Electronic Supplementary Material (ESI) for Chemical Communications

Universal Labeling of 5′-Triphosphate RNAs by Artificial RNA Ligase Enzyme with Broad Substrate Specificity

*John C. Haugner III and Burckhard Seelig**

Department of Biochemistry, Molecular Biology and Biophysics & BioTechnology Institute, University of Minnesota, 1479 Gortner Ave, Saint Paul, MN 55108 USA.

Supporting Information

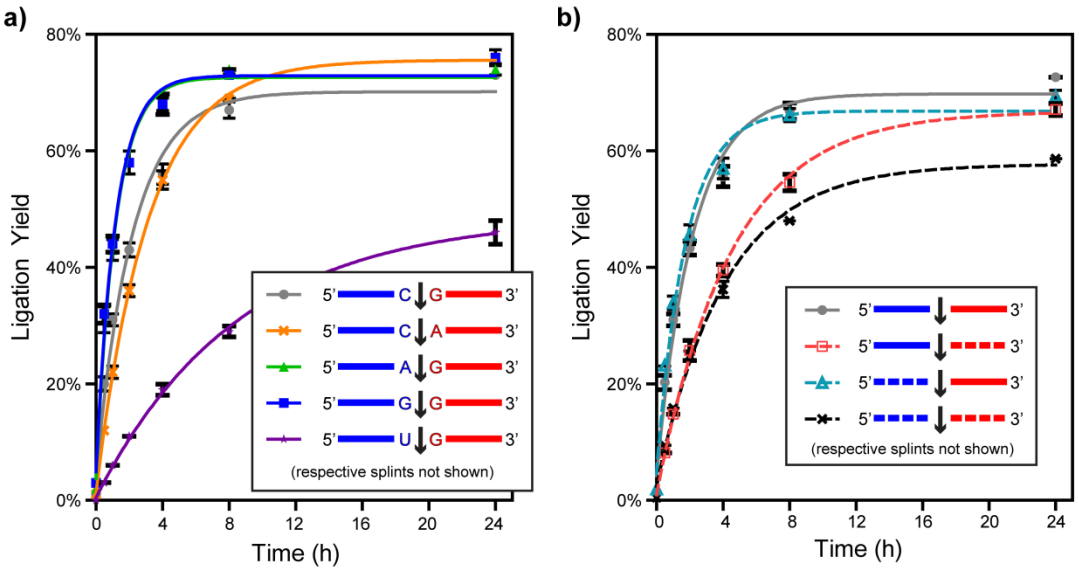

Fig. S1 Probing the sequence specificity of RNA ligase 10C with various substrate combinations. The PPP-RNA substrates and the RNA-OH substrates are shown in red and blue, respectively. The data are the mean of 3 replicates ± SD and were fit to first-order kinetics $Y=Y_f(1-e^{-kt})$ with Y_f=final yield and $k=k_{obs}$. The arrow represents the site of ligation and the complementary splint required for the reaction is omitted for clarity. **a)** Single nucleotide changes next to the ligation site. The C↓G substrates are the original two sequences that were used in the selection and evolution of RNA ligase 10C. Rates of ligation (k_{obs}) ranged from 0.12 ± 0.01 h^{-1} for U↓G to 0.85 ±0.05 h^{-1} for A↓G. While not all theoretically possible substrate nucleotide combinations were tested, the combination U↓A is likely to have the lowest reaction rate. This rate can be estimated by multiplying the k_{obs} (U↓G) by the fraction of k_{obs} (C↓A) / k_{obs} (C↓G) which yields an approximate value of 0.085 h^{-1}, assuming that the reaction rate contributions from nucleotides on both sides of the ligation site are independent. **b)** Change of whole sequence of ligation substrates. Progress of ligation is shown for combinations of the original substrate sequences used during the selection and evolution of ligase 10C (solid gray line) and substrate sequences that are completely unrelated, but have the same length and GC content (dashed lines). For detailed sequence information see **Table S1**.

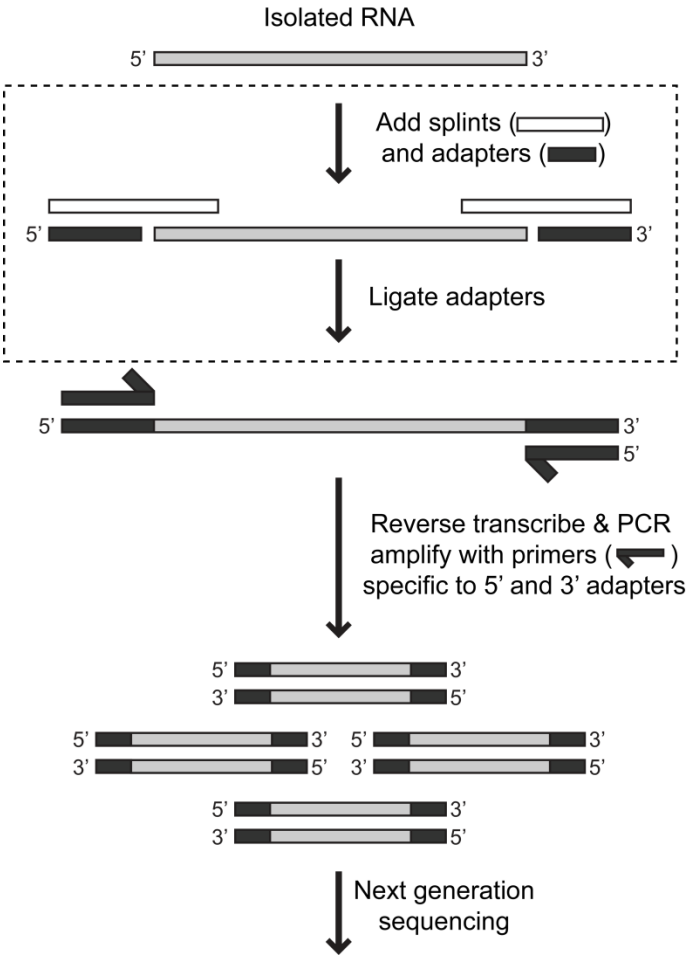

Fig. S2 General method for the modification of RNA samples necessary for next generation sequencing. RNA is typically not sequenced directly, but first converted to DNA through reverse transcription. Adapters add the needed terminal constant regions that facilitate the annealing of primers and reveal the orientation of the original RNA sequence. The artificial RNA ligase enzyme can be used during the "Ligate adapters" step to ligate the adapter to those "Isolated RNA" molecules that have a 5′-triphosphate group utilizing degenerate splints (dashed box). Ligase 10C can also ligate the second adapter to the 3′-terminus of the "Isolated RNA" if an adapter with a 5′-triphosphate is used.

Table S1. Oligonucleotide substrate combinations used in the sequence specificity and application experiments.

Substrate combinations	RNA-OH	PPP-RNA[a]	DNA Splint
Variation at ligation site[b]			
[c]	5′–CUAACGUUCG**C**	5′–**G**GAGACUCUUU	5′–GAGTCTCCGCGAACGT
	5′–CUAACGUUCG**C**	5′–**A**GAGACUCUUU	5′–GAGTCTCTGCGAACGT
	5′–CUAACGUUCG**A**	5′–**G**GAGACUCUUU	5′–GAGTCTCCTCGAACGT
	5′–CUAACGUUCG**G**	5′–**G**GAGACUCUUU	5′–GAGTCTCCCCGAACGT
	5′–CUAACGUUCG**U**	5′–**G**GAGACUCUUU	5′–GAGTCTCCACGAACGT
Variation of whole sequence			
[c]	5′–CUAACGUUCGC	5′–GGAGACUCUUU	5′–GAGTCTCCGCGAACGT
	5′–GCAUGUCAGCA	5′–AGGCCUAUCAA	5′–ATAGGCCTTGCTGACA
	5′–CUAACGUUCGC	5′–AGGCCUAUCAA	5′–ATAGGCCTGCGAACGT
	5′–GCAUGUCAGCA	5′–GGAGACUCUUU	5′–GAGTCTCCTGCTGACA
Secondary siRNA			
	5′–ACGUUCGA	5′–GGAGACUCUUU	5′–GAGTCTCCTCGAACGT
	5′–ACGUUCGA	5′–GCAGAAACUGGAACCCAGGU	5′–GTTTCTGCTCGAACGT
	5′–ACGUUCGA	5′–GGAACCCAGGUGUUGGUCUUUG	5′–TGGGTTCCTCGAACGT

[a] All oligonucleotides in this column carry a 5′-terminal triphosphate, which is not shown here to simplify the table.
[b] Nucleotides that were varied are underlined and shown in bold.
[c] This combination of substrate sequences is identical to the original substrates used in the selection and evolution of the RNA ligase enzymes.

Table S2. Melting temperatures (Tm) calculated for the hybridization of each PPP-RNA substrate with each of the three different splints used in the application experiment. The Tm values for fully complementary PPP-RNA/splint combinations are close to room temperature or above (shown in bold). All combinations that would result in mismatches yielded calculated Tm values that were substantially below room temperature (shown in parentheses and italics). In practical terms, this means that the fully complementary combinations are stable under the reaction conditions used in the application experiment, whereas the mismatched combinations are unlikely to hybridize.

PPP-RNA	Tm (by Pasteur[a])			Tm (by Stratagene[b])		
	Splint-1	Splint-2	Splint-3	Splint-1	Splint-2	Splint-3
Original (S1)	**30.1 °C**	-	-	**22.8 °C**	*(-19.9 °C)*	*(-27.3 °C)*
siRNA (S2)	-	**18.3 °C**	-	*(-14.8 °C)*	**17.6 °C**	*(-63.0 °C)*
siRNA (S3)	-	-	**30.2 °C**	*(-27.3 °C)*	*(-57.4 °C)*	**22.8 °C**

[a] Tm values were calculated using the Melting 4.1f calculator hosted by Mobyle@Pasteur (http://mobyle.pasteur.fr/cgi-bin/portal.py?#forms::melting) with adjustments made for the RNA/DNA duplex, [Na+] and [oligonucleotide].

[b] Tm values were calculated using Stratagene's QuikChange® Primer Design Program (Tm = 81.5 + 0.41(%GC) - (675/N) - % mismatch, where N = total number of bases. https://www.genomics.agilent.com/CollectionSubpage.aspx?PageType=Tool&SubPageType=ToolQCPD&PageID=15) Note that this calculator assumes a DNA/DNA duplex. The calculator was used because no suitable RNA/DNA calculator was available that considers mismatches. While the real Tm values for RNA/DNA duplexes are higher than those calculated for the respective DNA/DNA duplexes, the general trend between different sequences is expected to be similar.

Materials and Methods

All chemicals were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise stated.

Expression & Purification of RNA Ligase 10C:

RNA Ligase 10C was expressed and purified as previously published.[1]

Preparation of Oligonucleotides:

The α-^{32}P-labeled PPP-RNA substrates were prepared by *in vitro* transcription using T7 RNA polymerase as previously published.[1] The inactive α-^{32}P-labeled P-RNA substrates were prepared by treating PPP-RNA with 5′ RNA Polyphosphatase from Epicentre (Madison, WI) followed by phenol/chloroform extraction or PAGE gel purification to remove the Polyphosphatase. The RNA-OH substrates were purchased from Dharmacon (Lafayette, CO) and prepared according to the manufacturer's protocol. DNA splints were purchased from Integrated DNA Technologies (Coralville, IA). All oligonucleotides were dissolved in ultra-pure water and concentrations determined by UV absorbance.

Ligation Assay:

1 μM PPP-RNA, 3 μM RNA-OH and 6 μM DNA splint (Table S1) were combined in a buffer containing 20 mM HEPES pH 7.5, 100 mM NaCl, 100 μM $ZnCl_2$. The oligonucleotides were annealed by heating the solution to 60 °C for 3 minutes and allowing it to cool at room temperature for 10 min. A stock of 50 μM RNA Ligase 10C in buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl, 100 μM $ZnCl_2$ and 0.5 mM β-mercaptoethanol was added to the oligonucleotide mix to a final concentration of 10 μM enzyme. The ligation reactions were incubated at room temperature for the indicated times, and quenched with two volumes of 20 mM EDTA in 8 M urea, heated to 95 °C for 4 min and separated by 20% denaturing PAGE gel. The gel was analyzed using GE Healthcare (Amersham Bioscience) Phosphorimager and ImageQuant software (Amersham Bioscience).

References

1. F.-A. Chao, A. Morelli, J. C. Haugner, III, L. Churchfield, L. N. Hagmann, L. Shi, L. R. Masterson, R. Sarangi, G. Veglia and B. Seelig, *Nat. Chem. Biol.*, 2013, **9**, 81-83.