Optimizing splinted ligation of highly structured small RNAs

WOLFRAM C. KURSCHAT, JULIUS MÜLLER, RICHARD WOMBACHER, and MARK HELM

Ruprecht-Karls Universität Heidelberg, Institute of Pharmacy and Molecular Biotechnology, Department of Chemistry, D-69120 Heidelberg, Germany

ABSTRACT

The synthesis of highly structured small RNAs containing nonstandard nucleotides is of high interest for structural and functional investigations. A general approach is the joining, by T4 DNA ligase-mediated splinted ligation, of two or more RNA fragments, each of which may contain its own set of modified nucleotides. The RNA fragments hybridize with a complementary DNA splint to form a ternary ligation-competent-complex (LCC), which is then turned over by the DNA ligase. We studied the formation of the LCC and its precursors using size exclusion chromatography combined with a fluorescence detector. The spatial proximity of two cyanine-dye-labeled RNA fragments in LCCs was detected by monitoring FRET. An observed correlation of LCC formation and ligation yields suggests the use of long splints to stabilize LCCs. Splint oligos of increasing length, which in general appear to reduce the number of different hybridization intermediate species found in a reaction mixture, were applied to the synthesis by T4-DNA-ligation of two highly structured target molecules, one a 73mer tRNA, the other a 49mer synthetic ribozyme. A stable LCC could be isolated and turned over with > 95% ligation efficiency. In conclusion, the use of long splints presents a generally applicable means to overcome the low propensity of highly structured RNAs for hybridization, and thus to significantly improve ligation efficiencies.

Keywords: microsurgery; T4-DNA ligase; splint; ligation; hybridization; RNA structure; FRET

INTRODUCTION

The introduction of modified nucleotides into small structured RNAs is important for structural and functional studies, for example, by cross-linking (Moore and Query 1998; Reed and Chiara 1999; Yu 1999) or fluorescence resonance energy transfer (FRET) (Kim et al. 2002). A frequent strategy for the introduction of modified nucleotides involves the use of short RNA fragments, each carrying desired modifications, and subsequent joining of the fragments by the splint ligation method. Such ligation reactions can be programmed to form specific products by including bridging splint DNA templates to create RNA:RNA/DNA complexes which are nicked specifically at the desired site of ligation (Moore and Query 2000). These complexes are substrates of the T4 DNA ligase,

and are therefore termed ligation-competent-complexes (LCCs) (Fig. 1A) throughout this paper. Previous studies reported splints that extend 8 to 10 nucleotides (nt) on either side of the junction as sufficient for quantitative ligation, and that ligation efficiencies fall off dramatically with shorter templates (Moore and Query 2000). Although the use of longer splints (Stephens et al. 2000) and disrupter oligos (Strobel and Cech 1993) has been reported, a systematic evaluation of the benefits of longer splints or of optimized ligation of highly structured RNAs is not available. Such a study is presented here in the context of the synthesis of two highly structured RNAs. Site-specific labels were introduced into each fragment by replacement of a uridine by a 5-C6-Amino-2′-deoxythymidine that was conjugated to an NHS-derivative of a cyanine dye (Edwards and Sigurdsson 2005). The use of such labels allows facile and, by choice of the Cy3/Cy5 pair, orthogonal detection, meaning that the presence of one dye does not affect independent quantitative detection of the respective other dye in electrophoresis and chromatography. Size exclusion chromatography (SEC) was used to study the formation of different stable RNA:DNA intermediates (Fig. 1A). The Cy3/Cy5 pair incidentally also fulfills the prerequisites for FRET, a distance-dependent interaction between the electronic excited states of two dye molecules, in which excitation is transferred from a donor dye to an
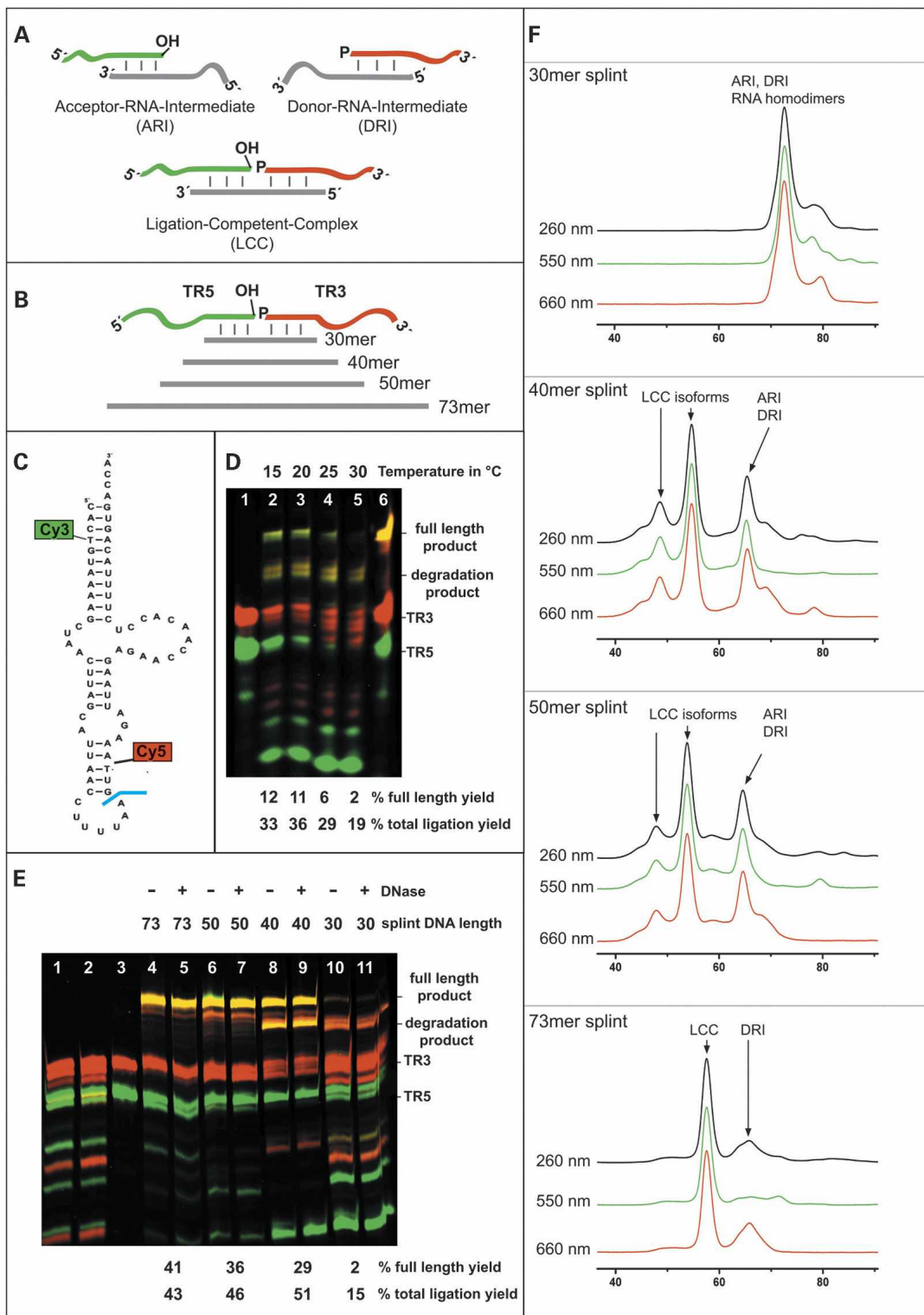
FIGURE 1. Splint ligation of mitochondrial tRNA[Lys]. (*A*) composition of ARI, DRI, and LCC; phosphates necessary for ligation are indicated by "p," and hydroxyl groups necessary for ligation are indicated by "OH".; DNA splints are in light gray; the phosphate acceptor RNA is in green and the donor RNA in red throughout the figure. (*B*) Splints of 30, 40, 50, and 73 nucleotides and RNA fragments TR5 and TR3. (*C*) tRNA[Lys] with attachment sites for Cy3 and Cy5. The ligation site is indicated in blue. (*D*) PAGE of ligation reactions mediated by a 30mer splint at different temperatures. (Lane *1*) Unligated fragments TR3,TR5. (Lanes *2–5*) Ligation reactions at 15°C, 20°C, 25°C, and 30°C. (Lane *6*) Contains size markers, TR3, TR5, and full-length product originating from an unrelated synthesis. Total ligation yield and full-length product yield are indicated *below* the PAGE. (*E*) Dependence of ligation yield on splint DNA length. (Lane *1*) Standard ligation without cDNA template. (Lane *2*) Standard ligation to which T4-DNA ligase was added after incubation with DNase. (Lane *3*) Unligated fragments TR3, TR5. (Lanes *4–7*) Standard ligations with splints of different length. Samples in lanes *5,7,9,11* were incubated with DNase after the ligation reaction. (*F*) SEC profiles of hybridization mixtures monitored by absorption at different wavelengths. The length of the splint is indicated in each panel. ARIs, DRIs, RNA dimers, LCCs, and structural isoforms as identified by accompanying titration experiments are indicated.

acceptor dye without emission of a photon. FRET has been used to observe formation of macromolecular complexes (Kim et al. 2002), and was used here to detect LCCs in SEC chromatograms.

RESULTS

Influence of temperature and splint length on ligation efficiency of tRNA^{Lys}

In a first series of experiments, a system of RNA fragments and splints (Fig. 1B) was designed to yield the human mitochondrial tRNA^{Lys} sequence of 73 nt, including two dye labels (Fig. 1C). The unmodified RNA adopts an atypical non-cloverleaf structure (Helm et al. 1998). Prior to ligation, the downstream fragment was 5′-phosphorylated by T4 polynucleotide kinase (T4-PNK) under conditions allowing both reactions to be carried out in a one-pot reaction without intermediate workup, yielding quantitative phosphorylation, and, as will be demonstrated later, quantitative ligation. Ligation of the dye-labeled RNA fragments TR5 and TR3 was first mediated by a 30mer splint DNA with a sequence complementarity of 15 nt on either side. Figure 1D shows the fluorescence scan of a PAGE analysis of ligation reactions performed at temperatures ranging from 15°C to 30°C. Overlay of Cy3 and Cy5 fluorescence scans identifies several bands as ligation products containing both dyes, of which the upper band represents the full-length product. The faster-moving products originate from degradation of the full-length ligation product, which is a known problem in splint ligation reactions (Moore and Query 1998, 2000). A mock incubation without ligase, which does not show any degradation (data not shown), identifies the RNase activity as associated with the ligase preparation. Quantification of ligation efficiencies was complicated by the fact that extensive degradation of ligation product may result in small fragments indistinguishable from those resulting from degradation of the starting material. Trace amounts (2%) of full-length product produced under standard conditions at 30°C (lane 5 in Fig. 1D) can only be barely distinguished from the background signal. The overall yield of ligation products, including degradation products, is 19% (quantified as described in Materials and Methods). Lower temperatures gradually increase the yield of full-length product to 12% at 15°C. These low yields may be attributed to structural features of the RNA fragments themselves, which enter into competition with, and thereby disfavor, hybridization. The positive influence of low temperature suggests that LCC stabilization is critical for ligation efficiency. Therefore, longer splints up to the full length of the target RNA were examined. The observed influence was quite pronounced (Fig. 1E). The yield of full-length product increases with splint length, increasing to 29% (40mer) and 36% (50mer) to reach 41% for the 73mer splint. When samples were split into two aliquots after ligation and one was incubated with

DNase I, no influence of the DNase treatment on degradation was detected. Degradation is most prominent with short splint DNA and decreases as the length of the splint oligo increases. Degradation products of the starting material, that is, the RNA fragments TR3 and TR5 are also observed in control reactions without splint (Fig. 1E, lane1). The full-length 73mer splint DNA protects most of the RNA, including the starting material, from degradation.

The above data indicate a splint DNA of maximum length as well as low temperature as parameters to improve full-length product formation. Because these parameters also favor hybridization of RNA fragments and splint DNA, thus suggesting a crucial role of LCC stability, we investigated the latter by SEC.

Separation of nucleic acids complexes by SEC

The contents of a reaction mixture before the addition of T4 DNA ligase were analyzed by SEC in combination with a diode array detector and a fluorescence detector. Monitoring absorbance at 550 nm and 660 nm allowed unambiguous and orthogonal identification of the two RNA fragments through the absorption spectra of the Cy3 and Cy5 dyes, while RNA fragments as well as the splint DNA contributed to absorption at 260 nm. A surprising number of different intermediates and complexes were detected and partially identified, based on retention times, the ratio of absorbance between the different absorption channels, and accompanying titration experiments, which are not shown here because of the large data volume. Intermediates formed by hybridization of splint DNA with an RNA fragment are abbreviated as DRI (donor-RNA-intermediate) and ARI (acceptor-RNA-intermediate), respectively, in the chromatograms (Fig. 1F), which also show some SEC-based experimental evidence of RNA dimers and several structural isoforms of LCCs. The RNA fragments TR3 and TR5 did not form any stable complex detectable by SEC. Stochiometric 1:1:1 mixtures, containing 2 μM of each RNA and splints of varying length, were annealed and analyzed by SEC. Most of the resulting chromatograms, displayed in Figure 1F, show a surprisingly high number of peaks. No LCC was detected in a mixture including the 30mer splint. This suggests that the low yield and the limited reproducibility of ligation reactions might be due to the failure of the 30mer splint DNA to form a stable LCC. A hybridization mixture now including a 40mer splint DNA shows a dramatic increase of nucleic acid complexes of high molecular weight, including several peaks below 60 min, corroborating the higher ligation yields obtained with 40mer splint DNA. The shape of the peaks is of identical aspect in all three absorption channels, suggesting that these peaks correspond to LCCs of different structural isoforms. Most importantly, no significant amounts of unhybridized RNA fragments TR5 or TR3 were detected, suggesting that a hybridization sequence of at least 20 complementary nucleotides is important for formation of complexes stable on the

SEC time scale. Further elongation of the splint to 50 nt resulted in only minor changes of the chromatogram. However, structural isoforms disappear when the 50mer splint is replaced with a 73mer splint, giving rise to one predominant peak. This presumed LCC peak was collected from the HPLC column, concentrated under native conditions, and submitted to a ligation reaction and subsequent analysis. Over 95% ligation yields were observed, which is striking in comparison to the incomplete ligation under similar conditions in the standard reaction mixture (Fig. 1E, lanes 4,5). These observations suggest that (1) the observed and isolated peak did indeed constitute the LCC; (2) the experimental conditions are optimal to permit quantitative ligation; and (3) an important factor for ligation efficiency under standard conditions may be formation of a stable LCC. The conclusion drawn from these data confirms our initial hypothesis, namely that to devise a strategy for optimal yield of full-length product, one must strive for maximum stability of the LCC.

Synthesis of double-labeled 49mer Diels-Alderase ribozyme

To verify the parameters established from the previous tRNA ligation experiments, synthesis of another highly structured RNA by splint-mediated ligation was optimized. The Diels-Alderase ribozyme (Fig. 2A) is a highly structured 49mer RNA that catalyses the Diels-Alder reaction between anthracen and maleimide and adopts a λ-shaped nested pseudo-knot structure (Serganov et al. 2005). Ligation product formation in the presence of a 30mer splint increased from 3% at 30°C gradually to 12% at 15°C. Use of longer splints at 30°C produced a strongly increased yield of full-length product ranging from 53% for the 40mer to 58% for the 49mer, as well as decreasing amounts of degradation products (Fig. 2C). Performing ligation with a 49mer splint at 15°C did not improve the yield any further.

Cy3 and Cy5 are separated by 21 nt in the primary sequence of this ribozyme. Assuming an A-form helix in an LCC, the six-carbon spacers, to which the dyes are attached, are separated by 60–65 Å. The dyes are thus in close enough

proximity to cause a FRET signal, as seen in the respective SEC traces (Fig. 2C). The SEC chromatograms share principal features with those of the tRNA system, namely that there is no LCC formation with a 30mer splint, and the full-length splint causes the simplest chromatogram with the highest yield of LCC formation. Differences compared to the tRNA system include the formation of RNA heterodimers in the absence of splint (Fig. 2C) and a complex with a
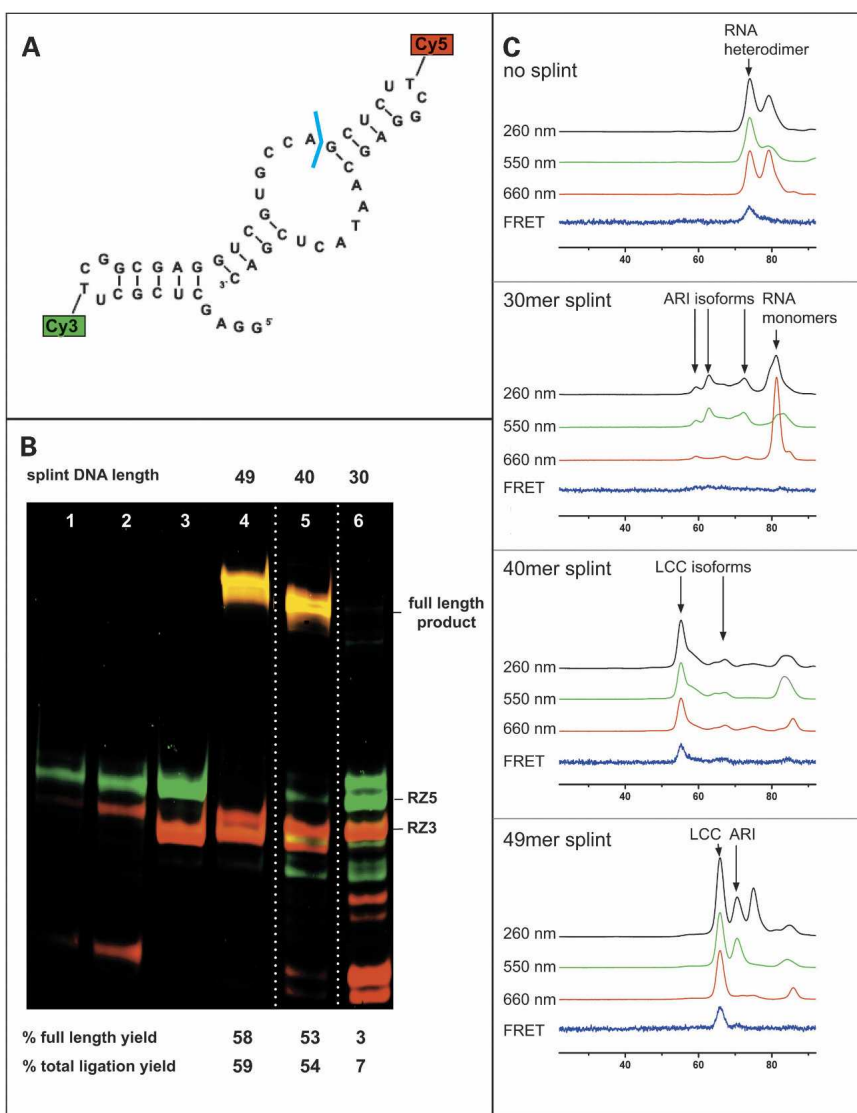

FIGURE 2. Splint ligation of a 49mer ribozyme. Three splint DNAs of 30, 40, and 49 nucleotides (nt) and the RNA fragments RZ5 (green) and RZ3 (red) were used in ligation studies. The color code is maintained throughout the figure. (*A*) Target ribozyme with attachment sites for Cy3 and Cy5. The ligation site is indicated in blue. (*B*) Dependence of ligation yield on splint DNA length. (Lane *1*) Standard ligation without cDNA template. (Lane *2*) Standard ligation to which T4-DNA ligase was added after incubation with DNase. (Lane *3*) Unligated fragments RZ3, RZ5. (Lanes *4–6*) Standard ligations with splint oligos of 49, 40, and 30 nt in length. As indicated by white lines *between* lanes *4* and *6*, these lanes stem from the same gel, but lanes containing nonrelevant experiments have been omitted. (*C*) SEC profiles of hybridization mixtures monitored by absorption at different wavelengths and by FRET detection. The length of the splint is indicated in each panel. ARIs, DRIs, RNA dimers, LCCs, and structural isoforms are indicated.

molecular weight high above that of the anticipated LCC in the presence of a 40mer splint. The corresponding FRET signal as well as the ligation yields obtained with the 40mer splint suggest that this complex may contain the LCC as a substructure. Ligation of an isolated LCC peak containing a 49mer splint yielded > 75% ligation product.

DISCUSSION

Here we investigated the synthesis of highly structured RNAs by splinted ligation. Conditions for the phosphorylation of chemically synthesized fragments and subsequent ligation were designed for a one-pot reaction, permitting both reactions to proceed to completion, as demonstrated by the quantitative turnover of an isolated LCC. Any suboptimal ligation yield under these conditions should therefore be due to suboptimal formation of stable LCCs, caused by nonoptimized stoichiometry, temperature, or splint length. SEC permits facile screening of these parameters, as it detects nucleic acid complexes which are stable on the time scale of a chromatography run, i.e., ~1 h. The technique also provides some indirect evidence of interactions which may be stable only on a shorter time scale. High amounts of ligation product were only obtained in reaction mixtures which also showed LCC formation in SEC analysis. This correlation marks SEC as a fast and convenient automated method for the screening of different nucleic acid constructs and stoichiometry. In comparison to, for example, native gel electrophoresis, LCC analysis displayed several additional advantages, in particular exact duplication of temperature and buffer conditions used in ligation reactions, including in particular a stable concentration of magnesium ions, which is very important for RNA tertiary structure. SEC analysis also reveals different structural isoforms of hybridization intermediates and LCCs, which can conveniently be recovered in ligation buffer directly from the column for further analysis. Fluorescence detection allows the identification of certain LCCs by FRET, provided that both fluorescent dyes are in close spatial proximity. This was the case for the presented LCC of the 49mer ribozyme, but not for the LCC of the tRNA.

Most experimental papers as well as a number of method review papers on splint-mediated ligation by T4-DNA ligase recommend splints of 15–30-nt total length, spanning eight to 15 nt on each RNA fragment (Strobel and Cech 1993; Reed and Chiara 1999; Moore and Query 2000). Although some authors suggest longer oligos which extend to 15–20 nt (Yu 1999; Frilander and Turunen 2005) or longer (Moore and Query 1998) on either side of the ligation site, experimental data are not given. In both of the present cases of highly structured RNAs, ligations at 30°C were inefficient with 30mer splints but significantly improved with 40mer splints, thus identifying a critical threshold between 15 and 20 nt of complementary sequence with each of the RNA fragments. This threshold may be lower at lower incubation

temperatures, as indicated by increased ligation yield with 30mer splints at 15°C. Regardless of the temperature, the present investigation supports the choice of splint oligos of maximum length in general, for several reasons. Similar to the previously reported use of disrupter oligos (Strobel and Cech 1993) designed to break up tertiary interactions in the RNA substrate, longer-splint DNAs tend to destabilize features of secondary structure in the RNA fragments. The general stabilization of the LCC results in (1) increased yield of LCC and ensuing ligation product; (2) the possibility to isolate a stable LCC; (3) reduction of non-ligation-competent intermediates; and (4) protection from degradation of ligation products in stable DNA-RNA duplexes.

MATERIAL AND METHODS

DNA splint oligos TR30DNA (5′-CTCTTAATCTTTAACTTAAA AGGTTAATGC-3′), TR40DNA (5′-TGGTTCTCTTAATCTTTAAC TTAAAAGGTTAATGCTAAGT-3′), TR50DNA (5′-GGTGTTGGT TCTCTTAATCTTTAACTTAAAAGGTTAATGCTAAGTTAGCT-3′), TR73DNA (5′-TGGTCACTGTAAAGAGGTGTTGGTTCTCTTAAT CTTTAACTTAAAAGGTTAATGCTAAGTTAGCTTTACAG TG-3′), RZ30 (5′-ATTGCTCCGAAGAGCTGGCACGACCTCGCC-3′), RZ40 (5′-CGAGTATTGCTCCGAAGAGCTGGCACGACCTCGCCGAA GC-3′), and RZ49 (5′-GTCGAGTATTGCTCCGAAGAGCTGGC ACGACCTCGCCGAAGCGAGCTCC-3′), and RNA fragments TR5 (5′-CACdT^{Cy3}GUAAAGCUAACUUAGCAUUAACCUUUUAA-3′) TR3 (5′-GUdT^{Cy5}AAAGAUUAAGAAGAACCAACACCUCUUUAC AGUCACCA-biotin-3′), RZ5 (5′-GGA GCUCGC UdT^{Cy3}CGGCGA GGUCGUGCCA-3′), and RZ3 (5′-GC UCUdT^{Cy5}CGGAGCAAUAC UCGAC-biotin-3′) were obtained from IBA (dT^{Cy3} = 2′-deoxythymi-dine-C6-Amino-NHS-Cy3; dT^{Cy5} = 2′-deoxythymidine-C6-Amino-NHS-Cy5). The Förster radius of the Cy3-Cy5 FRET pair is reported to be 56 Å by the manufacturer (Amersham Biosciences). Standard kinase-ligation (KL) buffer contained 50 mM TrisHCl pH 7.4 and 10 mM MgCl₂. For kinase and ligation reactions, ATP and DTT were added to final concentrations of 5 mM each. Nucleic acids were diluted in KL buffer, their UV-absorption spectra measured on a Varian Cary 50 Scan UV-spectrometer, and their concentrations readjusted before use. Donor RNAs (10 μM) were phosphorylated at 30°C for 30 min in the presence of 1.5 units/μL of T4-PNK (Fermentas). Prior to ligation or SEC analysis, 100 μL samples were submitted to heating at 75°C for 2 min, followed by cooling for 20 min at room temperature. The SEC analysis was conducted on a Pharmacia Superdex 200 10/300 GL column (Amersham Biosciences) mounted on an Agilent 1100 HPLC system. Separations were performed at 30°C at a flow rate of 0.2 mL/min KL buffer for 120 min. Nucleic acids were detected with a diode array detector by monitoring absorptions at 260 nm, 550 nm, and 660 nm. For FRET detection, the excitation wave-length of a fluorescence detector was set to 525 nm and the emission was monitored at 675 nm. FRET traces were corrected for Cy5 background emission. Concentration of LCC peaks was performed with a 3000-kDa MWCO spin filter (Vivascience). Standard ligation conditions included T4 DNA ligase (a stock solution of 30 Weiss units/μL was from Fermentas) at a final concentration of 1 unit/μL and incubation for 3 h at 30°C unless specified otherwise. DNA digestion was performed by incubation with 5 units of DNase I (Roche Applied Science) for 15 min at 30°C. For

analysis of ligation reactions, 20 μL of the ligation batch was mixed with 20 μL of dye-free loading buffer (90% formamide, 1xTBE), heat-denatured, and submitted to urea-PAGE. Dye-labeled RNAs were visualized by fluorescence scanning on a Typhoon 9400 (Amersham Biosciences) with default Cy3/Cy5 settings at a voltage of 400 V. Full-length yields were determined as the fraction of fluorescence of full-length product divided by the fluorescence of the entire lane, using Image Quant 5.2 software. Orthogonal quantification, that is, independent quantification of the same band with Cy3 or Cy5 fluorescence, respectively, yielded results consistent within a 2%–3% range. The total ligation yield was calculated by the same procedure, but using the fluorescence of all products longer than the starting material. Because extensively degraded ligation products, being shorter than the starting material, appear in the lower part of the gel and evade quantification, the results of the orthogonal quantification were not consistent. For all orthogonal quantifications, the lower of both values is given in the text and figures.

ACKNOWLEDGMENTS

REFERENCES

Edwards, T.E. and Sigurdsson, S.T. 2005. Modified RNA as tools in RNA biochemistry. In *Handbook of RNA biochemistry* (eds. R.K. Hartmann et al.), pp. 112–128. Wiley, Weinheim, Germany.

Frilander, M.F. and Turunen, J.J. 2005. RNA Ligation using T4 DNA ligase. In *Handbook of RNA biochemistry* (eds. R.K. Hartmann et al.), pp. 36–52. Wiley, Weinheim, Germany.

Helm, M., Brule, H., Degoul, F., Cepanec, C., Leroux, J.P., Giege, R., and Florentz, C. 1998. The presence of modified nucleotides is required for cloverleaf folding of a human mitochondrial tRNA. *Nucleic Acids Res.* **26:** 1636–1643.

Kim, H.D., Nienhaus, G.U., Ha, T., Orr, J.W., Williamson, J.R., and Chu, S. 2002. Mg2+-dependent conformational change of RNA studied by fluorescence correlation and FRET on immobilized single molecules. *Proc. Natl. Acad. Sci.* **99:** 4284–4289.

Moore, M.J. and Query, C.C. 1998. Use of site-specifically modified RNAs constructed by RNA ligation. In *RNA: protein interactions. A practical approach* (ed. C.W.J. Smith), pp. 75–108. Oxford University Press, Oxford.

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

Reed, R. and Chiara, M.D. 1999. Identification of RNA-protein contacts within functional ribonucleoprotein complexes by RNA site-specific labeling and UV crosslinking. *Methods* **18:** 3–12.

Serganov, A., Keiper, S., Malinina, L., Tereshko, V., Skripkin, E., Hobartner, C., Polonskaia, A., Phan, A.T., Wombacher, R., Micura, R., et al. 2005. Structural basis for Diels-Alder ribozyme-catalyzed carbon–carbon bond formation. *Nat. Struct. Mol. Biol.* **12:** 218–224.

Stephens, O.M., Yi-Brunozzi, H.Y., and Beal, P.A. 2000. Analysis of the RNA-editing reaction of ADAR2 with structural and fluorescent analogues of the GluR-B R/G editing site. *Biochemistry* **39:** 12243–12251.

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.

Yu, Y.T. 1999. Construction of 4-thiouridine site-specifically substituted RNAs for cross-linking studies. *Methods* **18:** 13–21.