7.5). The RNA-protein mix was then incubated for 15 min in a 1.5-ml conical tube. Unless otherwise mentioned, all incubations were done with gentle rotation at room temperature. In a separate tube, 325 μl of GSH-Sepharose 4B resin slurry (GE Healthcare) was washed twice with 800 μl of PBS buffer. The RNA-protein mix was added to the washed resin and incubated for 15 min, centrifuged 1 min at 1500*g* and the load supernatant was kept for quantitative analysis on gel (LS). The pelleted resin was washed three times with 800 μl equilibration buffer. These and all subsequent washes involved incubation for 5 min and centrifugation for 1 min at 1500*g*. The wash supernatants were kept for quantitative analysis (W1, W2, W3). Elution of the desired RNA (RSA_{U65C}) was induced by leaving the pelleted resin at 37°C for 10 min in 800 μl Elution buffer (20 mM Tris buffer pH = 7.6, 10 mM $MgCl_2$ and 500 μM GlcN6P) and transferring for 5 min at room temperature prior to centrifugation. The elution supernatant contained the RNA of interest (E1). The pelleted resin was washed two times with 800 μl equilibration buffer, and the elution-wash supernatants were kept for quantitative analysis (E2, E3). To prevent loss of resins during the wash and elution steps, all supernatants were centrifuged for 1 min at 1500*g* and the minute amount of pelleted resin was mixed with the buffer used for the next purification step. To remove RNA left on the resin after elution, the pelleted resin was washed with 800 μl of 2.5 M NaCl. The supernatant was kept for quantitative analysis (NaCl). For complete matrix regeneration, the GSH-Sepharose resin was subsequently washed with PBS, 20 mM L-glutathione pH = 8 in PBS and then with 20% ethanol for storage.

Quantitative analysis of small-scale affinity batch purifications

For quantitative analysis, each small-scale affinity batch purification was performed at least three times, and purifications made from the same ARiBo-fused RSA_{U65C} precursor were performed from the same transcription reaction. Aliquots from the various steps of purification were analyzed by denaturing-gel electrophoresis. Care was taken to load gels with sample volumes corresponding to precise amounts of RSA_{U65C}-ARiBo-fusion RNA present in the transcription reaction. The gels were stained for 10 min in a SYBR Gold (Invitrogen) solution [1:10 000 dilution in TBE buffer (200 mM Tris-Base, 200 mM boric acid and 4 mM EDTA)] and scanned on a Molecular Imager FX densitometer (Bio-Rad). The band intensities were analyzed using the QuantityOne software (version 4.4.1 from Bio-Rad).

For each gel, several control lanes were loaded with known amounts of RNA to derive three standard curves that were used to determine the quantity (in ng) of RSA_{U65C} (N_{RNA}), ARiBo tag (N_{ARiBo}) and RSA_{U65C}-ARiBo fusion RNA (N_{Fusion}) at each purification step. For the RSA_{U65C} standard curve, the quantities of RNA loaded on the gel were obtained from OD_{260} measurements. For the standard curves of the ARiBo-tag and the RSA_{U65C}-ARiBo-fusion RNAs, the quantities of RNA loaded on the gel were calculated from the quantity of RSA_{U65C} detected in transcription reactions treated with GlcN6P to achieve ≥99% ARiBo tag cleavage.

The 'percentage of unbound RNA' was calculated using the equation $[(\Sigma N_{Fusion})/I_{Fusion}] \times 100\%$, where ΣN_{Fusion} represents the total amount of fusion RNA (ng) detected in lanes LS, W1, W2 and W3, and I_{Fusion} represents the input of the same RNA in equivalent volumes of affinity batch purification (250 ng). The percentage of unbound RNA is given as a minimum value, since it is only based on the amount of fusion RNA that migrates as such on the gel; slower migrating species have been observed in certain cases and probably represent some forms of RNA aggregates, but were not quantified. The 'percentage of cleavage in solution' was determined from a control lane in which the transcription reaction was treated with GlcN6P (lane 19) using the equation $\{(N_{ARiBo}/nt_{ARiBo}) / [(N_{ARiBo}/nt_{ARiBo}) + (N_{Fusion}/nt_{Fusion})]\} \times 100\%$, where nt_{ARiBo} and nt_{Fusion} represent the number of nucleotides for the ARiBo-tag and fusion RNAs, respectively. The same equation was used to calculate 'the percentage of cleavage on the resin', although this was derived from the NaCl lane (lane 20). The 'percentage of RNA' eluted was calculated using the equation $[(\Sigma N_{RNA})/I_{RNA}] \times 100\%$, where ΣN_{RNA} represents the total amount of RSA_{U65C} (ng) detected in lanes E1, E2 and E3, and I_{RNA} represents the calculated amount of RSA_{U65C} expected from 100% cleavage in equivalent volumes of transcription (100 ng). The 'percentage of RNA purity' was calculated from the E1 lane (lane 9) using the equation $[N_{RNA}/(N_{RNA}+N_{ARiBo})] \times 100\%$.

Large-scale affinity batch purification of ARiBo-fused RNAs

A large-scale purification of ARiBo1-fused RSA_{U65C} was processed in 50-ml conical tubes, similarly to the small-scale purification with the λN^{+}-L^{+}-GST fusion protein, but increasing all volumes approximately 30 times (25-ml wash and elution buffers). In addition, an alkaline phosphatase step was inserted between the first (W1) and second washes (W2). This consisted of a 4-h incubation at 37°C in 25 ml of CIP buffer (50 mM HEPES pH = 8.5 and 0.1 mM EDTA) with 130 U of calf intestine alkaline phosphatase (Roche) per micromole of RNA, followed by a 5-min incubation at room temperature prior to centrifugation. The supernatant was kept for analysis (CIP). The purified RNA (E1, E2 and E3 fractions) was concentrated with Amicon Ultra-15 centrifugal filter devices (Millipore) and exchanged in NMR buffer (10 mM sodium cacodylate pH = 6.5, 50 mM KCl, 5 mM $MgCl_2$, 0.05 mM NaN_3 in 90% H_2O/10% D_2O).

In vitro transcription of RNA and purification by denaturing gel electrophoresis

RSA_{U65C} was also synthesized as an RSA_{U65C}-VS precursor containing a *Varkud* Satellite (VS) ribozyme substrate at its 3′-end. *In vitro* transcription and purification of RSA_{U65C} by denaturing-gel electrophoresis was performed as described earlier (21), except that the HPLC purification step using a DNA-Pac100 column heated at 65°C was replaced by gravity-flow anion-exchange chromatography with DEAE-Sephacel at room temperature.

NMR spectroscopy studies

The 1D ^{1}H flip-back watergate spectra were collected at 15°C on a Varian ^{Unity}INOVA 500 MHz spectrometer equipped with a pulse-field gradient unit and an actively-shielded z gradient $^{1}H\{^{13}C/^{15}N\}$ triple resonance probe.

RESULTS AND DISCUSSION

General scheme for affinity purification of RNA using ARiBo tags

To develop an efficient affinity-purification procedure that maximizes RNA yield and purity, we selected some of the most reliable reagents available for RNA immobilization and elution. For RNA immobilization, we exploited GST/λN-fusion proteins attached to a GSH-Sepharose matrix. The natural λN and its cognate *boxB* RNA form a very stable and specific interaction (K_D ~2–20 nM), and increased stability can be obtained using engineered λN peptides [$K_D \geq 10$ pM; (19,22)]. The GST/GSH-Sepharose system is one of the most affordable, versatile and commonly used affinity methods for purification of recombinant proteins expressed in *Escherichia coli* (*E. coli*). The GST/GSH-Sepharose interaction is compatible with all commonly used aqueous buffers, yet easily reversible by addition of free glutathione. For RNA elution, we exploited the *glmS* ribozyme from *Bacillus anthracis* (23–25). The *glmS* ribozyme self-cleaves quickly and efficiently when activated by GlcN6P, and displays very low background activity in the absence of GlcN6P, Tris and related compounds (23,25–27). Since it requires only a single adenine nucleotide at its 5′-end, the *glmS* ribozyme is also ideally suited to cleave the 3′-end of a transcript (23). We combined this Activatable Ribozyme with the λ*BoxB* RNA to create a novel affinity tag, termed the ARiBo tag.

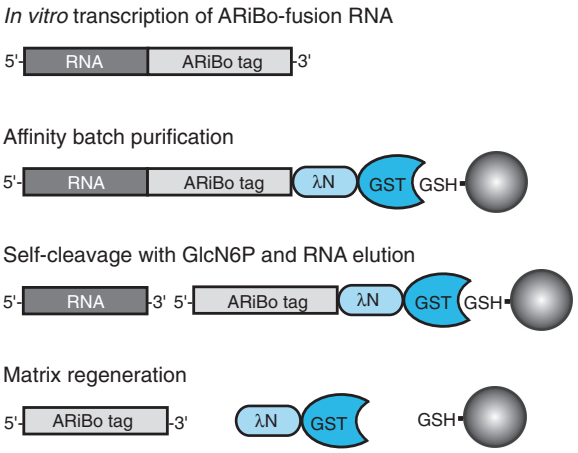

Figure 1. General strategy for affinity batch purification of RNA based on the λ*boxB*/λN-peptide interaction. In this schematic, the RNA is fused to an ARiBo tag and purified via binding to a λN peptide fused to a GST protein, for immobilization on GSH-Sepharose beads. RNA elution is triggered by addition of GlcN6P, which activates the *glmS* ribozyme of the ARiBo tag. The resin can be regenerated by step-wise incubations with 2.5 M NaCl and 20 mM GSH.

The general purification strategy is outlined in Figure 1. The RNA of interest is first transcribed with an ARiBo tag at its 3′-end. The ARiBo-fusion RNA is then bound to a GST/λN-fusion protein, and the resulting complex is captured on GSH-Sepharose resin. After washing to remove impurities, the RNA is eluted by self-cleavage of the *glmS* ribozyme following the addition of GlcN6P to activate the ribozyme. As needed, the resin can be regenerated using 2.5 M NaCl to remove the affinity tag and 20 mM GSH to liberate the GST/λN-fusion protein.

Development of an optimal affinity batch purification method

Although affinity purifications are often performed in a gravity-column format, we explored a batch method, because it is time efficient, ideally suited for purification from crude preparations and easily amenable to enzymatic RNA processing and high-throughput applications. To develop the method, we purified a mutant of the pbuE adenine riboswitch aptamer from *B. subtilis* (RSA_{U65C}; Figure 2A), because we have previously purified this non-coding RNA by a standard gel method (21). Several ARiBo tags (Figure 2B) and GST/λN-fusion proteins (Figure 3) were tested to develop an optimum protocol for affinity batch purification of RNA.

Initially, we performed affinity batch purification using the ARiBo1 tag (Figure 2B) and the GST-λN fusion

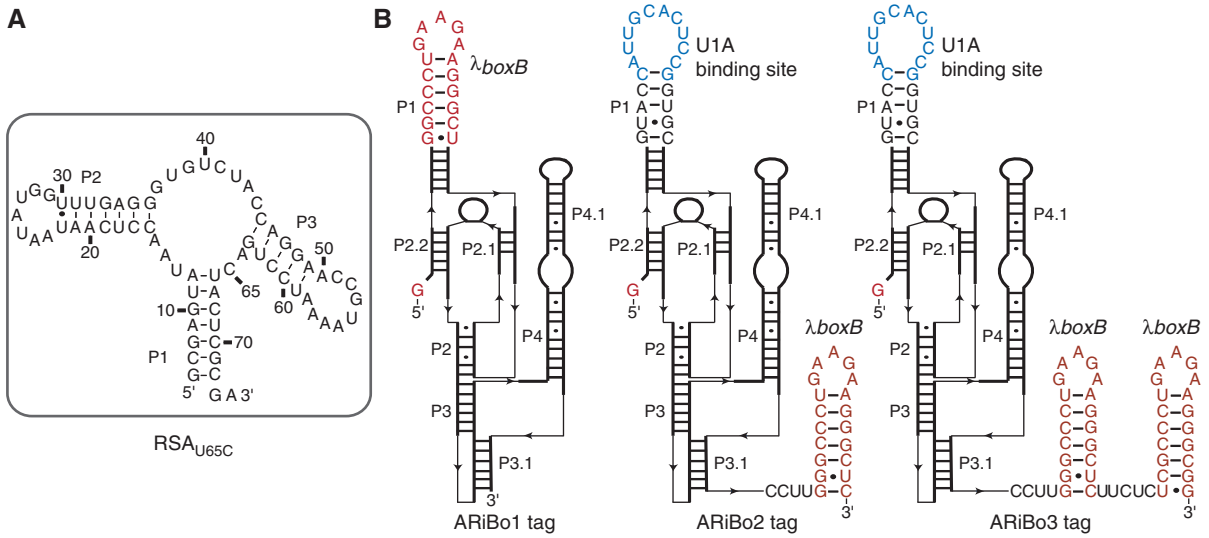

Figure 2. Primary and secondary structures of (**A**) the RNA of interest, the U65C mutant of the *B. subtilis* pbuE adenine riboswitch aptamer, and (**B**) the ARiBo tags tested in this study. The ARiBo tags contain the *B. anthracis glmS* sequence (24,25) except for P1 substitutions and 3′ extensions with the λ*boxB* RNA sequence or the U1A binding site, as shown. In the ARiBo-fused RNAs, the GA at the 3′-end of RSA$_{U65C}$ (A) connects to the G at the 5′-end of each ARiBo tag (G+1 of the *glmS* ribozyme in B).

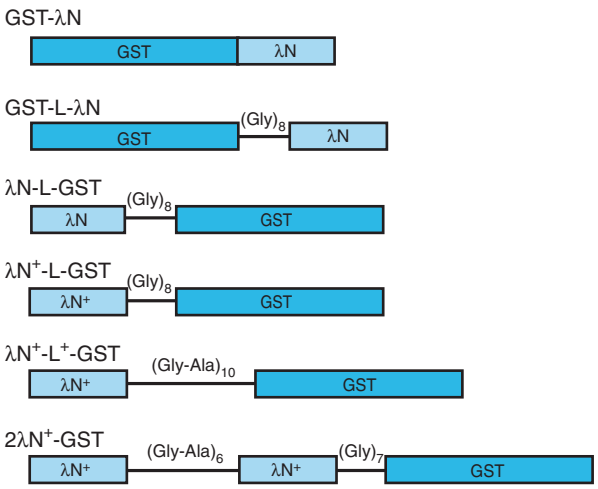

Figure 3. Nomenclature and schematic diagram of the different GST/λN fusion proteins tested in this study. The λN peptide contains the first 22 residues of the λN protein and the λN$^+$ peptide is a G1N2K4 triple mutant of the λN peptide (19).

protein in which the GST was directly fused to the N terminus of λN (Figure 3). Examination of aliquots collected at the various steps of the purification revealed that the eluted RNA was contaminated with the affinity tag and the yield was rather poor (Figure 4A). Similar results were obtained using a GST-L-λN fusion protein, in which a linker sequence was inserted between the C terminus of the GST and the N terminus of λN (Figure 3; results not shown). Although, the results with the GST-λN and GST-L-λN fusion proteins showed promise, there was a need for improvement.

In the NMR structure of the λ*boxB*/λN complex, the N-peptide is oriented such that C-terminal residues interact with the *boxB* loop, while N-terminal residues interact with the opposite part of the *boxB* stem (22). Thus, the GST fused to the N terminus of λN likely interferes with *boxB* binding, especially in the context of the ARiBo tag where the *boxB*/P1 stem is connected to the *glmS* ribozyme (25). In contrast, the GST fused to the C terminus of λN would position the GST protein away from both the P1 stem and the *glmS* ribozyme and this should be less likely to interfere with *boxB* binding. Thus, in an attempt to improve the performance of the method, the GST was attached to the C terminus of λN. Several such GST/λN-fusion proteins were tested (λN-L-GST, λN$^+$-L-GST, λN$^+$-L$^+$-GST and 2λN$^+$-GST; Figure 3), which differ according to the N-peptide sequence, the linker sequence and the number (1 or 2) of λN$^+$ peptide repeats (Figure 3). The N-peptide sequence was either the wild-type (λN) or a high-affinity variant (λN$^+$) sequence [Supplementary Figure S1A; (19)]. All these fusion proteins in which GST is fused to the C terminus of N provided improved RNA yield and purity, as illustrated for the λN$^+$-L$^+$-GST fusion protein in Figure 4B.

The optimal tethering system

To systematically evaluate the performance of the method using each GST/λN fusion protein, we performed a quantitative analysis of each step of the procedure (Table 1). The efficiency of RNA capture was evaluated from the percentage of unbound fusion RNA in the load supernatant and washes. The high percentage of unbound fusion RNA observed for the GST-λN protein (≥44%) compared to other GST/λN fusion proteins (≥8–22%) indicates a lower efficiency of RNA capture for the GST-λN protein. However, the percentage of RNA self-cleavage on the resin was very efficient for all GST/λN fusion proteins; it was only slightly lower on the resin (93–99%) than in the transcription reaction (>99%). Hence, immobilization of the RNA on the resin using any GST/λN fusion protein did not significantly affect the efficiency of self-cleavage by the *glmS* ribozyme. The percentage of RNA eluted with respect to the expected yield from the transcription

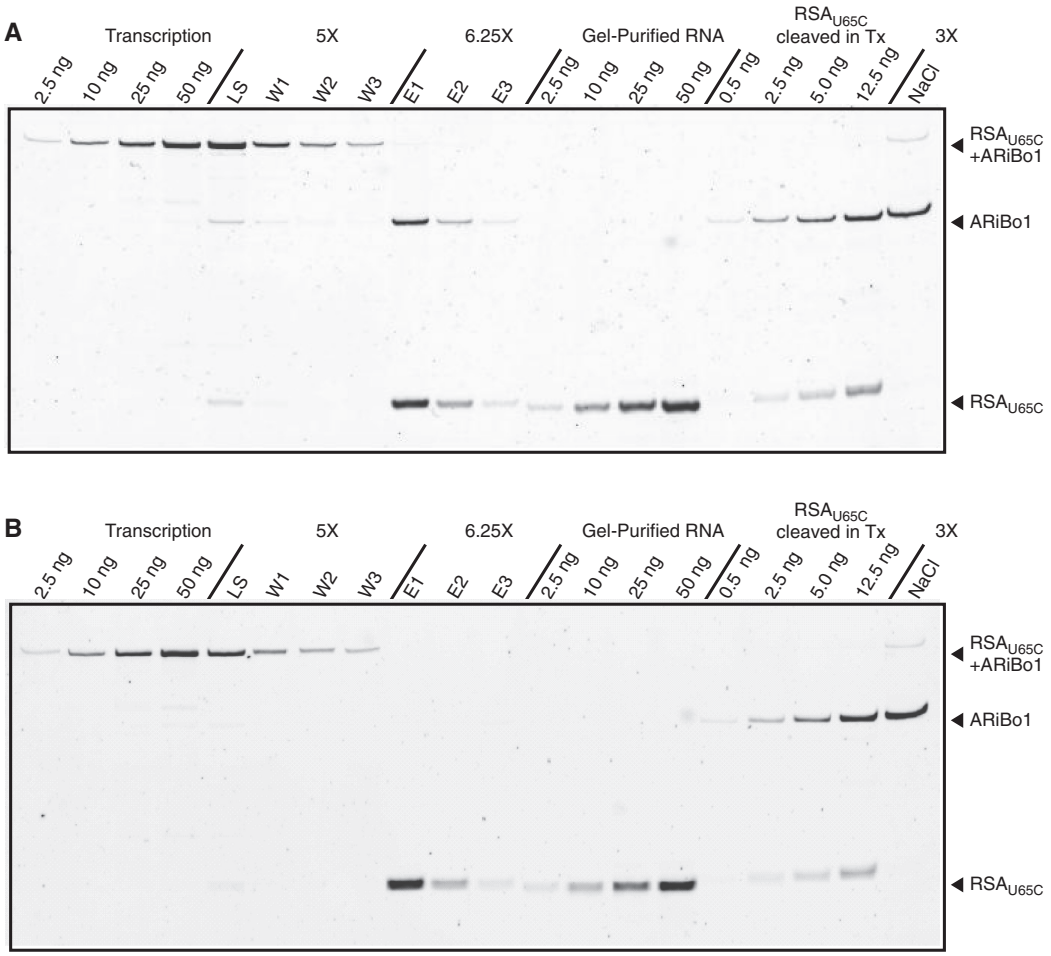

Figure 4. Typical small-scale affinity batch purifications of RSA_{U65C} analyzed on a denaturing polyacrylamide gel stained with SYBR Gold. The RSA_{U65C} was synthesized as an ARiBo1-fused RNA. The GST/λN fusion protein used was either GST-λN (**A**) or λN^+-L^+-GST (**B**). Aliquots from each purification step were loaded on the gel (LS: loading supernatant; W1–3: washes; E1–3: elutions; and NaCl: matrix regeneration with 2.5 M NaCl) in amounts shown, where 1× correspond to 50 ng of ARiBo-fused RSA_{U65C} present in the transcription reaction or 16 ng of RSA_{U65C} to be purified. In addition, standard amounts of ARiBo-fused RSA_{U65C} from the transcription reaction, purified RSA_{U65C} and RSA_{U65C} cleaved in the transcription reaction were loaded for quantitative analysis of the purification ('Materials and Methods' section). Bands corresponding to the ARiBo-fused RSA_{U65C}, the ARiBo1 tag and the desired RNA (RSA_{U65C}) are indicated on the right of the gel.

Table 1. Results of affinity batch purification of RSA_{U65C} using the ARiBo1 tag and different GST/λN fusion proteins

Fusion proteins	GST-λN	GST-L-λN	λN-L-GST	λN^+-L-GST	λN^+-L^+-GST	$2\lambda N^+$-GST
Unbound RNA (%)	≥44 ± 4	≥20 ± 4	≥17 ± 1	≥22 ± 3	≥13 ± 1	≥8 ± 1
Cleavage in solution (%)	99.9 ± 0.1	99.8 ± 0.1	99.8 ± 0.1	99.7 ± 0.2	99.6 ± 0.2	99.6 ± 0.1
Cleavage on the resin (%)	97.8 ± 0.7	93 ± 3	96 ± 1	97 ± 2	98.7 ± 0.5	95 ± 2
RNA eluted (%)	39 ± 2	49 ± 2	57 ± 3	58 ± 2	64.6 ± 0.7	61 ± 3
RNA purity estimate (%)	70 ± 1	56 ± 2	96.2 ± 0.4	99.1 ± 0.3	99.86 ± 0.09	99.8 ± 0.2

reaction was <50% for the GST-λN (39 ± 2%) and GST-L-λN (49 ± 2%) and ranged between 54 and 65% for λN-L-GST, λN^+-L-GST, λN^+-L^+-GST and $2\lambda N^+$-GST. The percentage of RNA purity with respect to the main RNA contaminant in the sample (the ARiBo tag), was poor for GST-λN (70 ± 1%) and GST-L-λN (56 ± 2%), but >95% for λN-L-GST, λN^+-L-GST, λN^+-L^+-GST and $2\lambda N^+$-GST. Thus, the quantitative analysis confirms that the GST-λN and GST-L-λN fusion proteins are not effective for affinity batch purification, in agreement with our structural analysis of the λ*boxB*/λN complex. In contrast, when the GST was fused to the C terminus of λN the method provided higher level of RNA yield and purity. Greater than 99% purity was systematically obtained using GST/λN fusion proteins engineered with the G1N2K4 λN peptide variant that binds *boxB* RNA with picomolar affinity (19). The λN^+-L^+-GST fusion protein offers optimum performance

with an RNA yield of 64.6 ± 0.7% and RNA purity of 99.86 ± 0.09%.

A major advantage of this new affinity purification procedure is the high-purity level that is attained with λN^+–L^+-GST. This fusion protein is likely compatible with both the high-affinity GST/GSH-Sepharose and *λboxB*/λN^+ interactions, preventing leakage of the ARiBo tag during elution. In addition, this fusion protein is very stable and large quantities are easily purified (~25 mg/l media). For optimal performance and practicality, we used a 5-fold molar excess of λN^+-L^+-GST with respect to RNA. Reducing this fusion protein:RNA ratio from 5:1 to 4:1 and 3:1 did not affect the purity, but resulted in lower yield from 65% to 54% and 43%, respectively (Supplementary Table S1). With a 2:1 ratio, both the purity (97%) and yield (25%) were reduced (Supplementary Table S1). Thus, the lower 3:1 ratio could be employed when a high level of purity is required, but yield is not critical. We also attempted to further improve yields by increasing the fusion protein:RNA ratio up to 10:1. A maximum RNA yield of 75 ± 4% with purity of 99.74 ± 0.05% was obtained by using a 8:1 fusion protein:RNA ratio (not shown). However, we find that a 5:1 fusion protein:RNA ratio is more practical and economical for general use, especially for large-scale purifications of RNA.

The optimal ARiBo tag

The ARiBo1 tag was created to minimize the size of the affinity tag by incorporating the *λboxB* RNA in the variable apical P1 stem-loop of the *glmS* ribozyme (23,25,28). A small tag may be desirable to improve the yield of *in vitro* transcriptions in the presence of limiting, modified and/or expensive nucleotides (NTPs), for example when using isotopically-labeled NTPs for NMR studies (29,30) or non-standard NTPs for structure–function studies (31). To evaluate the efficiency of the ARiBo1 tag, we carried out control affinity purifications with the λN^+-L^+-GST fusion protein using the ARiBo2 and ARiBo3 tags, in which either one or two *boxB* RNAs were respectively positioned at the 3′-end of the *glmS* ribozyme (Figure 3). The quantitative analysis indicates that the ARiBo1 tag provides similar purity, but higher RNA yields than either the ARiBo2 or ARiBo3 tags (Table 2). Thus, by engineering a minimal affinity tag that combines the *λboxB* RNA and *glmS* ribozyme elements at the structural level rather than in a sequential manner, a more efficient affinity tag was designed.

Table 2. Results of affinity batch purification of RSA_{U65C} using the λN^+-L^+-GST fusion protein and different ARiBo tags

ARiBo tags	ARiBo1	ARiBo2	ARiBo3
Unbound RNA (%)	≥13 ± 1	≥23 ± 2	≥16 ± 3
Cleavage in solution (%)	99.6 ± 0.2	99.4 ± 0.3	99.2 ± 0.1
Cleavage on the resin (%)	98.7 ± 0.5	96 ± 3	97 ± 2
RNA eluted (%)	64.6 ± 0.7	49 ± 2	59 ± 2
RNA purity estimate (%)	99.86 ± 0.09	99.8 ± 0.2	99.87 ± 0.06

Examination of sequence requirements at the 3′-end of the RNA of interest

Previous studies on the *glmS* ribozyme indicate that an adenine immediately upstream of the cleavage site (N−1 position) is conserved in bacteria, minimally required for self-cleavage, and its mutation to guanine leads to reduced cleavage activity (23,27,28). The crystal structure of the *B. anthracis glmS* ribozyme bound to GlcN6P revealed that the A−1 forms two hydrogen bonds with G57 (25). Since, G57 also directly interacts with GlcN6P and the scissile phosphate, A−1 may help organize the cleavage site for catalysis (25).

To further examine the sequence requirements at the cleavage site of the *glmS* ribozyme, we modified the ARiBo1-fusion RSA_{U65C} such that the 2-nt GA linker ($G_{73}A_{74}$ in Figure 2A) was replaced by GG, GC, GU or a single A and tested for self-cleavage directly in the transcription reaction (Figure 5). Under standard cleavage conditions (0.5 mM GlcN6P, 20 mM Tris pH = 7.6 and 10 mM $MgCl_2$ for 15 min at 37°C), 99% cleavage was obtained for the GA and A linkers, but lower cleavage yields were obtained for the GG (84%), GU (11%) and GC (11%) linkers. Under these conditions, the cleavage efficiency is GA ~ A > GG > GC ~ GU (Figure 5). Structurally, a guanine may be able to partially substitute for A−1, however the C and U bases are smaller and may not be able to interact with G57, possibly affecting optimal binding of GlcN6P. By slightly increasing the GlcN6P concentration (2 mM) to favor GlcN6P binding, >99% cleavage was obtained in 30 min for the GG linker. By increasing the GlcN6P concentration even more (10 mM), 98% cleavage was obtained in 2 h for both the GU and GC linkers (Figure 5). Complete deletion of the linker resulted in very poor cleavage under these conditions (11%, not shown). In summary, a minimal linker of at least one unpaired nucleotide is compatible with efficient *glmS* self-cleavage (23), and the affinity purification could be adapted to allow for any of the four standard nucleotides at the 3′-end of the RNA of interest.

Affinity batch purification versus standard gel purification

A large-scale affinity purification was also performed to compare our novel procedure with a standard purification by denaturing-gel electrophoresis. Gel purification from a 5-ml transcription reaction was completed in ~6 days (excluding buffer exchange) and produced highly pure RSA_{U65C} (≥99%) at a yield of 0.58 mg/ml of transcription. To obtain an equivalent product by affinity purification with the ARiBo1-fused RSA_{U65C}, a 4-h alkaline phosphatase step was added between the first two washes. This step is necessary to produce RNA with a homogeneous 5′-end, which is particularly important in NMR studies of RNA (3), but is not required for several other applications. Moreover, it could be performed over a shorter time period by increasing the phosphatase concentration as generally done for small-scale applications. Affinity batch purification of RSA_{U65C} was completed in 7 h and produced highly pure RNA (99%) at a yield of 0.61 mg/ml (27 nmoles/ml) of transcription. However, the procedure can be completed in 3 h if the

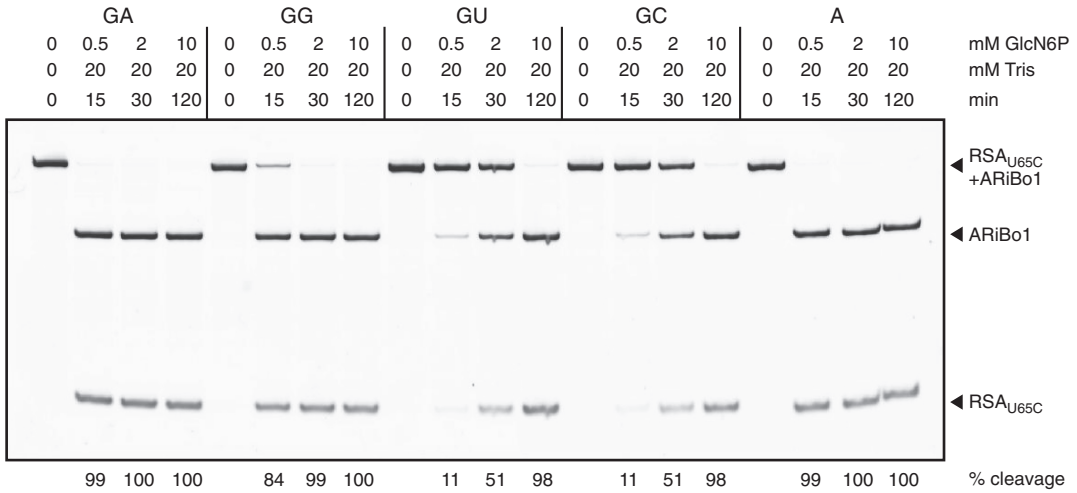

Figure 5. Sequence requirement at the 3′-end of the desired RNA for cleavage by the *glmS* ribozyme of the ARiBo1 tag. ARiBo1-fused RNAs with the original RSA_{U65C} sequence (GA linker; Figure 2A) or carrying mutations at the 3′-end (GG, GU, GC and A linkers) were cleaved under different conditions. All cleavage reactions were performed at 37°C in solution containing a 1:50 dilution of the transcription reaction, 10 mM $MgCl_2$, 20 mM Tris buffer pH = 7.6, but contained different concentrations of GlcN6P and were incubated for different amounts of time (in min), as indicated above each lane. The mobility of the RNA precursor and products is marked with arrows on the right side of the gel. The percentage of cleavage for each condition is given below the gel.

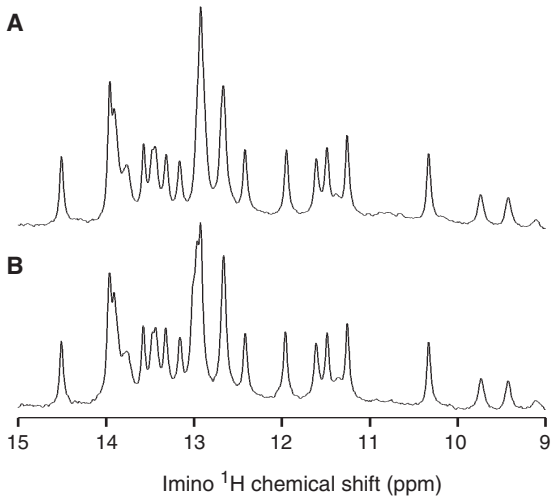

Figure 6. Imino regions of the 1D ^{1}H NMR spectra of 0.35 mM RSA_{U65C} purified by (**A**) affinity batch purification and (**B**) a standard purification protocol based on denaturing gel electrophoresis.

optional phosphatase step is omitted. The 1D imino ^{1}H NMR spectrum of the affinity-purified RNA is essentially identical to that of the gel-purified RNA (Figure 6), both being compatible with the compact three-dimensional structure of RSA_{U65C} (21). Thus, our affinity purification method produces highly pure native RNA with a yield comparable to a standard denaturing gel method, but in a significantly shorter period of time. Affinity purification could be easily scaled up to a 25–50 ml transcription reaction and completed within a working day by a single individual, whereas this would require 2–3 weeks using denaturing gels. Standard gel purification of milligram quantities of RNA is fairly time consuming because it generally involves several preparative gels, and the number of gels is dependent on the amount of RNA to be purified. It also involves additional steps to recover the RNA from the gel (e.g. electroelution or crush and soak), and to remove acrylamide contaminants [e.g. ion-exchange chromatography; (2–4,6)].

CONCLUSION

In this work, we established an affinity batch purification procedure for RNA based on the high-affinity λ*boxB*/λN peptide and GSH/GST-Sepharose interactions for immobilization and on the activable *glmS* ribozyme for elution. As other protocols described thus far for affinity purification of RNA, purification can be achieved within a few hours and allows for production of RNAs with homogenous 3′-ends (13–17). An important aspect of our procedure is that it reliably provides highly pure RNA with very good yields. Although purity levels >90% were previously reported for an affinity purification protocol where the affinity tag contains a stem-loop from the signal recognition particle and an imidazole-activated HDV ribozyme (13,16), yields of ribozyme cleavage from the support were ≤83% and absolute yields of purification were variable (between 0.5 and 24 nmoles/ml transcription). In addition, several technical problems were reported with this procedure, including the instability of the affinity protein and the high concentration of imidazole required for RNA elution, which leads to RNA degradation (14). For improved procedures that rely on the *glmS* ribozyme for RNA elution (14,15,17), purity levels were reported in only one case, where the affinity tag was designed to bind a biotinylated single-stranded DNA for immobilization on streptavidin-coated magnetic beads (17). Although purity levels >95% were obtained, absolute yields of RNA were rather low [~0.5 nmol/ml transcription; (17)]. Here, we also demonstrated that our procedure is compatible with alkaline phosphatase treatment for production of homogeneous 5′-OH termini. The chemically homogeneous RNA samples obtained with our

novel non-denaturing procedure are extremely valuable in obtaining diffracting crystals for X-ray structure determination and in providing non ambiguous data in biophysical studies, such as NMR, isothermal titration calorimetry and single molecule fluorescence energy transfer studies. In addition, we showed that all four standard bases at the 3′-end of the target RNA are compatible with *glmS* ribozyme cleavage, and that a minimal linker of one unpaired nucleotide is needed between the target RNA and the ARiBo tag. This procedure is very reliable and has been used successfully in our laboratory for the preparation of several RNAs other than RSA_{U65C}, including small RNA hairpins, the *Neurospora* VS ribozyme and several forms of pre-microRNAs.

The incentive for developing an affinity purification method in our laboratory was to simplify and reduce the time for purification of large quantities of RNA, which is generally the rate-limiting step of biophysical and structural studies. However, we have found this procedure also extremely useful for small-scale preparation of RNA, where multiple samples can be quickly purified in tandem. This novel procedure should be adaptable to a wide range of applications that require RNA purification and/or immobilization, including large-scale and small-scale purifications of *in vitro* transcribed RNA, isolation of RNA-associated complexes from living cells and high-throughput applications. For example, high-throughput purification of RNA would be valuable for screening RNAs for crystallization in structural genomics approaches and for screening ligand-binding RNAs in drug discovery efforts.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Joyce Li for suggesting several years ago that alternative methods should be developed for purification of RNA. They also thank Stephen Michnick for critical reading of the article.

FUNDING

Canadian Institutes for Health Research (CIHR to P.L.) (MOP-64341, MOP-86502). P.L. holds a Canada Research Chair in Structural Biology and Engineering of RNA. Funding for open access charge: CIHR.

Conflict of interest statement. P.L. is an inventor on a provisional patent related to this RNA affinity purification procedure.

REFERENCES

1. Milligan,J.F., Groebe,D.R., Witherell,G.W. and Uhlenbeck,O.C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.*, **15**, 8783–8798.
2. Wyatt,J.R., Chastain,M. and Puglisi,J.D. (1991) Synthesis and purification of large amounts of RNA oligonucleotides. *Biotechniques*, **11**, 764–769.
3. Legault,P. (1995) Ph.D. Structural studies of ribozymes by heteronuclear NMR spectroscopy. Thesis. University of Colorado at Boulder Boulder.
4. Doudna,J.A. (1997) Preparation of homogeneous ribozyme RNA for crystallization. *Methods Mol. Biol.*, **74**, 365–370.
5. Uhlenbeck,O.C. (1995) Keeping RNA happy. *RNA*, **1**, 4–6.
6. Lukavsky,P.J. and Puglisi,J.D. (2004) Large-scale preparation and purification of polyacrylamide-free RNA oligonucleotides. *RNA*, **10**, 889–893.
7. Shields,T.P., Mollova,E., Ste Marie,L., Hansen,M.R. and Pardi,A. (1999) High-performance liquid chromatography purification of homogenous-length RNA produced by trans cleavage with a hammerhead ribozyme. *RNA*, **5**, 1259–1267.
8. Kim,I., McKenna,S.A., Puglisi,E.V. and Puglisi,J.D. (2007) Rapid purification of RNAs using fast performance liquid chromatography (FPLC). *RNA*, **13**, 289–294.
9. McKenna,S.A., Kim,I., Puglisi,E.V., Lindhout,D.A., Aitken,C.E., Marshall,R.A. and Puglisi,J.D. (2007) Purification and characterization of transcribed RNAs using gel filtration chromatography. *Nature Protocols*, **2**, 3270–3277.
10. Keel,A.Y., Easton,L.E., Lukavsky,P.J. and Kieft,J.S. (2009) *Methods in Enzymology, Vol. 469: Biophysical, Chemical, and Functional Probes of Rna Structure, Interactions and Folding, Pt B*, Vol. 469. Academic Press, London, UK, pp. 3–25.
11. Easton,L.E., Shibata,Y. and Lukavsky,P.J. (2010) Rapid, nondenaturing RNA purification using weak anion-exchange fast performance liquid chromatography. *RNA*, **16**, 647–653.
12. Cheong,H.K., Hwang,E., Lee,C., Choi,B.S. and Cheong,C. (2004) Rapid preparation of RNA samples for NMR spectroscopy and X-ray crystallography. *Nucleic Acids Res.*, **32**, e84.
13. Kieft,J.S. and Batey,R.T. (2004) A general method for rapid and nondenaturing purification of RNAs. *RNA*, **10**, 988–995.
14. Batey,R.T. and Kieft,J.S. (2007) Improved native affinity purification of RNA. *RNA*, **13**, 1384–1389.
15. Boese,B.J., Corbino,K. and Breaker,R.R. (2008) In vitro selection and characterization of cellulose-binding RNA aptamers using isothermal amplification. *Nucleos. Nucleot. Nuc.*, **27**, 949–966.
16. Vicens,Q., Gooding,A.R., Duarte,L.F. and Batey,R.T. (2009) Preparation of group I introns for biochemical studies and crystallization assays by native affinity purification. *Plos One*, **4**, 1–14.
17. Pereira,M.J., Behera,V. and Walter,N.G. (2010) Nondenaturing purification of co-transcriptionally folded RNA avoids common folding heterogeneity. *Plos One*, **5**, 1–7.
18. Mogridge,J., Legault,P., Li,J., Van Oene,M.D., Kay,L.E. and Greenblatt,J. (1998) Independent ligand-induced folding of the RNA-binding domain and two functionally distinct antitermination regions in the phage lambda N protein. *Mol. Cell*, **1**, 265–275.
19. Austin,R.J., Xia,T., Ren,J., Takahashi,T.T. and Roberts,R.W. (2002) Designed arginine-rich RNA-binding peptides with picomolar affinity. *J. Am. Chem. Soc.*, **124**, 10966–10967.
20. Rastogi,T. and Collins,R.A. (1998) Smaller, faster ribozymes reveal the catalytic core of *Neurospora* VS RNA. *J. Mol. Biol.*, **277**, 215–224.
21. Delfosse,V., Bouchard,P., Bonneau,E., Dagenais,P., Lemay,J.F., Lafontaine,D.A. and Legault,P. (2010) Riboswitch structure: an internal residue mimicking the purine ligand. *Nucleic Acids Res.*, **38**, 2057–2068.
22. Legault,P., Li,J., Mogridge,J., Kay,L.E. and Greenblatt,J. (1998) NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell*, **93**, 289–299.
23. Winkler,W.C., Nahvi,A., Roth,A., Collins,J.A. and Breaker,R.R. (2004) Control of gene expression by a natural metabolite-responsive ribozyme. *Nature*, **428**, 281–286.
24. Wilkinson,S.R. and Been,M.D. (2005) A pseudoknot in the 3′ non-core region of the glmS ribozyme enhances self-cleavage activity. *RNA*, **11**, 1788–1794.

25. Cochrane,J.C., Lipchock,S.V. and Strobel,S.A. (2007) Structural investigation of the GlmS ribozyme bound to its catalytic cofactor. *Chem. Biol.*, **14**, 97–105.
26. McCarthy,T.J., Plog,M.A., Floy,S.A., Jansen,J.A., Soukup,J.K. and Soukup,G.A. (2005) Ligand requirements for glmS ribozyme self-cleavage. *Chem. Biol.*, **12**, 1221–1226.
27. Roth,A., Nahvi,A., Lee,M., Jona,I. and Breaker,R.R. (2006) Characteristics of the glmS ribozyme suggest only structural roles for divalent metal ions. *RNA*, **12**, 607–619.
28. Barrick,J.E., Corbino,K.A., Winkler,W.C., Nahvi,A., Mandal,M., Collins,J., Lee,M., Roth,A., Sudarsan,N., Jona,I. *et al.* (2004) New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc. Natl Acad. Sci. USA*, **101**, 6421–6426.
29. Nikonowicz,E.P., Sirr,A., Legault,P., Jucker,F.M., Baer,L.M. and Pardi,A. (1992) Preparation of ^{13}C and ^{15}N labelled RNAs for heteronuclear multidimensional NMR studies. *Nucleic Acids Res.*, **20**, 4507–4513.
30. Batey,R.T., Inada,M., Kujawinski,E., Puglisi,J.D. and Williamson,J.R. (1992) Preparation of isotopically labeled ribonucleotides for multidimensional NMR spectroscopy of RNA. *Nucleic Acids Res.*, **20**, 4515–4523.
31. Padilla,R. and Sousa,R. (1999) Efficient synthesis of nucleic acids heavily modified with non-canonical ribose 2′-groups using a mutant T7 RNA polymerase (RNAP). *Nucleic Acids Res.*, **27**, 1561–1563.