ChemComm

RSC Publishing

COMMUNICATION

Universal labeling of 5′-triphosphate RNAs by artificial RNA ligase enzyme with broad substrate specificity†

Cite this: *Chem. Commun.*, 2013, **49**, 7322

Received 14th June 2013,
Accepted 3rd July 2013

DOI: 10.1039/c3cc44454f

www.rsc.org/chemcomm

John C. Haugner III and Burckhard Seelig*

An artificial RNA ligase specific to RNA with a 5′-triphosphate (PPP-RNA) exhibits broad sequence specificity on model substrates and secondary siRNAs with direct applications in the identification of PPP-RNAs through sequencing.

The development of novel enzymes to manipulate nucleic acids has been a driving force in pioneering methods for molecular biology, genomics and transcriptomics. One such type of enzymes are ligases, which join together two strands of nucleic acids by forming a phosphodiester bond with a 3′–5′ linkage. Many known natural ligases couple the phosphodiester bond formation with ATP hydrolysis through a series of group transfers.[1] Ligases have been used extensively in the analysis of large mixtures of RNA to add adapters of known sequence to the RNAs,[2] thus enabling reverse transcription, PCR amplification and ultimately their identification by microarray analysis or sequencing.[3] As all known natural ligase enzymes require a 5′-monophosphate terminus (P-RNA) for the ligation reaction to occur, RNA mixtures are typically enzymatically modified to uniformly introduce the 5′-monophosphate in preparation for ligation and sequencing. This can prove problematic for interpreting subsequent RNA sequencing data as classes of RNA characterized by a different 5′-phosphorylation can no longer be easily distinguished in sequencing data. Identifying primary transcripts, also called immature RNA, can be particularly difficult because they have a 5′-triphosphate (PPP-RNA) and cannot be selectively purified from other classes of RNAs. While the triphosphate of eukaryotic mRNA is eventually capped, other classes of RNA retain the 5′-triphosphate as one of their only identifying features. This includes prokaryote mRNA, some viral mRNAs in eukaryotes,[4] transcripts from mitochondria[5] and chloroplasts,[6] and secondary siRNAs.[7] To identify these RNAs from cell isolates, methods exist which enrich for PPP-RNAs prior to sequencing, but they commonly identify false positives and require laborious independent verification with a complementary method.[8] Therefore, an alternative, simplified method for identifying PPP-RNAs with improved accuracy would be highly desirable.

The splinted, 5′-triphosphate-dependent ligation of two RNA strands has repeatedly been used as a model reaction to generate novel catalysts. Both artificial ribozymes[9] and deoxyribozymes[10] have been generated that catalyze this reaction. However, these catalytic nucleic acids often have strict sequence requirements near the ligation site, limiting the choice of PPP-RNA sequences that can be ligated. Furthermore, these nucleic acid enzymes act as both catalyst and complementary splint, which requires a new catalyst to be designed and synthesized specifically for each new substrate sequence. Previously, we reported the *de novo* selection and evolution of triphosphate dependent artificial RNA ligases.[11] These protein enzymes ligate a 5′-triphosphate RNA substrate to the 3′-hydroxyl of a second RNA exclusively forming a 3′–5′ linkage while both substrates are aligned by a complementary oligonucleotide splint (Fig. 1). The artificial ligases are zinc-dependent metalloproteins of ~10 kDa and were derived from the two zinc finger protein hRXRα.[11a] We solved the structure of RNA ligase 10C, a highly active variant,

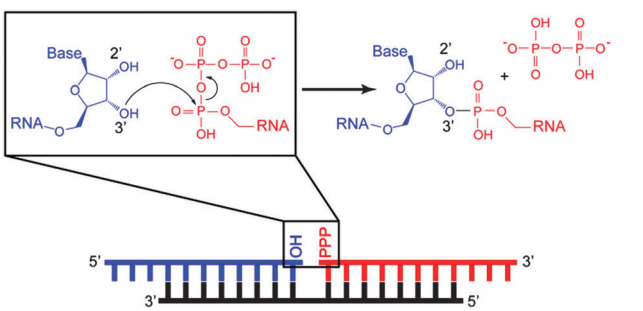

Fig. 1 Artificial RNA ligase catalyzing the formation of a 3′–5′-phosphodiester bond between a 5′-triphosphate RNA (red) and 3′-hydroxyl RNA (blue) substrate. The proposed mechanism of the reaction is shown. A complementary splint DNA oligonucleotide (black) anneals to the RNA and is necessary for ligation.

Department of Biochemistry, Molecular Biology and Biophysics & BioTechnology Institute, University of Minnesota, 1479 Gortner Ave, Saint Paul, MN 55108, USA. E-mail: seelig@umn.edu; Fax: +1-612-625-5780; Tel: +1-612-626-6281

† Electronic supplementary information (ESI) available: Detailed experiments examining sequence specificity, schematic of general RNA-sequencing protocol, table of oligonucleotide substrates, calculated T_m values for PPP-RNA–splint hybrids of the application experiments, materials and methods. See DOI: 10.1039/c3cc44454f

revealing a novel protein fold with highly unusual structural features.[12]

To determine if this enzyme could be used for the general identification of PPP-RNAs, we investigated the sequence specificity of ligase 10C. This artificial enzyme had been generated to ligate a single substrate pair, but had never been challenged to evolve an ability to ligate alternative RNA sequences. To determine the sequence specificity of 10C, we varied the nucleotides adjacent to the ligation site. We also tested a substrate pair of entirely unrelated sequence to assess the impact of large scale changes in sequence. Finally, to extend 10C for practical applications, we investigated whether this ligase can perform selective ligation in a mixture of secondary siRNAs. These experiments demonstrate RNA ligase 10C exhibits broad sequence specificity and therefore should be a valuable tool for the whole-cell isolation and identification of PPP-RNA.

RNA ligase 10C was assayed for the ligation of chemically synthesized RNA oligonucleotides to PPP-RNA substrates synthesized and ^{32}P-labeled by *in vitro* transcription. Assay conditions were chosen to mimic conditions used in commercial RNA isolation and sequencing kits.[13] Under these conditions, 10C was able to ligate the original substrates used during its evolution as well as an unrelated sequence, even though the nucleotides flanking the ligation site were completely different (Fig. 2). Ligation rates were similar for these two substrates with the k_{obs} of 0.45 ± 0.01 h^{-1} for the original substrates and 0.24 ± 0.01 h^{-1} for the substrate pair with unrelated sequence. We expanded upon these experiments by testing different nucleotides at the 3′-hydroxyl and 5′-triphosphate sides of the ligation site independently (Table S1, ESI†). As we were limited to the transcription by T7 RNA polymerase, we only created PPP-RNA substrates beginning with a G or A.[14] For all substrates tested, the k_{obs} were similar within a 10-fold range (Fig. S1, ESI†). By comparison, the most commonly used natural enzyme T4 DNA ligase has been shown to have an approximately 3-fold difference in k_{cat} in the ligation of nicked DNA depending on the sequence surrounding the ligation site.[15] If the ligation using 10C is allowed to proceed overnight as is commonly done with commercial ligases, the differences in yield for different substrates are reduced to <3-fold. Although the ligase 10C was originally selected to ligate only one specific pair of substrate sequences, the enzyme does not require any specific nucleotide sequence near the ligation site.

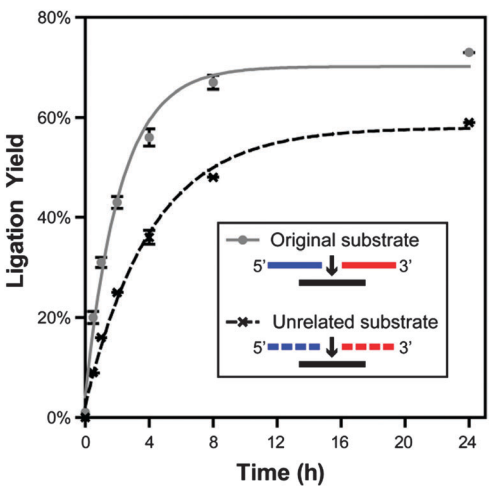

Fig. 2 Ligation of two pairs of RNA substrates by artificial ligase 10C. The arrow represents the site of ligation. The data were fit to first-order kinetics $Y = Y_f(1-e^{-kt})$ where Y_f = final yield and $k = k_{obs}$. The solid lines show the progress of the reaction with the substrates that were originally used to generate ligase 10C (5′ CUAACGUUCGC↓GGAGACUCUUU). Dashed lines indicate the ligation of a second substrate pair of unrelated sequence (5′ GCAUGUCAGCA↓AGGCCUAUCAA) that has the same GC content. Data are the mean of 3 replicates ± SD.

The ligation rate of ligase 10C is comparable to previously reported deoxyribozymes[10a] which catalyze the same reaction under similar reaction conditions, but is lower than the typical rate observed for natural ligase protein enzymes.[16] Nevertheless, the current activity of 10C is sufficient for immediate applications, and engineering efforts to improve the catalytic efficiency of 10C are underway. In contrast to previously generated nucleic acid enzymes,[9,10] ligase 10C can ligate all different RNA sequences without the need to tailor-make a modified catalyst for each substrate because the complementary splint is not part of, or attached to, the catalyst.

To directly demonstrate the utility of 10C, we tested the ligase with two secondary short interfering RNAs (siRNAs). siRNAs are involved in the regulation of gene expression through the RNAi pathway. Secondary siRNAs are a unique class of siRNA that were identified in *C. elegans* and have a 5′-triphosphate as they are synthesized by an RNA-dependent RNA polymerase to amplify the RNAi response.[7,17] We chose two previously published siRNAs (S2, S3 in Fig. 3a)[7a] which we transcribed with T7 polymerase. Those siRNAs lack secondary structure that could interfere in base pairing to the complementary DNA splint. We also included the original PPP-RNA substrate (S1) used to evolve ligase 10C as it shares a similar 5′-sequence with the secondary siRNAs. As expected, ligase 10C was capable of ligating 5′-triphosphorylated versions of all three substrates to an adapter sequence 8 nucleotides in length. While the 8 nt adapter was chosen to facilitate separation of RNAs in the gel, longer adapters of constant sequence can be used to enable subsequent RNA sequencing. Substrates of the same sequence but with a 5′-monophosphate instead of the 5′-triphosphate were not ligated (Fig. 3b).

We then combined the three 5′-triphosphorylated substrates in a single reaction and tested if individual splints could specifically promote the ligation of their complementary substrates in a mixture of similar sequences. No cross-reactivity between splints was detected, despite sequence identity of the substrates in three of the four nucleotides adjacent to the ligation site (Fig. 3b). The discrimination between substrates with a complementary splint and those with imperfect complementarity is >100-fold, using the limit of detection of the assay as a conservative estimate. Calculated melting temperatures (T_m) of the splint–substrate combinations indicate that only the matching sequence pairs will hybridize at the temperature and ionic strength of the ligation mixture (Table S2, ESI†). By using splints with relatively short overhangs (*e.g.* 8 base pairs), the hybridization stringency increases dramatically as even single mismatches can reduce the T_m by more than 10 °C. Therefore, the ligation of sequencing adapters to biologically

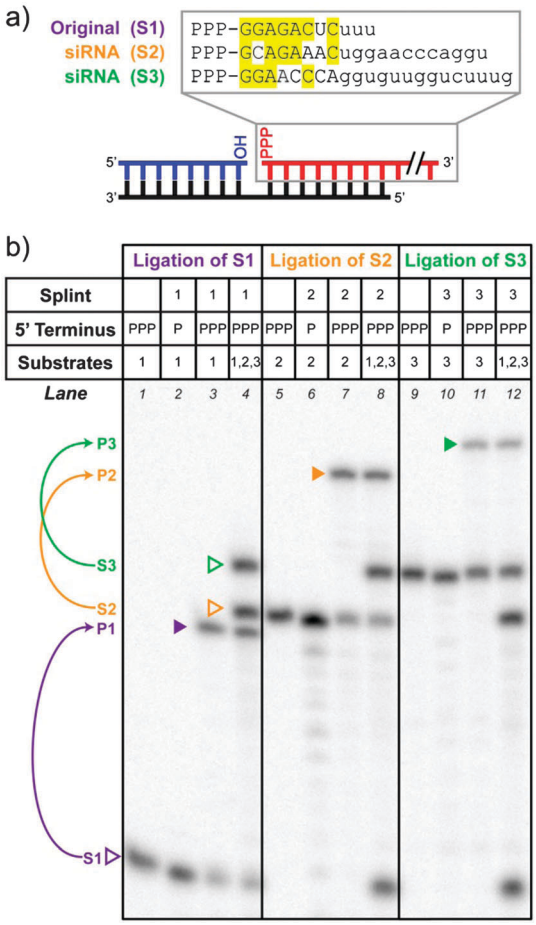

Fig. 3 Application of the artificial RNA ligase enzyme to selectively ligate secondary siRNA. (a) Alignment of the original ligation substrate (S1) and secondary siRNA substrates (S2) and (S3). Sequence similarities are highlighted in yellow. For each of the three substrates, a matching complementary splint oligodeoxynucleotide (black) was used to ligate them to a 3′-hydroxy-substrate ''adapter''. The adapter is identical in all three cases (blue). The first eight nucleotides of each substrate base-pair to its respective splint and are capitalized. (b) Denaturing PAGE gel depicting ligation experiments with substrates S1 to S3 to form products P1 to P3, demonstrating the necessity of a triphosphate (PPP) at the 5′-terminus (lanes 3, 7, 11). No ligation is observed for substrates with a 5′-monophosphate (P) (lanes 2, 6, 10). Lanes 4, 8 and 12 show the selective ligation of a single substrate in a mixture of substrates with highly similar sequences, which is dependent on the presence of the correct complementary splint. The substrate bands are marked by empty triangles and the product bands by filled triangles.

relevant RNAs by ligase 10C can be controlled through the use of a splint which specifically base pairs to the desired PPP-RNA substrate.

The artificial RNA ligase enzyme 10C has a potential application in the sequencing of PPP-RNAs (Fig. S2, ESI†). Deep RNA sequencing (RNA-seq) is replacing traditional microarray technology for large scale RNA analysis in part because of its increased dynamic range and lower detection limit. RNA-seq has been used to characterize specific RNA subpopulations, including those which are of low cellular abundance.[3] One of the primary challenges in the analysis of specific RNA classes is the minimization of false positive results due to sample contamination with other classes of RNA or cleavage products.[3] For PPP-RNAs, ligase 10C offers a simple and direct route for the ligation of adapters without false positives as the triphosphate provides the energy needed to drive bond formation. We envision that, in addition to sequencing specific sequences, total PPP-RNA could be analyzed by using oligonucleotide splints with a short degenerate sequence overhang. This would allow for ligation of any RNA sequence and has been used previously for RNA-seq in a similar fashion.[13a]

In summary, these results highlight the value of the artificial RNA ligase 10C for the direct isolation and identification of virtually any PPP-RNA. Efforts are underway to utilize ligase 10C in RNA sequencing efforts to evaluate the efficacy of 10C to identify a broad range of PPP-RNAs from biological samples. Finally, this exemplary application of the artificial RNA ligase highlights the potential of *in vitro* selection and evolution methods for the creation of useful artificial biocatalysts.[18]

This work was supported in part by the US National Aeronautics and Space Administration (NASA) Agreement no. NNX09AH70A through the NASA Astrobiology Institute–Ames Research Center, the US National Institutes of Health (NIH) (T32 GM08347) and by a grant from the Biocatalysis Initiative of the BioTechnology Institute at the University of Minnesota (to B.S.). We thank M. D. Lane, D. Morrone and M. V. Golynskiy for comments on the manuscript.

Notes and references

1 (*a*) I. R. Lehman, *Science*, 1974, **186**, 790–797; (*b*) S. Shuman, *J. Biol. Chem.*, 2009, **284**, 17365–17369.
2 (*a*) M. J. Moore and P. A. Sharp, *Science*, 1992, **256**, 992–997; (*b*) M. J. Moore and C. C. Query, in *Methods Enzymol.*, Academic Press, 2000, vol. 317, pp. 109–123.
3 F. Ozsolak and P. M. Milos, *Nat. Rev. Genet.*, 2011, **12**, 87–98.
4 V. Hornung, J. Ellegast, S. Kim, K. Brzózka, A. Jung, H. Kato, H. Poeck, S. Akira, K.-K. Conzelmann, M. Schlee, S. Endres and G. Hartmann, *Science*, 2006, **314**, 994–997.
5 D. A. Clayton, *Annu. Rev. Biochem.*, 1984, **53**, 573–594.
6 M. Sugiura, *Plant Mol. Biol.*, 1992, **19**, 149–168.
7 (*a*) T. Sijen, F. A. Steiner, K. L. Thijssen and R. H. A. Plasterk, *Science*, 2007, **315**, 244–247; (*b*) J. Pak and A. Fire, *Science*, 2007, **315**, 241–244.
8 Y. Qiu, B.-K. Cho, Y. S. Park, D. Lovley, B. Ø. Palsson and K. Zengler, *Genome Res.*, 2010, **20**, 1304–1311.
9 (*a*) D. P. Bartel and J. W. Szostak, *Science*, 1993, **261**, 1411–1418; (*b*) L. Jaeger, M. C. Wright and G. F. Joyce, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 14712–14717; (*c*) Y. Ikawa, K. Tsuda, S. Matsumura and T. Inoue, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 13750–13755.
10 (*a*) W. E. Purtha, R. L. Coppins, M. K. Smalley and S. K. Silverman, *J. Am. Chem. Soc.*, 2005, **127**, 13124–13125; (*b*) A. Flynn-Charlebois, Y. M. Wang, T. K. Prior, I. Rashid, K. A. Hoadley, R. L. Coppins, A. C. Wolf and S. K. Silverman, *J. Am. Chem. Soc.*, 2003, **125**, 2444–2454.
11 (*a*) B. Seelig and J. W. Szostak, *Nature*, 2007, **448**, 828–831; (*b*) B. Seelig, *Nat. Protocols*, 2011, **6**, 540–552.
12 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.
13 (*a*) S. R. Kuersten, *US Pat.*, US20100279305A1, 2010; (*b*) An excess of enzyme was used to reduce ligation times.
14 5′-Triphosphate RNA is commonly synthesized using T7 polymerase, which strongly disfavors the incorporation of 5′-C or U.
15 G. J. S. Lohman, L. Chen and T. C. Evans, *J. Biol. Chem.*, 2011, **286**, 44187–44196.
16 While k_{obs} of ligase 10C is approximately 3 orders of magnitude lower than k_{cat} of T4 DNA ligase, ligations are completed after overnight incubation.
17 K. Aoki, H. Moriguchi, T. Yoshioka, K. Okawa and H. Tabara, *EMBO J.*, 2007, **26**, 5007–5019.
18 (*a*) M. V. Golynskiy and B. Seelig, *Trends Biotechnol.*, 2010, **28**, 340–345; (*b*) M. V. Golynskiy, J. C. Haugner III, A. Morelli, D. Morrone and B. Seelig, *Methods Mol. Biol.*, 2013, **978**, 73–92.

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

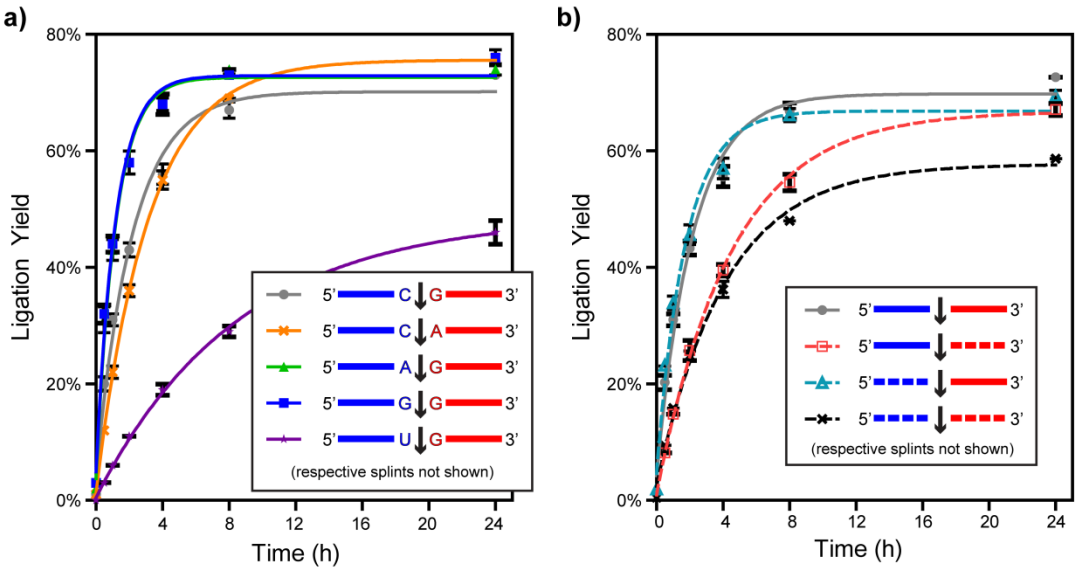

Fig. S1 Probing the sequence specificity of RNA ligase 10C with various substrate combinations. The PPP-RNA substrates and the RNA-OH substrates are shown in red and blue, respectively. The data are the mean of 3 replicates ± SD and were fit to first-order kinetics $Y=Y_f(1-e^{-kt})$ with Y_f=final yield 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**.

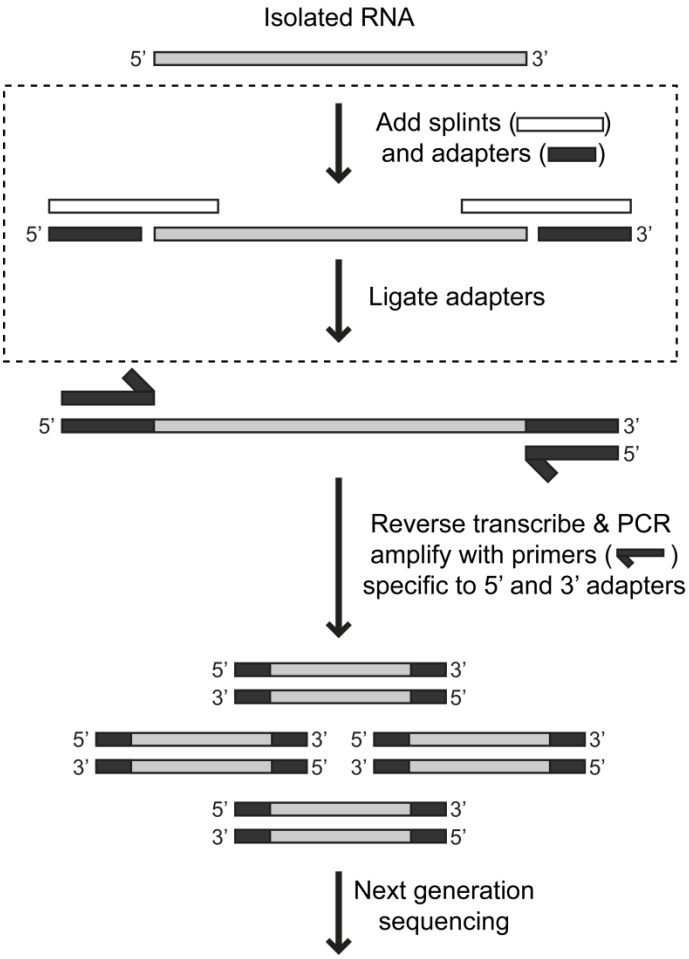

Fig. S2 General method for the modification of RNA samples necessary for next generation sequencing. RNA is typically not sequenced directly, but first converted to DNA through reverse transcription. Adapters add the needed terminal constant regions that facilitate the annealing of primers and reveal 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.