CHAPTER THREE

METHODS OF SITE-SPECIFIC LABELING OF RNA WITH FLUORESCENT DYES

Sergey Solomatin* *and* Daniel Herschlag†

Contents

1. Introduction 48
2. Design of Labeled RNA Constructs 49
   2.1. Selection of labeling sites and construct assembly methods 49
   2.2. Design of the construct assembly 49
   2.3. Selection of dyes for single molecule fluorescence studies 51
3. Dye Labeling of RNA Fragments 53
4. Notes on *In Vitro* Transcription with T7 RNA Polymerase 55
5. Assembly of Labeled RNA Constructs 56
6. Examples of Protocols 58
   6.1. Protocol 1: Labeling of RNA oligos by with fluorescent dyes (NHS ester form) 58
   6.2. Protocol 2: Ligation of large RNAs with T4 DNA ligase 61
Acknowledgments 66
References 66

Abstract

Single molecule fluorescence techniques offer unique insights into mechanisms of conformational changes of RNA. Knowing how to make fluorescently labeled RNA molecules and understanding potential limitations of different labeling strategies is essential for successful implementation of single molecule fluorescence techniques. This chapter offers a step by step overview of the process of obtaining RNA constructs ready for single molecule measurements. Several alternative methods are described for each step, and ways of troubleshooting the most common problems, in particular, splinted RNA ligation, are suggested.

* Department of Biochemistry, Stanford University, Stanford, California, USA
† Departments of Biochemistry and Chemistry, Stanford University, Stanford, California, USA

Methods in Enzymology, Volume 469
ISSN 0076-6879, DOI: 10.1016/S0076-6879(09)69003-0
© 2009 Elsevier Inc.
All rights reserved.

Vilber EX1073

1. Introduction

Over the past two decades, an ever increasing appreciation of multiple roles RNAs play in biology has led to an increasing interest in understanding the fundamental behavior of RNA. A broad variety of experimental approaches has been applied to studying structure and dynamics of RNA, including native gel electrophoresis, NMR, small angle X-ray scattering (SAXS), chemical structure probing, atomic force microscopy, UV, and fluorescence spectroscopy.

Single molecule methods have become the latest frontier in studies of RNA dynamics (Weiss, 2000; Zhuang, 2005), revealing unique information about the behavior of RNA molecules hidden from bulk experiments by ensemble averaging (Downey *et al* , 2006; Ha *et al* , 1999; Hodak *et al* , 2005; Lee *et al* , 2007; Qu *et al* , 2008; Russell *et al* , 2002; Xie *et al* , 2004; Zhuang *et al* , 2000, 2002). Fluorescence-based techniques have two characteristics that make them particularly useful in single molecule implementation. First, they allow increased throughput of single molecule measurements through simultaneous observation of hundreds of individual molecules at the same time. Analysis of data from thousands of individual molecules is essential for bridging the gap between observations made on individual molecules and traditional bulk measurements made with ensembles of $\sim 10^{20}$ molecules. Second, an ability to label different parts of RNA molecules allows one to study the dynamics at a submolecular level and obtain increasingly detailed information about the RNA structure and motions.

To take full advantage of these traits, one needs to incorporate fluorescent dyes site-specifically into an RNA of interest and to be able to do this at different selected position on an RNA molecule. This chapter outlines general strategies of preparing dye-labeled RNA constructs, in particular concentrating on double-labeled constructs for single molecule fluorescence resonance energy transfer (smFRET) measurements (Ha, 2001; Roy *et al* , 2008). These constructs are designed such that the dynamics of interest is revealed through changes of the distance between two dye labels, donor and acceptor, which in turn result in changes of the energy transfer efficiency from the donor to the acceptor. Anticorrelated changes of the donor and acceptor fluorescence, resulting from changes in energy transfer efficiency, are easy to distinguish from uncorrelated fluctuations of the fluorescence intensity arising from multiple possible sources. Because of this trait, FRET is the technique that has been most widely used for studying RNA dynamics (Weiss, 2000; Zhuang, 2005).

A step by step overview of the process of obtaining RNA constructs ready for single molecule FRET is presented here. For each step, several alternative methods described in the literature are suggested. Means of

troubleshooting of the most common problems, in particular, splinted RNA ligation, are suggested based on the literature and on the authors' personal experience.

2. Design of Labeled RNA Constructs

2.1. Selection of labeling sites and construct assembly methods

Design of a labeled RNA construct is aimed at achieving two goals:

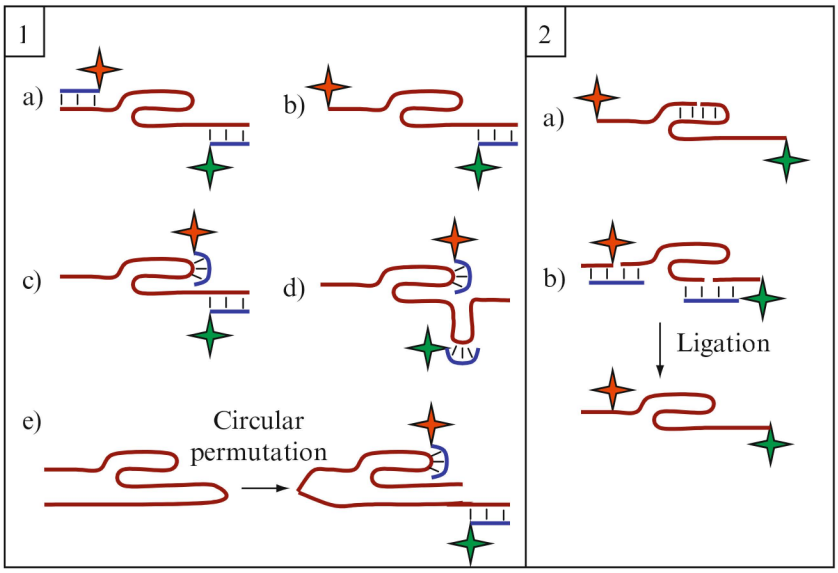
1. Obtaining the best possible FRET signal.
2. Minimally perturbing the behavior of the RNA.

Observation of a FRET signal requires two dyes to be positioned on the RNA within the range of efficient energy transfer determined by the Förster distance for a particular dye pair (typically, 3–6 nm). A crystal structure is a great starting point for identifying appropriate labeling positions. A simple heuristic rule is to choose labeling sites that are remote in the secondary structure, but close in the tertiary structure. However, one should be careful not to perturb residues that might be involved in forming long-range tertiary interactions (e.g., Brion and Westhof, 1997, and references therein). Such residues are likely to conform to the rule above, but modifying them can severely destabilize the native structure of the RNA. On the other hand, labeling base paired regions adjacent to residues that make long-range tertiary contacts is safer, and it ensures that FRET will be observed. If a crystal structure is not available, phylogenetic (Jaeger *et al* , 1994; Michel and Westhof, 1990), cross-linking (Chen *et al* , 1998), and biochemical (Lehnert *et al* , 1996) data can help guide the search for best labeling sites.

2.2. Design of the construct assembly

Labeling with two different dyes typically requires the final RNA construct to be assembled from at least two fragments (Scheme 3.1). Depending on the choice of labeling sites, one of the following assembly strategies can be pursued:

1. RNA is made (synthetically or by *in vitro* transcription) in one piece, and labeled by base pairing to complementary-labeled oligos:
   - **(a)** Two oligos are bound at the 5′ and 3′ ends
   - **(b)** One of RNA ends is labeled directly, and an oligo is bound at the other end
   - **(c)** One oligo is bound at an end and the other at an internal site



Scheme 3.1 Design strategies for making labeled RNA constructs. Numeration corresponds to the outline numbering in the text. Brown lines designate RNA, blue lines designate DNA oligos, stars designate fluorescent dyes.

(d) Two oligos are bound at internal sites.
(e) New 5′ and 3′ ends are designed by circular permutations, and an oligo is bound at an internal site.

2. RNA is split into several fragments:

(a) RNA is reassembled by base pairing of the fragments.
(b) RNA is reassembled by covalent joining of the fragments.

Assembly method 2.b (Lee *et al* , 2007; Sattin *et al* , 2008) reproduces exactly the same RNA as of the original unlabeled construct and can be considered the least perturbing method of labeling, as long as the positions of the dyes were appropriately chosen. However, it also remains the most technically challenging method, and it imposes the strictest requirements on the purity of RNA fragments, as discussed below.

Other methods of assembly (1.a–e and 2.a) are easier to implement, but they restrict the choice of labeling sites to the ends of the molecule (1.a–c and 2.a), and/or require assumptions that modifications of the sequence—that is, binding of oligos at the ends and at internal sites (1.b–d), circular permutations (1.e) or nicks in the continuous backbone (2.a)—do not affect the dynamics of interest. Testing such assumptions can be nontrivial.

The modular architecture of RNA structure lends some support to the assumption that adding duplexes at the ends of the molecule will not generally perturb its behavior (Brion and Westhof, 1997; Tinoco and Bustamante, 1999). Single molecule studies demonstrated that the overall folding rate, substrate docking rate and catalysis by the Tetrahymena group I

ribozyme were the same for the construct labeled by and oligo annealed to a 3′ end extension as for unmodified construct (Russell *et al* , 2002; Zhuang *et al* , 2000, 2002). This method of labeling has been most widely used, but it has an obvious limitation of placing dyes in the vicinity of the ends of the molecule.

In principle, the ends of an RNA molecule can be moved by making circular permutations (Pan, 2000). As it is known that circular permutations can strongly affect folding mechanisms of RNA (Lease *et al* , 2007; Pan *et al* , 1999), this approach may be mostly useful for studying structure and local dynamics of RNA. Also, after a circular permutation the ends of the molecule are expected to be right next to each other, so that the second dye has to be placed at some internal position.

To place a dye at internal positions without breaking the RNA backbone, the Pan lab developed a method that involves replacing nonessential hairpin loops within RNA sequences with larger loops with specific sequences that are hybridized to labeled DNA oligos (Smith *et al* , 2005). These modifications had little effect on structure, as assayed by chemical footprinting, or catalytic efficiency of the catalytic domain of RNase P. The same method was successfully employed for studying ribosome dynamics (Dorywalska *et al* , 2005).

2.3. Selection of dyes for single molecule fluorescence studies

Single molecule FRET experiments push the limits of sensitivity and time resolution of the detection systems, because the goal of these experiments is to get as much information as possible from the weakest possible light source. Good photophysical properties of dyes are essential for getting the most out of these experiments. The following properties are highly desirable for smFRET dyes: (1) high extinction coefficient and quantum yield (i.e., most of the excitation light is converted into useful signal); (2) high stability against photobleaching (i.e., each molecule can be observed for a long time); (3) stable fluorescent signal (i.e., no chemical or conformational transformations of the dye leading to large fluctuations of fluorescence such as blinking); (4) good spectral overlap of donor emission and acceptor excitation (allowing high maximum FRET efficiency); (5) good spectral separation of donor and acceptor emissions (i.e., easy to optically separate two signals and calculate the actual value of FRET); (6) donor and acceptor emission in the range of high quantum efficiency of the detection systems (e.g., for some CCDs quantum efficiency falls off sharply outside of the ~450–850 nm window).

Currently, one can choose from a broad variety of organic fluorophores covering the entire optical spectrum from UV to near IR that are commercially available (see Table 3.1). A lot of early smFRET work was performed

Table 3.1 Spectral properties of several selected fluorophores for RNA labeling

Dye	λ_{ex} (nm)	λ_{em} (nm)	ε (M^{-1} cm^{-1})	Φ
Cyanine fluorphores				
Cy2	489	506	150,000	0.12
Cy3	548	562	150,000	0.16–0.39
Cy5	649	670	250,000	0.28
Cy5.5	675	694	190,000	0.23
Alexa fluorophores				
Alexa 350	346	445	19,000	–
Alexa 430	430	545	15,000	–
Alexa 488	494	517	73,000	0.92
Alexa 532	530	555	81,000	0.61
Alexa 555	555	572	155,000	0.1
Alexa 647	651	672	270,000	0.33
Alexa 700	702	723	205,000	0.25
ATTC fluorophores				
ATTO 425	436	484	45,000	0.90
ATTO 532	532	553	115,000	0.90
ATTO 647	645	669	120,000	0.20
ATTO 700	700	719	120,000	0.25
Cther fluorophores				
Fluorescein (0.1 *M* NaOH)	495	519	79,000	0.79–0.95
Tetramethyl rhodamine	557	576	103,000	0.2
Texas Red	589	615	139,000	0.9

with the cyanine dyes Cy3 and Cy5, and, while several lines of dyes were marketed recently as superior to cyanine dyes, the Cy dyes are still widely used because they have a highly desirable combination of properties. The fluorescence signal of Cy dyes is strong, long-lived and stable in oxygen-depleted environments in the presence of stabilizing agents,[1] and they have well-separated emission spectra in the range ideal for most detection systems.

Manufacturers of Alexa (Invitrogen), Dylight (Thermo Fisher Scientific), and ATTO (ATTO-Tec, also available from Sigma-Aldrich) dyes offer a broad choice of fluorophores covering the entire visible spectrum in small increments. Red dyes from these lines are reported to be significantly

[1] Trolox (Rasnik *et al.*, 2006) is an exceptionally good one, but other compounds, such as β-mercaptoethanol, *n*-propyl gallate, ascorbic acid, or chloramphenicol, have been used (Widengren *et al.*, 2007).

more photostable than Cy5.[2] Furthermore, ATTO dyes display exceptionally high brightness and have low intersystem crossing rates, which should favorably reflect on the stability of their fluorescence signal.

Furthermore, some of the older, nonbranded, dyes, for example, Texas Red (Ha *et al*., 1996), fluorescein (Xie *et al*., 2004), or tetramethyl rhodamine (Lang *et al*., 2004), can work perfectly well in single molecule applications, and these dyes typically cost less than those noted above.

All of the above dyes can be purchased in *N*-succinimide activated form and used for labeling of amino groups (see below). The choice of fluorophores are more limited for labeling of thiol groups with maleimide derivatives of dyes, and especially for full synthesis of oligos using dye-labeled phosphoramidites.

3. Dye Labeling of RNA Fragments

Any strategy of assembly of a labeled RNA construct for smFRET requires obtaining two dye-labeled RNA fragments, typically two oligos. Dye-labeled RNA fragments can be obtained in one of the following ways:

1. Commercially labeled oligos.
2. In-house labeling of commercially synthesized oligos with amino- or thiol groups.
3. Full in-house oligonucleotide synthesis with 5′ end incorporation of dye phosphoramidites.
4. Direct RNA labeling.

One should keep in mind that not every labeling method is compatible with assembly strategies that require ligation of the oligos, for example enzymatic ligation with T4 DNA ligase would require 5′ monophosphate group and a free hydroxyl group at the 3′ of each junction to be ligated, whereas RNAs labeled by dye phosphoramidite incorporation or direct labeling at the 5′ and 3′ ends will have different end groups most likely incompatible with the ligation.

1. *Commercially labeled oligos*. Several commercial oligos manufacturers (e.g., International DNA Technology, or IDT, www.idtdna.com, Dharmacon, www.dharmacon.com, Gene Link, www.genelink.com) provide an option of purchasing a custom RNA oligos with a variety of dyes incorporated at the 5′ end, 3′ end or internally (currently only at uracils). While undeniably convenient, this strategy often is not the most

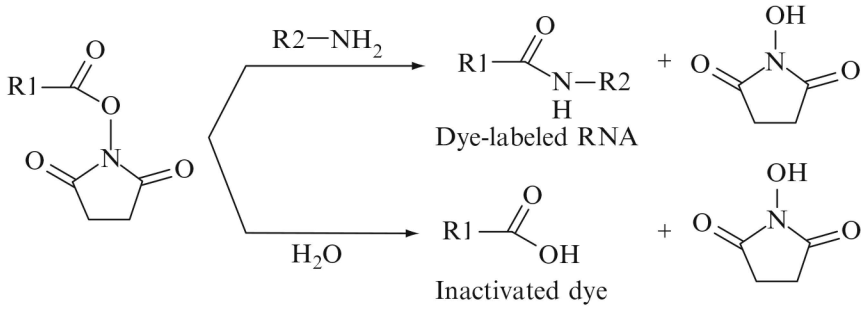
[2] These comparisons (by the manufacturers) were probably made in oxygen-rich environments, as the differences appear to be small in oxygen-depleted solutions in the presence of Trolox (unpublished observations).

cost-effective because of higher prices and low yields of supplied oligos. Furthermore, the choice of dyes is limited to those few that available from the RNA manufacturer, and currently the selection is narrow for dyes to be incorporated at 3′ and internal positions.

2. *Labeling of amino- or thiol-modified oligos.* In-house labeling of commercially synthesized oligos carrying reactive groups at specified positions combines the convenience of not having to synthesize the full oligo sequence with much broader flexibility of choosing one's favorite dyes and placing them at any position along the oligo sequence. These factors, along with the lower cost of synthesis, makes in house labeling strategy our current favorite.

Most fluorescent dyes are available as *N*-hydroxysuccinimide (NHS) or maleimide derivatives, ready for conjugation to primary amino- or thiol groups, respectively. As natural RNAs lack strongly nucleophilic aliphatic primary amino groups, the use of NHS-derivatives is typically easier and more straightforward than thiol modification by maleimides, which may require working in reducing or oxygen-free environments. Amino groups can be incorporated into custom RNA oligos during synthesis at either of the ends (through an aliphatic linker), internally at the uridine base as 5-aminoallyl uridine or at the backbone (Uni-Link^{TM} by IDT).[3]

The labeling reaction (Scheme 3.2) is easy to perform (see below), as long as certain precautions are taken. It is necessary to make sure that amines are not present in the reaction mix (e.g., TRIS-based buffers are incompatible with NHS labeling) and that the pH is optimal for labeling. At low pH most of the amino groups are protonated and their reactivity is low, but at high pH NHS esters can be hydrolyzed faster than they can react with the oligo, resulting in low yields. Carbonate or phosphate-based buffers at pH 8–9 typically work well.



Scheme 3.2 The labeling reaction of amino-modified RNA (R2) with NHS-activated dye (R1) and a competing side reaction of NHS hydrolysis.

[3] 2′-Amino ribose can also be used, but amino group in this position is much less reactive than a typical aliphatic amine.

Purification of the labeled oligo from excess dye can be accomplished by a combination of ethanol precipitation and PAGE or HPLC purification. Changes in the oligo mobility resulting from dye labeling are usually sufficient to purify the labeled oligo from the unlabeled material. Accomplishing this purification is beneficial for single molecule experiments, as it decreases the number of molecules that are not labeled, or labeled with a single dye only.

3. *Full in-house oligonucleotide synthesis with 5′ end incorporation of dye phosphoramidites.* Due to wide availability of commercially synthesized oligonucleotides, full synthesis of RNA oligos in individual labs is not frequently carried out these days and choices of commercially available dye phosphoramidites are limited. Thermo Fisher Scientific (www.thermo.com) provides a selection of DyLight phosphoramidites that have spectral properties analogous to Cy3, Cy5, and Cy5.5 dyes.
4. *Direct RNA labeling.* Several methods of labeling the directly labeling transcribed RNA molecules been reported in the literature, but have not yet found wide applications. Dyes at the 5′ were incorporated cotranscriptionally by initiating with a dye–guanosine conjugate (Fang *et al* , 1999), at the 3′ end through oxidation of the terminal ribose to aldehyde form by sodium periodate with subsequent reaction with hydrazine derivatives of dyes (Proudnikov and Mirzabekov, 1996), and at the 2′ hydroxyl in the middle of an RNA by deoxyribozyme-catalyzed ligation (Baum and Scott, 2007).

4. Notes on *In Vitro* Transcription with T7 RNA Polymerase

RNA fragments longer than 40–50 nucleotides are most often synthesized by *in vitro* transcription with T7 RNA polymerase. Heterogeneity of transcripts obtained by this method must be recognized by researchers aiming to use such transcripts for ligations, as it can lead to serious artifacts that are especially notable in single molecule experiments.

T7 RNA polymerase strongly prefers guanosine at the first and second transcribed positions, otherwise the transcription yields drastically decrease (Milligan and Uhlenbeck, 1989). However, if T7 polymerase encounters four or more guanosines in a row at the start site, it generates heterogeneity at the 5′ end (Pleiss *et al* , 1998). A proper choice of the starting sequence appears to be sufficient to avoid this problem. However, heterogeneity at the 3′ end is the rule rather than the exception for any RNA sequence transcribed by T7 polymerase.

Run-off *in vitro* transcripts almost invariably contain a significant fraction of nontemplated nucleotides at the 3′ end, with $n + 1$, $n + 2$, and $n + 3$

transcripts being the main contaminants (Milligan and Uhlenbeck, 1989). For transcripts that are larger than ~100 nt this contamination is not easy to recognize and essentially impossible to purify away. To obtain clean 3′ ends, it is best to extend the transcribed sequence beyond the intended end and then cleave the RNA at the desired site.

RNA can be extended with (a) a sequence that encodes one of small ribozymes (Ferre-D'Amare and Doudna, 1996; Price *et al*., 1995); (b) a recognition sequence for a DNAzyme (Santoro and Joyce, 1997); or (c) with a sequence complementary to a DNA oligo with subsequent cleavage of the hybrid by RNase H (Stone *et al*., 2007, see Akiyama and Stone, Chapter 2, this volume).

Extending the 3′ end sequence with a *cis*-cleaving hammerhead ribozyme is an easy and efficient way to obtain "clean" 3′ ends, as this gives cotranscriptional cleavage. However, it does place certain constraints on the sequence at cleavage site (Birikh *et al*., 1997). If the desired 3′ end sequence is not compatible with the hammerhead cleavage, another small ribozyme or other methods mentioned above can be used.

Small ribozymes and DNAzymes leave 2′–3′ cyclic phosphate at the site of cleavage, and this group must be removed before the ligation. This task can be accomplished by treating cleaved products with polynucleotide kinase (PNK) in the absence of ATP (Schurer *et al*., 2002).

5. Assembly of Labeled RNA Constructs

1. *Noncovalent assembly by Watson–Creek base pairing.* Assembly of larger RNA constructs through base pairing of complementary oligos, or oligos and *in vitro* transcripts, is the most widely used method in single molecule fluorescence field (Ha *et al*., 1999; Hodak *et al*., 2005; Xie *et al*., 2004; Zhuang *et al*., 2002). Its biggest advantage is the ease of the procedure, which usually involves simple mixing of the solutions of oligonucleotides and annealing via some combination of heating and cooling steps. Upon annealing, the efficiency of construct assembly can be tested using nondenaturing acrylamide gels. Purification of the fully assembled constructs can also be done using nondenaturing PAGE.
2. *Covalent incorporation of labeled oligos.* Full-length-labeled RNA molecules of essentially any size can be obtained by joining together labeled RNA oligos (typically obtained by synthesis) and either synthetic or *in vitro* transcribed RNA fragments comprising the rest of the sequence. Joining several RNA fragments into a single chain can be done using (i) protein ligases, (ii) deoxyribozyme ligase, and (iii) chemical ligations.

   (a) *Enzymatic ligation* using T4 DNA ligase (Moore and Sharp, 1992), remains the most often used process, offering significant advantages

over RNA ligase I as discussed by Moore and Query (2000). Recently discovered RNA ligase II (Ho and Shuman, 2002) does not suffer from many of the limitations of RNA ligase I and may become an efficient alternative to T4 DNA ligase.

Heterogeneity of fragments that need to be ligated is one of the biggest obstacles for enzymatic ligations. Heterogeneity of *in vitro* transcripts was discussed above, and it must be avoided. Synthetic oligos are always contaminated by shorter products ($n - 1$, $n - 2$, etc.) because of <100% coupling efficiency in synthesis, and these contaminants must also be purified away. This purification is typically accomplished together with purification of labeled from unlabeled material, but it is best to check the purified material on a denaturing gel. Templated nature of the reactions with DNA ligase and RNA ligase II ensures that these products would not be ligated efficiently (Moore and Query, 2000). However, T4 DNA ligase does not perfectly exclude nontemplated nucleotides at the end from the ligation (K. Travers, W. Zhao, D. Herschlag, unpublished data), which results in incorporation of extra bases in the middle of the sequence—a very undesirable outcome. The exact specificity of T4 RNA ligase II toward templated nucleotides is not known to the authors. Even if they are not ligated, the presence of shorter products will limit the annealing efficiency, increasing the amount of enzyme consumed in the reaction and decreasing the yields.

Even with purified fragments, ligation efficiency is often low, especially when multiple fragments need to be ligated. In many cases, the yield of the ligation is limited by the extent of formation of correctly annealed complexes between all RNA fragments and DNA splints. The most likely explanation for the poor annealing efficiency is intramolecular structure formation by the RNA fragments. Long splints (Kurschat *et al*., 2005), ''disrupter'' oligos (Strobel and Cech, 1993), and optimization of the annealing protocols can be tried for improving annealing efficiency, but, in some cases, choosing different sites for the ligation junctions is the only method that results in yield improvement.

Purification of the full-length-labeled RNA can be performed by PAGE. Even small amount of the ligated product can be detected by fluorescent scanners, and detection of fluorescence from both dyes can be used to identify the correct band.[4]

(b) *Deoxyribozymes* (see S. Silverman, Chapter 5, this volume) that are capable of efficient ligation of RNA fragments have recently been evolved in the Silverman laboratory (Purtha *et al*., 2005). These

[4] As a rule rather than an exception, multiple products of incomplete ligations will be observed on the gel. It is therefore not recommended to have either of the terminal fragments very short and unlabeled, as incomplete products missing these fragments would be very difficult to distinguish and separate from the full-length RNA, especially for longer molecules.

deoxyribozymes catalyze the formation of the natural 3′–5′ linkages and have reasonably broad sequence requirements at the site of cleavage, so they have a large potential as alternatives to protein ligases.

(c) *Chemical ligations* employ reagents such as carbodiimide, or bromocyane (BrCN, also known as cyanogen bromide) (Dolinnaya *et al*., 1988; Fedorova *et al*., 1996). These ligations typically work better with DNA than with RNA substrates and are prone to formation of nonnatural linkages.

6. Examples of Protocols

6.1. Protocol 1: Labeling of RNA oligos by with fluorescent dyes (NHS ester form)

6.1.1. Protocol outline

1. Prepare reagents
   - (a) Dry dimethyl sulfoxide (DMSO)
   - (b) Buffer exchange (and concentrate) RNA oligos
   - (c) Pour the purification gel
2. Run the labeling reaction
3. Purify products
   - (a) Precipitation
   - (b) PAGE purification[5]
   - (c) Second precipitation and/or desalting step
   - (d) Analysis

Reagents	Equipment
RNA oligo (5-aminoallyl uridine and 5′-monophosphate modified)[a]	Microcon columns (YM-3 or YM-10, Millipore)
Dye, NHS-derivative (powder)	Table top centrifuge
Dimethyl sulfoxide (DMSO)	Floor centrifuge
Molecular sieves, class 3A	Centrifuge tubes (e.g., Oakridge tubes fro Nalgene)
Phosphate buffer (500 m*M*, pH 8)	Gel electrophoresis box
Sodium acetate (3 *M*, pH 5.2)	Sterile scalpels (alternatively, one can use flame sterilized razor blades)
Ethanol (95% or 200 proof)	Glass rods

[5] HPLC purification can be used instead of PAGE.

Loading solution (90% formamide, trace amounts of bromphenol blue and xylene cyanol)	
Acrylamide (29:1, 20%, 7 *M* urea)	
TBE buffer (100 m*M* Tris–base, 83 m*M* boric acid, 1 m*M* EDTA), autoclaved	

[a] 5′-Monophosphate is only needed if the 5′ end of the oligo will be ligated.

6.1.2. Protocol

1. Preparation of the reagents

   1.1. *Drying DMSO* Dyes are expensive, typically, ~$300 per 1 mg, and rarely available packaged in small amounts that are convenient for test reactions or small scale labeling. Labeling of 100 nmol or less of an RNA oligo will require only a fraction of 1 mg of dye, and the excess dye solution can be stored for future use. The presence of residual water in DMSO can cause hydrolysis of the NHS ester, so it is important to keep DMSO anhydrous. This task is accomplished by drying DMSO over molecular sieves 3A.[6]

   1.2. *Buffer exchanging and concentrating the oligos* If the RNA oligos arrived from the manufacturer in lyophilized form and were deprotected before use, they may contain residual TEMED that will interfere with the labeling reaction. If the oligos are supplied in ''lab-ready'' form, the buffer might be TRIS based, which will also interfere with labeling. It is advised to desalt and buffer-exchange the oligos using Microcon columns (Millipore) or the equivalent. If the oligo concentration is less than 1 m*M*, it is also strongly advised to concentrate the solution at this step to 1 m*M* or higher. Higher concentrations of RNA result in better yields at the same dye-to-oligo ratio, as the labeling reaction is second order and can more effectively compete with hydrolysis of the active dye (see Scheme 3.2).

   1.3. *Pour a denaturing polyacrylamide gel.* Choose the percentage of the gel and its length according to the length of the oligo. For some dyes (e.g., Cy3 and Cy5), the labeled material has electrophoretic mobility approximately equal to an oligo that is longer by a single nucleotide. It is strongly suggested to run a test gel with a small aliquot of labeling reaction and an unlabeled oligo marker. The labeled and unlabeled bands can be visualized by staining using, for

[6] Molecular sieves should be activated according to manufacturer's instructions (typically by heating at 250 °C for 2 h or longer).

example, StainsAll (Sigma-Aldrich) and the percentage of acrylamide and the length of the gel can then be adjusted to obtain proper separation. For <30-mer oligos, 20% gels with ~20 cm plates provide adequate separation; 1.5 mm gels with 10–20 mm wells provide a good balance between band visibility and resolution on 10–30 nmol (oligo) scale; and 3 mm gels and/or wider wells should be used for labeling reactions on larger scales.

2. A sample labeling reaction for 100 nmol of an RNA oligo

Stock reagent	Volume	Final concentration
1 m*M* amino-modified RNA	100 μl	~500 μM
500 m*M* phosphate buffer, pH = 8	20 μl	~50 m*M*
DMSO (anhydrous)	100 μl	45%
Dye (1 mg)		~5 m*M*
Total volume	220 μl	

2.1. Combine RNA and phosphate buffer in a 2-ml Eppendorf tube.
2.2. Add DMSO to the tube of dye and pipette up and down vigorously to dissolve it.
2.3. Add the needed amount of dye solution to the RNA solution and mix well.
2.4. If there is dye remaining, it should be quickly frozen on dry ice and stored at − 80 °C to prevent hydrolysis of the NHS ester.
2.5. Incubate at 22 °C for 1 h.

3. Purification of the products

3.1. Precipitate RNA by adding 0.11 volumes of 3 *M* sodium acetate (pH 5.2), and 4 volumes of cold ethanol and incubate on dry ice for 30 min or longer.
3.2. Centrifuge at 14,000 rpm for 40 min at 4 °C, decant and save the supernatant, allow the pellet to dry under air.
3.3. Resuspend the pellet in a minimal volume of water, add 1× volume of the loading solution, load the gel and run as long as needed to achieve good separation of the labeled product from the unlabeled oligo (see 1.3).
3.4. Identify the correct band and excise it with a sterile scalpel.[7]
3.5. Place excised gel pieces into ice-chilled collection tubes and crush them thoroughly with glass rods, then quickly freeze on dry ice.

[7] Note that the excess free dye often aggregates and runs as multiple bands on the gel. To avoid misidentification of the bands, it is essential to perform the precipitation step (3.1–3.4), removing the excess dye, and to run a test gel (1.3) to establish the mobility of the labeled oligo.

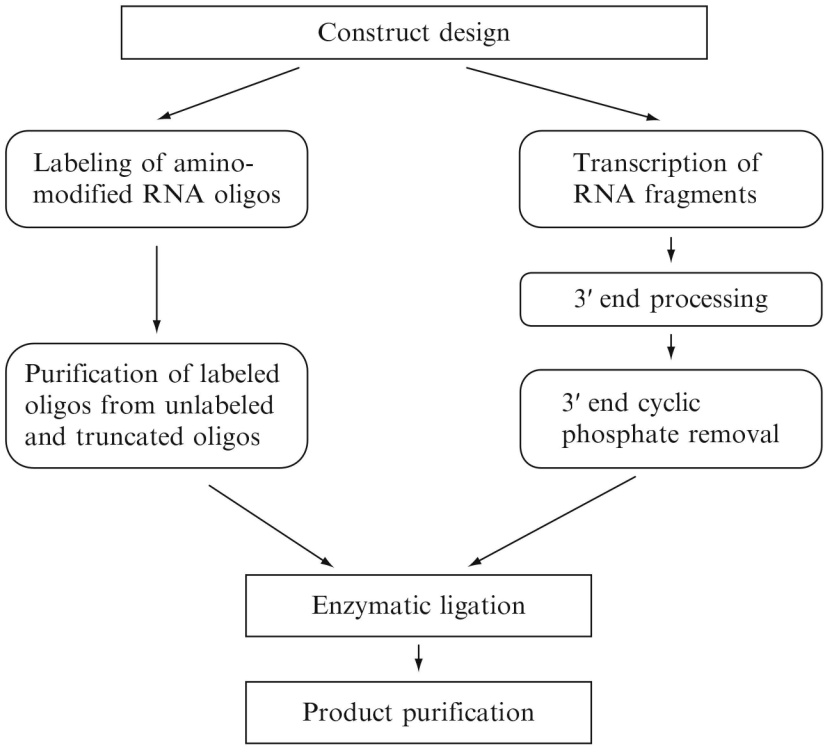
3.6. Add water (3× volume of the gel) and repeat freeze-thaw cycle three times.
3.7. Centrifuge for 10 min at 4000*g* to remove the gel fragments and take out the supernatant. Perform the second round of elution overnight on a shaker at 4 °C.
3.8. Combine the supernatants from the first and the second elution rounds, filter through 0.2 μm filters and precipitate RNA as in (3.1)–(3.4).
3.9. Resuspend in water or in a storage buffer of choice, and estimate the concentration of the dye by measuring the UV absorbance in the range of dye absorption, and of the RNA oligo by measuring the UV absorbance at 260 nm. The labeling efficiency can be calculated from the ratio of the dye concentration to the RNA concentration. Alternatively, the concentration of the dye can be estimated by measuring its fluorescence intensity and comparing it to a standard curve measured for a series of known dye concentrations.
3.10. It is recommended to test the purity of the product at the end of the purification procedure by running a denaturing PAGE gel and (1) scanning it on a fluorescent scanner (e.g., Typhoon system, Molecular Dynamics) to determine the extent of degradation (and contamination by shorter, for example, $n - 1$, $n - 2$, etc., labeled oligos), and (2) staining it with StainsAll to determine the extent of contamination with the unlabeled oligo.

6.2. Protocol 2: Ligation of large RNAs with T4 DNA ligase

Preparation of the labeled constructs by enzymatic ligation involves several steps of preparing ligation components, as shown in Scheme 3.3. This protocol describes only the last step in the overall procedure, as the preceding steps are standard in RNA biochemistry and corresponding protocols can be found elsewhere (see, e.g., Akiyama and Stone, Chapter 2, this volume).

6.2.1. Outline of the protocol

1. Anneal RNA fragments and DNA splints
2. Run the ligation reaction
   (a) Test the results of the ligation
3. Purify products by PAGE



Scheme 3.3 **Flow-chart for making labeled RNA constructs by enzymatic (T4 DNA ligase) ligation. Denaturing PAGE purifications after each step of RNA transcript preparation are performed, but now shown for clarity.**

Reagents	Equipment
TE buffer (10 m*M* Tris–HCl, pH 8, 1 m*M* EDTA)	PCR machine (Thermocycler) (or heating block) for annealing
NaCl (5 *M*)	Table top centrifuge
T4 DNA Ligase	Gel electrophoresis box
10× ligase buffer (660 m*M* Tris–HCl, pH 7.6, 80 m*M* $MgCl_2$, 100 m*M* DTT, 1 m*M* ATP, 0.04% Triton X-100)	Sterile scalpels (or flame sterilized razor blades)
RNA fragments[a] (≥10 μM is desirable)	Glass rods
DNA splint(s)	

[a] 5′-Ends that will be joined must be phosphorylated, and 3′ ends must have free hydroxyl groups.

6.2.2. Protocol

1. Annealing

Annealing efficiency is higher when all the components are in approximately stoichiometric amounts, and the concentration of each component is

higher than 1 μM ($\sim$10 μM is preferred). Excess amounts of middle fragments or splints will markedly decrease ligation efficiency because instead of a single splint joining the two ends at the junction, each end can bind a splint, so that no joining occurs. Therefore, concentrations of each component must be carefully measured. To achieve $\geq$10 μM final concentrations in the annealing mixture, solutions of RNA fragments might need to be concentrated, especially for multijunction ligations. This can be done using Microcon columns with an appropriate molecular weight cut-off (Table 3.2).

Mix the components together in a tube, place into a PCR machine and run the following program: 95 °C for 5 min; ramp down to 22 °C at −0.1 °C/min; keep at 4 °C. Alternatively, heat to 95 °C in heat block for 5 min, then let cool to room temp by leaving in turned off heat block. It will take 1.5–2 h to reach room temperature; can be varied if needed.

2. Run ligation reaction

Mix the components and incubate at room temperature overnight. Run a test gel with a small aliquote of the reaction mix (1 μl should be sufficient, as the Typhoon scanner can detect <10 fmol of dye) before stopping the ligation. Some junctions might require significantly longer incubation times, increase of the yield over as many as 72 h have been observed (unpublished data). RNA degradation may become a limiting factor over such long incubation times, so maintaining RNase-free conditions is essential. Nonenzymatic RNA degradation can be limited by lowering the pH (Table 3.3).

3. Purify products as described in the Protocol 1, Part 3.

Table 3.2 Annealing at 1 nmol scale for ligating *N* fragments with *M* splints

Stock reagent	Volume	Final concentration
RNA fragments (100 μM each)	$N \times$ 10 μl	$\sim$10 μM each
DNA splints (100 μM each)	$M \times$ 10 μl	$\sim$10 μM each
TE buffer	10 μl	0.1×
NaCl (5 *M*), optional	0–20 μl	0–1 *M*[a]
H_2O	100 − (*N* + M) × 10 − V_{NaCl}	
Total volume		100 μl

[a] The amount of NaCl added during the annealing step can be varied to achieve optimal annealing. However, T4 DNA ligase is inhibited at >200 m*M* NaCl, and consequently NaCl concentration in the ligation step should be <200 m*M*.

Table 3.3 Ligation at 1 nmol scale for *N* fragments

Stock reagent	Volume	Final concentration
Annealed complex (10 μM)	100 μl	~1 μM each
Ligase buffer	100 μl	~10 μM each
Superas⋆In (20 U/μl), optional	5 μl	0.1 U/μl
T4 DNA ligase (~200 μM)	$(N-1) \times 5$ μl	$(N-1)$ μM
H_2O	$800 - (N-1) \times 5$	
Total volume		1 ml

If the scale of the ligation is more than 10 nmol, the band of the full-length product on the gel will most likely be visible to a naked eye. If the ligation is performed on a smaller scale, the position of the ligated full-length RNA can be identified by scanning the gel on a Typhoon scanner in the fluorescent mode.

6.2.3. Troubleshooting of ligations

Typical yields of full-length ligated products after gel purification are ~40% for two-fragment ligations, ~20% for three-fragment ligations, and ~10% for four- and five-fragments ligations. If none of the full-length product is formed or the yield is too low, conditions of the annealing and ligation reaction can be optimized (see below). As multiple factors may contribute to low ligation yields, it is recommended that individual parameters affecting ligation efficiency are systematically tested. Figure 3.1 provides an example of systematic optimization of the ligation reaction that produced no product before optimization primarily due to a single junction not being ligated, most likely because of structure formation in the 3′-end fragment. Several reaction parameters ([ATP], $[Mg^{2+}]$, reaction time) had to be optimized to achieve ligation of the full-length product.

1. Vary ligation conditions:
   - **(a)** Concentration of the ligase (~1 mol of ligase per junction is typically required)
   - **(b)** Concentration of ATP. Although ATP is required, excess ATP may inhibit the reaction (Cherepanov and de Vries, 2003)
   - **(c)** Incubation time
   - **(d)** Incubation temperature
2. Test ligation of individual junctions
   - **(a)** Analysis of a test gel for the full length ligation often gives a clear indication as to which junction is not ligated efficiently. Identify which product corresponds to each band by its mobility and dye color.

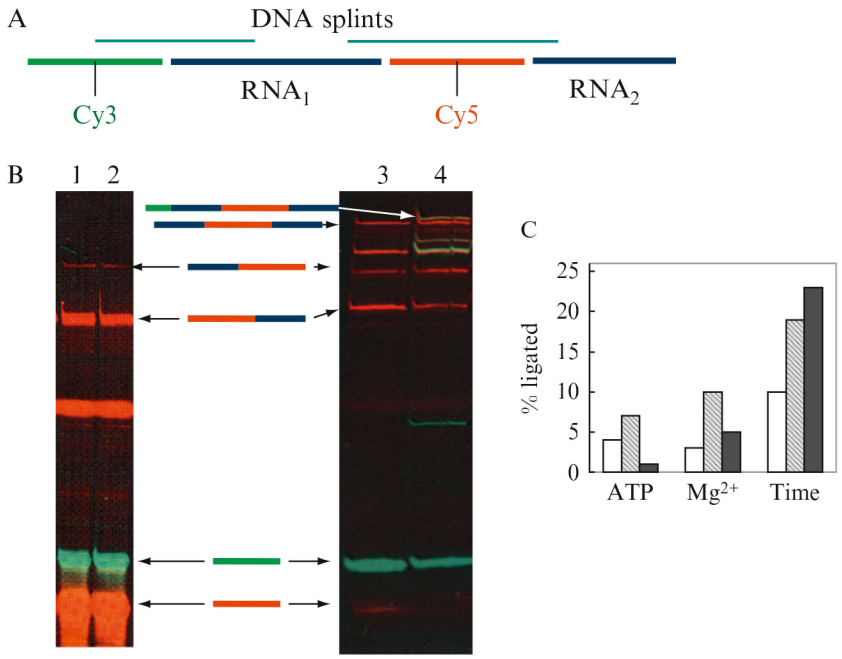


Figure 3.1 Optimization of the four-fragment ligations of double-labeled Tetrahymena group I ribozyme. (A) Design of the labeled construct. (B) Fluorescence scans of a denaturing PAGE of the ligation reactions before optimization (*left*) and after optimization (*right*). It is obvious that at the first junction Cy3-labeled oligo is not ligated to RNA_1 at all. Lanes 1 and 3 correspond to 12 h ligation time, lanes 2 and 4 correspond to 36 h ligation times. (C) Optimization of the ligation efficiency of the first junction. Parameters indicated on the *x*-axis were varied as follows: [ATP] was 10 μM (clear), 100 μM (hatched), 1 m*M* (black), for all reactions [Mg^{2+}] was 12 m*M*, reaction time 12 h; [Mg^{2+}] was 10 m*M* (clear), 20 m*M* (hatched), 30 m*M* (black), for all reactions [ATP] was 100 μM, reaction time 12 h; reaction time was 12 h (clear), 24 h (hatched), 36 h (black), for all reactions [ATP] was 100 μM, [Mg^{2+}] was 20 m*M*.

(b) If a particular "difficult" junction is identified, test if the 5′ side has the phosphate and the 3′ side has the free OH group (cyclic phosphate was efficiently removed). The presence of the 5′-phosphate in RNA oligos is easiest to confirm by the mass spectrometry. The presence of the free 3′-OH group can be detected by, for example, a primer extension assay with the Klenow fragment.

3. Test annealing by running the annealed complex on a nondenaturing gel and visualizing by fluorescent scanning. If low amount of fully annealed complex is observed, suggesting poor annealing efficiency, the following parameters should be optimized in parallel small scale reactions:

   (a) Concentration of RNA and DNA splints (middle fragments should *NEVER* be added in excess, since this *reduces* the concentration of the complexes.)
   (b) NaCl concentration
   (c) Additional salts to aid annealing (e.g., $MgCl_2$)
   (d) Rate of cooling, or several cycles of heating and cooling

ACKNOWLEDGMENTS

This work was supported by an NIH Program Project Grant (PO1-GM-066275). We thank W. Zhao, K. Travers, B. Sattin, and M. Forconi for sharing their experimental expertise.

REFERENCES

Baum, DanaA., and Scott, K. S. (2007). Deoxyribozyme-catalyzed labeling of RNA. *Angew Chem* **46,** 3502–3504.

Birikh, K. R., Heaton, P. A., and Eckstein, F. (1997). The structure, function and application of the hammerhead ribozyme. *Eur J Biochem* **245,** 1–16.

Brion, P., and Westhof, E. (1997). Hierarchy and dynamics of RNA folding. *Annu Rev Biophys Biomol Struct* **26,** 113–137.

Chen, J.-L., Nolan, J. M., Harris, M. E., and Pace, N. R. (1998). Comparative photocrosslinking analysis of the tertiary structures of *Escherichia coli* and *Bacillus subtilis* RNAse P RNAs. *EMBO J* **17,** 1515–1525.

Cherepanov, A. V., and de Vries, S. (2003). Kinetics and thermodynamics of nick sealing by T4 DNA ligase. *Eur J Biochem* **270,** 4315.

Dolinnaya, N. G., Sokolova, N. I., Gryaznova, O. I., and Shabarova, Z. A. (1988). Site-directed modification of DNA duplexes by chemical ligation. *Nucleic Acid Res* **16,** 3721–3738.

Dorywalska, M., Blanchard, S. C., Gonzalez, R. L., Kim, H. D., Chu, S., and Puglisi, J. D. (2005). Site-specific labeling of the ribosome for single-molecule spectroscopy. *Nucleic Acid Res* **33,** 182–189.

Downey, C. D., Fiore, J. L., Stoddard, C. D., Hodak, J. H., Nesbitt, D. J., and Pardi, A. (2006). Metal ion dependence, thermodynamics, and kinetics for intramolecular docking of a GAAA tetraloop and receptor connected by a flexible linker. *Biochemistry* **45,** 3664–3673.

Fang, X., Pan, T., and Sosnick, T. R. (1999). A thermodynamic framework and cooperativity in the tertiary folding of a Mg^{2+}-dependent ribozyme. *Biochemistry* **38,** 16840–16846.

Fedorova, O. A., Gottikh, M. B., Oretskaya, T. S., and Shabarova, Z. A. (1996). Cyanogen bromide-induced chemical ligation: Mechanism and optimization of the reaction conditions. *Nucleosides Nucleotides* **15,** 1137–1147.

Ferre-D'Amare, A. R., and Doudna, J. A. (1996). Use of *cis*- and *trans*-ribozymes to remove 5′ and 3′ heterogeneities from milligrams of *in vitro* transcribed RNA. *Nucleic Acid Res* **24,** 977–978.

Ha, T. (2001). Single-molecule fluorescence resonance energy transfer. *Methods* **25,** 78–86.

Ha, T., Enderle, T., Ogletree, D. F., Chemla, D. S., Selvin, P. R., and Weiss, S. (1996). Probing the interaction between two single molecules: Fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc Electrochem Soc* **93,** 6264–6268.

Ha, T., Zhuang, X., Kim, H. D., Orr, J. W., Williamson, J. R., and Chu, S. (1999). Ligand-induced conformational changes observed in single RNA molecules. *Proc Natl Acad Sci USA* **96,** 9077–9082.

Ho, C. K., and Shuman, S. (2002). Bacteriophage T4 RNA ligase 2 (gp24.1) exemplifies a family of RNA ligases found in all phylogenetic domains. *Proc Natl Acad Sci USA* **99,** 12709–12714.

Hodak, J. H., Fiore, J. L., Nesbitt, D. J., Downey, C. D., and Pardi, A. (2005). Docking kinetics and equilibrium of a GAAA tetraloop-receptor motif probed by single-molecule FRET. *Proc Natl Acad Sci USA* **102,** 10505–10510.

Jaeger, L., Michel, F., and Westhof, E. (1994). Involvement of a GNRA tetraloop in long-range RNA tertiary interactions. *J Mol Biol* **236,** 1271–1276.

Kurschat, W. C., Mueller, J., Wombacher, R., and Helm, M. (2005). Optimizing splinted ligation of highly structured small RNAs. *RNA* **11,** 1909–1914.

Lang, M. J., Fordyce, P. M., Engh, A. M., Neuman, K. C., and Block, S. M. (2004). Simultaneous, coincident optical trapping and single-molecule fluorescence. *Nat Methods* **1,** 133–139.

Lease, R. A., Adilakshmi, T., Heilman-Miller, S., and Woodson, S. A. (2007). Communication between RNA folding domains revealed by folding of circularly permuted ribozymes. *J Mol Biol* **373,** 197–210.

Lee, T.-H., Lapidus, L. J., Zhao, W., Travers, K. J., Herschlag, D., and Chu, S. (2007). Measuring the folding transition time of single RNA molecules. *Biophys J* **92,** 3275–3283.

Lehnert, V., Jaeger, L., Michele, F., and Westhof, E. (1996). New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: A complete 3D model of the Tetrahymena thermophila ribozyme. *Chem Biol* **3,** 993–1009.

Michel, F., and Westhof, E. (1990). Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J Mol Biol* **216,** 585–610.

Milligan, J. F., and Uhlenbeck, O. C. (1989). Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol* **180,** 51–62.

Moore, M., and Query, C. C. (2000). Joining of RNAs by splinted ligation. *Methods Enzymol* **317,** 109–123.

Moore, M. J., and Sharp, P. A. (1992). Site-specific modification of pre-mRNA: The 2′-hydroxyl groups at the splice sites. *Science* **256,** 992–997.

Pan, T. (2000). Probing RNA structure and function by circular permutations. *Methods Enzymol* **317,** 313–330.

Pan, T., Fang, X., and Sosnick, T. (1999). Pathway modulation, circular permutation and rapid RNA folding under kinetic control. *J Mol Biol* **286,** 721–731.

Pleiss, J. A., Derrick, M. L., and Uhlenbeck, O. C. (1998). T7 RNA polymerase produces 5′ end heterogeneity during *in vitro* transcription from certain templates. *RNA* **4,** 1313–1317.

Price, S. R., Ito, N., Oubridge, C., Avis, J. M., and Nagai, K. (1995). Crystallization of RNA-protein complexes I. Methods for the large-scale preparation of RNA suitable for crystallographic studies. *J Mol Biol* **249,** 398–408.

Proudnikov, D., and Mirzabekov, A. (1996). Chemical methods of DNA and RNA fluorescent labeling. *Nucleic Acid Res* **24,** 4535–4542.

Purtha, W. E., Coppins, R. L., Smalley, M. K., and Silverman, S. K. (2005). General deoxyribozyme-catalyzed synthesis of native 3′–5′ RNA linkages. *J Am Chem Soc* **127,** 13124–13125.

Qu, X., Smith, G. J., Lee, K. T., Sosnick, T. R., Pan, T., and Scherer, N. F. (2008). Single-molecule nonequilibrium periodic Mg^{2+}-concentration jump experiments reveal details of the early folding pathways of a large RNA. *Proc Natl Acad Sci USA* **105,** 6602–6607.

Rasnik, I., McKinney, S. A., and Ha, T. (2006). Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat Methods* **3,** 891–893.

Roy, R., Hohng, S., and Ha, T. (2008). A practical guide to single-molecule FRET. *Nat Methods* **5,** 507–516.

Russell, R., Zhuang, X., Babcock, H. P., Millett, I. S., Doniach, S., Chu, S., and Herschlag, D. (2002). Exploring the folding landscape of a structured RNA. *Proc Natl Acad Sci USA* **99,** 155–160.

Santoro, S. W., and Joyce, G. F. (1997). A general purpose RNA-cleaving DNA enzyme. *Proc Natl Acad Sci USA* **94,** 4262–4266.

Sattin, B. D., Zhao, W., Travers, K., Chu, S., and Herschlag, D. (2008). Direct measurement of tertiary contact cooperativity in RNA folding. *J Am Chem Soc* **130,** 6085–6087.

Schurer, H., Lang, K., Schuster, J., and Morl, M. (2002). A universal method to produce *in vitro* transcripts with homogeneous 3′ ends. *Nucleic Acid Res* **30,** e56.

Smith, G. J., Sosnick, T. R., Scherer, N. F., and Pan, T. A. O. (2005). Efficient fluorescence labeling of a large RNA through oligonucleotide hybridization. *RNA* **11,** 234–239.

Stone, M. D., Mihalusova, M., O'Connor, C. M., Prathapam, R., Collins, K., and Zhuang, X. (2007). Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature* **446,** 458–461.

Strobel, S. A., and Cech, T. R. (1993). Tertiary interactions with the internal guide sequence mediate docking of the P1 helix into the catalytic core of the Tetrahymena ribozyme. *Biochemistry* **32,** 13593–13604.

Tinoco, I., and Bustamante, C. (1999). How RNA folds. *J Mol Biol* **293,** 271–281.

Weiss, S. (2000). Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. *Nat Struct Mol Biol* **7,** 724–729.

Widengren, J., Chmyrov, A., Eggeling, C., Lofdahl, P.-A., and Seidel, C. A. M. (2007). Strategies to improve photostabilities in ultrasensitive fluorescence spectroscopy. *J Phys Chem* **111,** 429–440.

Xie, Z., Srividya, N., Sosnick, T. R., Pan, T., and Scherer, N. F. (2004). Single-molecule studies highlight conformational heterogeneity in the early folding steps of a large ribozyme. *Proc Natl Acad Sci USA* **101,** 534–539.

Zhuang, X. (2005). Single-molecule RNA science. *Annu Rev Biophys Biomol Struct* **34,** 399–414.

Zhuang, X., Bartley, L., Babcock, H. P., Russell, R., Ha, T., Herschlag, D., and Chu, S. (2000). A single-molecule study of RNA catalysis and folding. *Science* **288,** 2048–2051.

Zhuang, X., Kim, H., Pereira, M. J. B., Babcock, H. P., Walter, N. G., and Chu, S. (2002). Correlating structural dynamics and function in single ribozyme molecules. *Science* **296,** 1473–1476.