368–377 *Nucleic Acids Research, 2022, Vol. 50, No. 1* *Published online 20 December 2021*
https://doi.org/10.1093/nar/gkab1239

RNA self-splicing by engineered hairpin ribozyme variants

Robert Hieronymus, Jikang Zhu and Sabine Müller*

Institute of Biochemistry, University of Greifswald, Felix-Hausdorff-Str. 4, 17487 Greifswald, Germany

Received March 02, 2021; Revised November 26, 2021; Editorial Decision November 29, 2021; Accepted December 03, 2021

ABSTRACT

Small RNAs capable of self-cleavage and ligation might have been the precursors for the much more complex self-splicing group I and II introns in an early RNA world. Here, we demonstrate the activity of engineered hairpin ribozyme variants, which as self-splicing introns are removed from their parent RNA. In the process, two cleavage reactions are supported at the two intron-exon junctions, followed by ligation of the two generated exon fragments. As a result, the hairpin ribozyme, here acting as the self-splicing intron, is cut out. Two self-splicing hairpin ribozyme variants were investigated, one designed by hand, the other by a computer-aided approach. Both variants perform self-splicing, generating a cut-out intron and ligated exons.

INTRODUCTION

RNA splicing is a process of pre-mRNA maturation prior to translation, consisting of two transesterification reactions. As a result, the intron is cut out and the two remaining exons are ligated. This process is mediated in eukaryotes mainly by the spliceosome, a RNA–protein complex. The spliceosome does not exist in prokaryotes and archaea. Here, RNA splicing is mediated by the action of self-splicing group I or II introns. Hence, the group I and group II self-splicing introns are ribozymes that, in a similar way to the spliceosome, catalyse two transesterification reactions. Group II introns and the spliceosome share common architectural features and a common splicing mechanism. Furthermore, structural similarities between the proteins encoded by group II introns and the spliceosomal protein PrP8 suggest a common catalytic core, which substantiates the long-held assumption that group II introns and the spliceosome share a common evolutionary origin (1). It is unclear when introns and splicing arose in the evolution of life, but the efficiency of self-splicing introns suggests an origin in the RNA world (intron first or intron early hypothesis) (2). Perhaps self-splicing introns, as the selfish genes that they are seen as in a DNA world today, were just emerging as one of the first parasitic elements, inserting themselves into other RNA species to be replicated without regard to the overall effect on the host RNA (3). In the RNA world however, RNA organisms may have gained some real benefit from splicing. As seen in eukaryotes today, alternative splicing can be used to reorganise segments of a gene to obtain different phenotypes out of a probably limited genotype (4–6). In addition, exon-exon shuffling between genes within one RNA or trans-splicing between genes of two different RNAs has an evolutionary effect, which creates new combinations of genes (6), as strongly exemplified in the metazoan radiation (7). Furthermore, self-splicing introns may have been useful for transmission of genetic material between RNA populations through reverse splicing mechanisms (8). Excised introns may also have played a role in processes of RNA maturation or as regulatory factors, as seen today with miRNAs or snoRNAs abundantly found in introns (9). Lastly, splicing may lead to circularization of excised RNA (10), which may have been advantageous in terms of replication. End-to-end copying of a linear template requires the definition/identification of a specific initiation site, because replication could otherwise initiate anywhere, and complete replication would be rather unlikely. The circular RNA could have undergone rolling circle amplification without the need for a defined initiation site, and potentially may have led to a higher output of RNA copies.

Both, group I and II self-splicing introns, are rather large and highly-structured ribozymes, which facilitates their high efficiency and substrate generality, but also makes their appearance in the early RNA world highly improbable. Hence, other RNAs, as more simple precursors to these ribozymes, are more likely candidates to catalyse self-splicing in the RNA world. We recently demonstrated hairpin ribozyme (HPR) mediated RNA recombination (11,12). This included the rational design of a hairpin ribozyme variant, able to independently bind and cleave two different RNA substrates, followed by ligation of recombined fragments, thus demonstrating an evolutionary process by recombining non-functional RNAs to a functional hammer-

*To whom correspondence should be addressed. Tel: +49 3834 4204333; Fax: +49 3834 4204471; Email: smueller@uni-greifswald.de
Present addresses:
Sabine Müller, Jikang Zhu, Institut für Biochemie, Felix-Hausdorff-Str. 4, 17487 Greifswald, Germany.
Robert Hieronymus, Interdisciplinary Computing and Complex BioSystems, School of Computing, Newcastle University, Newcastle NE1 7RU, UK.

head ribozyme (11), and the functional hammerhead ribozyme further on to an effector responsive aptazyme (12). Ribozyme mediated RNA recombination has some similarities to the process of group I intron catalysed splicing, as demonstrated by Lehmann *et al*. by the construction of a recombinase ribozyme derived from the *azoarcus* group I intron (13,14). Both recombination and splicing occur by transesterification reactions, however, the mechanism of the transesterification catalysed by the hairpin ribozyme and the group I intron differ from each other. Reaction of the hairpin ribozyme is initiated by an internal attack of the 2′-OH on the neighbouring phosphate, leading to an in-line rejected 5′-OH and a 2′,3′-cyclic phosphate (cP) (15). The transesterification supported by the group I intron, is triggered by the nucleophilic attack of the 3′-OH of an external guanosine onto the phosphorous at the 5′-splice site. This results in an intron, which is elongated at its 5′-terminus by one guanosine and a 5′-exon carrying a 2′,3′-diol. The 3′-OH of the generated diol then attacks the phosphate at the 3′-splice site, resulting in ligation of the exons and excision of the intron RNA (16).

A hairpin ribozyme that performs the function of a self-splicing intron needs to catalyse three transesterification reactions: two cleavage reactions generating fragments with 5′-OH and 2′,3′-cP as intermediate products, and one ligation reaction between the 5′-OH of one fragment and the 2′,3′-cP of another. Due to the fact that transesterification is almost a net zero energy reaction, the additional reaction step may not be as large a drawback as it seems in comparison to the two-step self-splicing reaction supported by the group I intron. The high yields obtained in our hairpin ribozyme mediated recombination systems, where number and types of reaction steps are similar, supports this view (11).

MATERIALS AND METHODS

RNA synthesis

RNAs were generated by either phosphoramidite solid-phase synthesis or *in vitro* transcription. The phosphoramidite solid-phase synthesis was performed on a Gene Assembler Special DNA/RNA Synthesizer. Standard phosphoramidites and CPG supports were obtained from Link Technologies or ChemGenes. Amino linker phosphoramidites were used for the labelling reaction with ATTO680. For 5′-terminal labelling, we coupled 5′-TFA-Amino Modifier C6-CE Phosphoramidite (Link Technologies Ltd.) to the 5′-end, and for 3′-terminal labelling, we used 3′-Amino-Modifier C7 CPG 1000 (Link Technologies Ltd.) as starting material for oligonucleotide synthesis. Synthesized RNAs were incubated with a 1:1 (v/v) solution of 32% ammonia and 8 M ethanolic methylamine at 65°C for 40 min to remove base and phosphate protecting groups and to cleave the linkage to the solid-support. The released RNAs were lyophilized, and the obtained pellets were dissolved in a mixture of 600 μl TEA·3HF and 200 μl DMF to remove the 2′-*O*-TBDMS protecting groups. The mixture was incubated at 55°C for 90 min and the reaction was stopped by the addition of 200 μl water. The solution containing the deprotected oligonucleotide was transferred to a falcon tube, RNA was precipitated with butanol for 1 h or overnight and centrifuged at 6000 rpm for 30 min at 6°C. The RNA pellet was lyophilized and dissolved in gel loading buffer. RNAs were purified via preparative denaturing PAGE (10–15%). Bands containing the full-length RNA were cut out, reduced to small pieces and eluted passively in 0.3 M NaOAc (pH 5.4) over three steps. The solvent of the collected elution fractions was evaporated by vacuum centrifugation and the concentrated RNA was desalted via ethanol precipitation. For sequences of chemically synthesized RNAs see Supplementary Information (S7, Supplementary Table S1).

RNAs synthesized by in vitro transcription were generated from template DNA (biomers.net GmbH), using 1 μM antisense ssDNA template and 1.5 μM T7 promotor ssDNA (TAATACGACT CACTATAGGG) in HEPES buffer (20 mM NaHEPES (pH 7.5), 30 mM $MgCl_2$, 2 mM spermidine, 40 mM DTT). For the transcription of hairpin ribozyme strands 5′HPR-II and 5′HPR-V, hammerhead ribozyme sequences were incorporated in the DNA templates to cleave the starting sequence GGGAGA following the *in vitro* transcription, as described by Mörl *et al.* (17). The transcription mixture was heated to 90°C for 2 min and slowly cooled down to room temperature. 2 mM of NTP mix, 0.5 mM of extra GTP, 0.03 U pyrophosphatase and 20 U of T7 RNA polymerase were added to the mixture, followed by incubation at 37°C for 3–6 h. RNA was precipitated with 70% ethanol and purified via denaturing PAGE. Product containing bands were eluted by electroelution (18) and again precipitated in ethanol. For sequences of RNAs see Supplementary Information (S7, Supplementary Table S1).

ATTO680 labelling

5 nmol of amino-modified RNA was dissolved in 50 μl of a 50 mM HCO_3^- buffer (pH 8.2). 25 μg of the ATTO680-NHS ester (ATTO-TEC GmbH) was dissolved in an equal volume of DMF and was added to the RNA. The mixture was incubated for 1 h at room temperature, before precipitation of the RNA with 70% ethanol and purification via reversed phase HPLC (ÄKTA purifier; column: 250/4 nucleodur C18; gradient: 5% to 30% acetonitrile in 0.1 M triethylammonium acetate buffer (pH 7.5) over 60 min; flow rate: 0,5 ml/min).

Exon-exon ligation assay

Exon ligation was studied under varying conditions. RNA fragments corresponding to the extended 5′-Exon ($5'EX_{OV}$) and the truncated introns (5′IN[tr] and 3′IN[tr]) were incubated at stoichiometric concentration (50 nM) in Tris/HCl buffer (pH 7.5) in the presence of 10 or 50 mM $MgCl_2$, at 40 or 50°C, to allow cleavage to occur and to generate the 2′,3′-cyclic phosphate at the 3′-end of 5′EX *in situ*. After 120 min, a stoichiometric amount of the RNA fragments corresponding to the 3′-Exon (3′EX) was added to the reaction followed by denaturation at 90°C and slow cooling to 20, 30 or 40°C. The mixture was then incubated for another 120 min (Figure 2). All analyses were performed in triplicate measurement.

Splice reaction

Reaction conditions for the self-splicing hairpin ribozyme designed by hand (SH-I). Hairpin ribozyme strands 5′HPR and 3′HPR (both 5′-labelled with ATTO680) were added to 50 mM Tris/HCl buffer (pH 7.5), 50 mM $MgCl_2$ and 8% PEG4000 to give a final concentration of 50 nM of both RNAs in a volume of 20 μl. The reaction mixture was denatured at 90°C for 2 min and slowly (0.3°C/s) cooled down to 40°C. The mixture was incubated for 120 min, again denatured for 2 min at 90°C and further incubated for 120 min at 40°C. At the end of the reaction, an aliquot was taken and added to stop-mix solution (7 M urea, 50 mM EDTA).

Reaction conditions for the self-splicing hairpin ribozyme obtained by computer-aided design (SH-II). Hairpin ribozyme strands 5′HPR-II and 3′HPR-II (3′-labelled with ATTO680) were added to 50 nM Tris/HCl buffer (pH 7.5), 50 mM $MgCl_2$ and 8% PEG4000 to give a final concentration of 50 nM of both RNAs in a volume of 20 μl. The reaction mixture was denatured at 90°C for 2 min and slowly (0.3°C/s) cooled down to 40°C. The mixture was incubated for 120 min, again denatured for 2 min at 90°C and further incubated for 120 min at 10°C. At the end of the reaction an aliquot was taken and added to stop-mix solution (7 M urea, 50 mM EDTA).

All analyses were performed in triplicate measurement.

Analysis of cleavage/ligation/splicing

The ATTO680-labelled RNAs collected in the aliquots were analysed on a 12% denaturing PAGE on a LI-COR4300 DNA sequencer via fluorescence detection (700 nm channel). Signal intensities were determined with the software Image Studio Lite by LI-COR Bioscience. The portion of the non-cleaved, cleaved and/or ligated RNA species were calculated by the signal intensity of the band in relation to the sum of the signal intensity of every band in the respective lane. Cleavage yields are treated as the sum of the intensities of cleaved and ligated bands, as cleavage is a precondition for ligation.

Isolation of the splice product

For the isolation of ligated exons as the final splice product (5′EX-3′EX) we conducted a preparative reaction with 2 nmol of each hairpin ribozyme strand 3′HPR and 5′HPR (not ATTO680 labelled) solved in 50 mM Tris/HCl (pH 7.5), 50 mM $MgCl_2$ and 8% PEG4000 in an end volume of 115 μl. As self-degradation is common at high RNA concentration and denaturation temperatures, the RNA was mixed with water and Tris-buffer in the absence of $MgCl_2$. The mixture was then heated to 90°C for 2 min and cooled down to 40°C at a cooling rate of 0.2°C/s, followed by incubation for 15 min. $MgCl_2$ and PEG were added to the mixture and the reaction was allowed to proceed for 240 min at 40°C (SH-I), or for 120 min at 40°C and another 120 min at 10°C (SH-II). RNA was isolated from both reaction mixtures by ethanol precipitation, followed by purification via denaturing PAGE. The band at the same height as the 25 nt reference RNA was cut out, chopped into small pieces and passively eluted in three washing steps. The RNA in the collected washing solutions was again precipitated with ethanol and used for Sequence analysis.

Preparation for sequence analysis

For Sanger sequencing, the isolated ligated exons (5′EX-3′EX) were reverse transcribed into cDNA, using an extended RT primer (extended primer I reverse transcription (EPIR) for SH-I and extended primer II reverse transcription (EPIIR) for SH-II, Supplementary Table S1 in SI) to add a binding region for the sequencing primer at the 3′-end. 10 pmol of 5′EX-3′EX was mixed together with 11 pmol of RT primer in water, denatured at 65°C for 5 min and put on ice for 2 min. Subsequently, 0.5 mM dNTP mix, 0.2 μl RNase Inhibitor (Ribolock, Thermo Scientific), 1x PrimeScript buffer and 1 μl PrimeScript RTase (Takara Bio Inc.) was added to an end volume of 20 μl. The reaction was incubated overnight at 42°C. The RTase was inactivated by increasing the temperature to 70°C for 15 min. This was followed by the addition of RNase H and further incubation at 37°C for 1 h. An aliquot of the reaction was analysed via denaturing PAGE (10%, SYBR-Gold stained; SI, Supplementary Figure S6). An aliquot of 4 μl was taken from the crude RT reaction and added to 11 μl of a mixture containing 1x Pfu buffer, 0.2 mM dNTP mix, 9 pmol RT primer, 9 pmol forward primer (extended primer I forward (EPIF) for SH-I and extended primer II forward (EPIIF) for SH-II, Supplementary Table S1 in SI) and 2 U PfuPlus! polymerase (roboklon). For DNA extension and amplification, the following PCR cycle was used: 60 s, 94°C → 30x [30 s, 94°C → 35 s, 68°C → 10 s, 72°C] → 60 s, 72°C → hold at 4°C. An aliquot of the PCR reaction was analysed by denaturing PAGE (10%, SYBR-Gold stained; SI, Supplementary Figure S6). The generated dsDNA strands were mixed with SEQplus sequencing primer (M13 forward and reversed) from Generi Biotech, necessary for the sequencing of short DNA strands, and were sent to Eurofins for Sanger sequencing.

RESULTS AND DISCUSSION

Rational design of a self-splicing hairpin ribozyme

Our concept for rational design of a self-splicing ribozyme includes a hairpin ribozyme that is flanked by two exons at its 5′- and 3′-terminus (Figure 1). The intron-exon junctions are the sites for hairpin ribozyme mediated cleavage. To support a splice reaction, the hairpin ribozyme construct must be able to fold into two alternative conformations, both addressing one of the two possible cleavage sites. After first cleavage took place in either conformation, generating a 5′-OH and 2′,3′-cP, dissociation of the cleaved RNA fragment and refolding of the hairpin ribozyme into the alternative conformation allows the second cleavage reaction to occur, also generating a 5′-OH and 2′,3′-cP. Both, the 5′-exon (carrying the 2′,3′-cP) and the 3′-exon (carrying the 5′-OH) then hybridise at the hairpin ribozyme binding site and are ligated by the intron. Since the excised intron also contains a 5′-OH and a 2′,3′-cP at its termini, it competes with exon ligation, to form a circular intron by self-ligation (Figure 1). To favour exon-exon ligation and to repress circularization, we designed the binding strength of individual sequences in a way that intermolecular binding of exons is favoured over

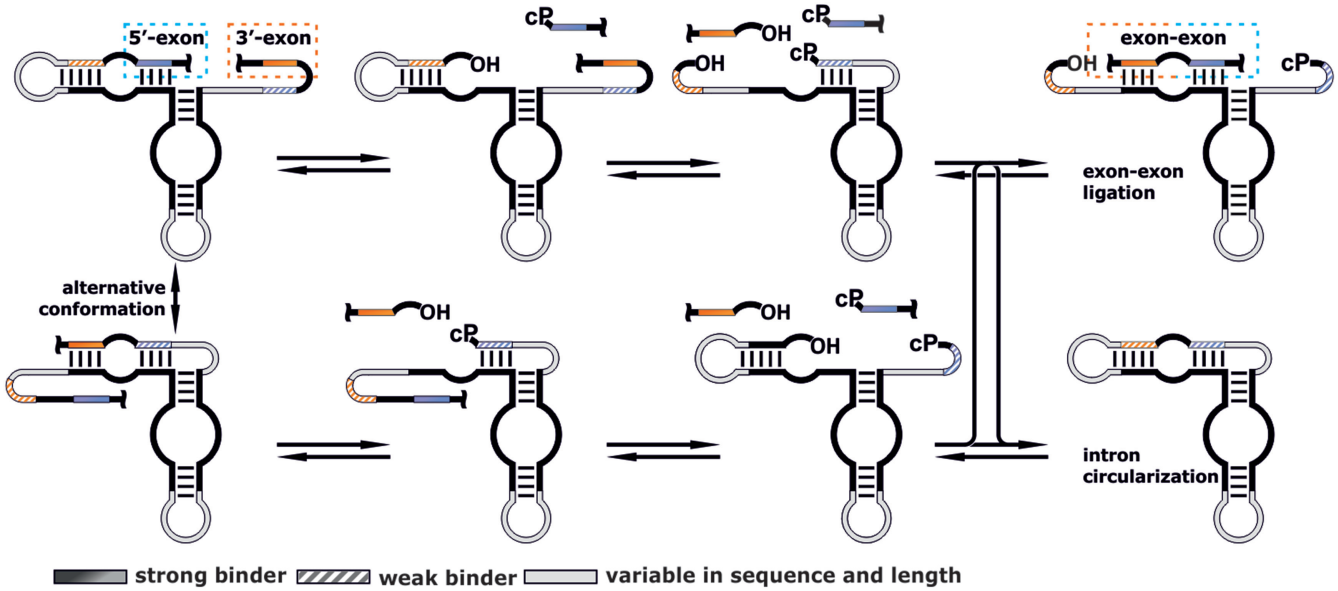

Figure 1. Scheme of hairpin ribozyme mediated self-splicing. The ribozyme in its initial state can fold into two alternative conformations. The hairpin ribozyme core structure is flanked at both termini by a 5′- and 3′-exon. The exon-intron boundaries correspond to the hairpin ribozyme cleavage sites. The two cleavage sites are flanked by a weak (in the intron, orange and blue striped lines) and strong binding site (in the exon, orange and blue filled lines), respectively. Light grey areas indicate regions variable in length and sequence. The first cleavage reaction generates a 5′-OH and a 2′,3′-cyclic phosphate terminus, and is followed by refolding of the ribozyme into the alternative conformation, which initiates the second cleavage reaction. For formation of the splice product, both the 5′- and 3′-exon bind to the hairpin ribozyme structure and are ligated. Cyclization of the intron is a competing side reaction.

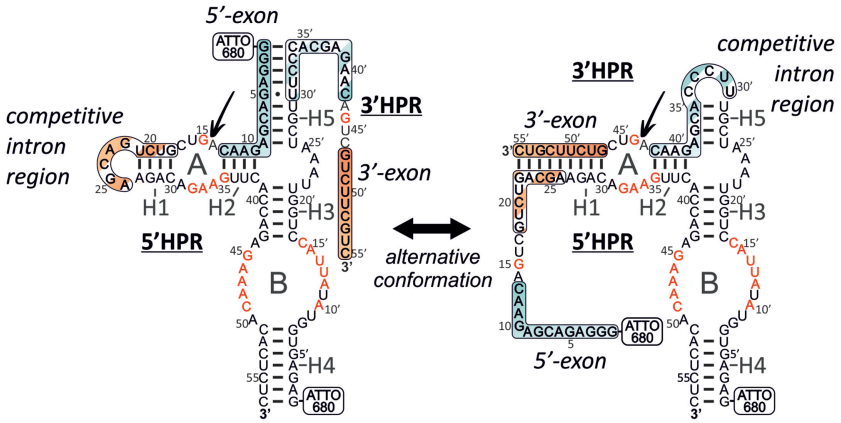

Figure 2. Sequence and secondary structures of the two alternative conformations of the rationally designed self-splicing hairpin ribozyme (SH-I). The arrow marks the cleavage/ligation site. Blue region: 5′-exon. Orange region: 3′-exon. Striped region: alternative binding sites containing competing intron regions (weaker hybridisation). Red nucleotides: highly conserved nucleotides in the hairpin ribozyme.

intramolecular binding of intron termini (filled and dashed lines in Figure 1). This also supports the desired preference of the hairpin ribozyme for cleavage or ligation, which is dependent on substrate binding strength (19). Thus, for weakly binding RNA substrates cleavage is favoured, while strongly binding RNA fragments are preferentially ligated.

It should be noted that the terms exon and intron are without functional meaning in this concept study, and are just used to address the different sequence parts of the hairpin ribozyme structures in the self-splicing reaction. Both intron and exons may be extended and/or varied in length and sequence in other designs or in a model RNA organism.

First, RNAs needed to be prepared, taking care that premature splicing is prevented. Due to the length of the desired sequences, synthesis of the full-length RNAs by solid phase phosphoramidite chemistry was a less suitable option. *In vitro* transcription appeared to be the better choice, however, required the presence of magnesium ions. This would lead to start of the splicing process already during transcription. In order to control the start of the splicing reaction, we refined our sequence design and divided the ribozyme into two RNA strands by removing the hairpin loop closing helix 4. This way, we obtained a controllable system, composed of two RNA strands, which only form an active complex, if both strands, '5′HPR' containing the 5′-exon and '3′HPR' containing the 3′-exon, are present (Figure 2).

The hairpin ribozyme is a superior candidate for rational design, because it has only a few highly conserved nucleotides in its sequence (highlighted in red in Figure 2) (20).

This allowed us to integrate into the design almost any sequence as long as the secondary structure with two loops and four helices is preserved. We chose the wild type sequence A•GUC (the dot marks the cleavage/ligation site) within loop A as boundary between 5′-exon and intron (Figure 2, conformation on the left) as well as between intron and 3′-exon (Figure 2, conformation on the right). Helix 1 and 2 next to loop A are formed by binding of the intron/exon parts to be processed, and have the same sequence in both alternative conformations. We integrated another helix (H5) as an extended binding site. Thus, in addition to formation of helix 2 in both conformations, the exons hybridise with the ribozyme to form a 9 bp helix 5 (between ribozyme and 5′-exon, Figure 2 on the left) or a 9 bp helix 1 (between ribozyme and 3′-exon, Figure 2 on the right) in either conformation, while helix 1 in the left structure and helix 5 in the right structure, both resulting from intramolecular folding of the intron have only 4 bp (Figure 2). This setup should favour re-hybridisation of the cut off exons and thus exon-exon ligation over ligation of the intron segments.

Every step in the design process was verified by secondary structure prediction with RNAstructure from the Mathews lab (21).

Studying conditions for best exon-exon ligation

For experimental validation, we chose a stepwise procedure. First, we investigated the exon-exon ligation reaction in a separate experiment. For this purpose, we separately synthesised the exon and intron fragments. The exons (5′EX and 3′EX) were generated by phosphoramidite solid-phase synthesis, and the intron segments (5′IN and 3′IN) by *in vitro* transcription. The 5′-exon was labelled at the 5′-terminus with ATTO680 to follow ligation via PAGE on a LI-COR 4300. To obtain the 2′,3′-cP required for the ligation reaction, the 5′-exon was extended at the 3′-end by a 7 nt long overhang (5′EX_{OV}) to enable a cleavage reaction by the hairpin ribozyme (assembled from 5′IN + 3′IN), and thus to generate the 2′,3′-cP *in situ* (Figure 3A, top). To look solely at exon binding and ligation, the intron fragments 5′IN and 3′IN, which assemble to form the active hairpin ribozyme, were truncated in order to prevent competition of intron parts with the exons for binding and ligation. In the following, these truncated introns are named 5′IN[tr] and 3′IN[tr] (Figure 3A).

Overall, higher Mg^{2+} concentrations in combination with lower reaction temperatures appear to favour the cleavage as well as the ligation reaction (Figure 3B). While with 10 mM $MgCl_2$ an increase in temperature leads to stronger cleavage, with 50 mM $MgCl_2$ a higher temperature is not necessary to achieve high cleavage rates in the first phase of the reaction (1–120 min). The highest exon-exon ligation yield (37%) was achieved at 50 mM $MgCl_2$ and 40°C over the entire 240 min reaction time. Interestingly, we observed a side reaction that occurs during the first 120 min of the reaction. Apparently, the 5′-exon carrying a 3′-terminal 2′,3′-cyclic phosphate (5′EX_{cP}) resulting from cleavage of the extended 5′-exon (5′EX_{OV}), reacts with the truncated 5′-intron segment (5′IN[tr]) to form a RNA of 53 nt in length (Figure 3A and C). Assuming ligation takes place as shown in Figure 3A, the ligation site would be A•GAC, which has been reported to be a poor cleavage site (22), but obviously has unexplored ligation qualities. In this construct, helix 1 consists of a single base pair and the whole structure would be stable only at low temperature or higher magnesium ion concentrations. With the addition of the 3′-exon (3′EX) to the reaction mix after the first 120 min, the side product disappears again, which is most likely a consequence of its re-cleavage, followed by competitive binding and ligation of the 3′-exon. Therefore, the observed formation of the side product can be neglected for exon-exon ligation. More efficient cleavage and ligation was achieved by the addition of the 3′-exon at the start of the reaction (50% exon-exon ligation). In this case, formation of the side product was completely abolished (Figures 4A and 5B).

Assembly study with RNAs differing in length

With the very encouraging results of exon-exon ligation at hand, we increased the length of our RNA species stepwise by integrating the non-truncated introns (5′IN and 3′IN) and finally the full-length hairpin ribozyme fragments (5′HPR and 3′HPR). Various combinations of those RNA species were tested for cleavage and ligation activities (Figures 4 and 5), culminating in the full-length self-splicing ribozyme formed by the hybridisation of 5′HPR and 3′HPR (Figure 4G). Progress of reaction was followed by using 5′-labelled versions of either the extended 5′-exon (5′EX_{OV}) or ribozyme strand 5′HPR (coloured blue in Figures 4 and 5). To determine if the position of the dye affects the reaction, the assays were repeated with 3′-labelled 3′-exon as control (coloured orange in Figures 4 and 5). Furthermore, 5′-labelled ribozyme strand 3′HPR was used to observe the competing intron-intron ligation in reaction schemes F and G (Figure 4F and G).

Exon-exon ligation yield decreases from 50% to 22% (in Figure 5A and B, lanes/columns A and B) when the non-truncated 5′-intron segment (5′IN) is used instead of the truncated version of the 5′-intron (5′IN[tr]), due to the competition of the extended 5′-end of 5′IN with the extended 5′-exon (5′EX_{OV}) for binding (Figure 4B). The reduction of the yield can be assigned to a side reaction, which is the ligation of the 5′-intron with cleaved 5′EX_{OV} to form ribozyme strand 5′HPR (37% yield, lane B in Figure 5A and B). The generation of 5′HPR is probably favoured by the lower entropy cost for ligation in this three-stranded system as compared to the complex consisting of four RNAs needed for exon-exon ligation, although this would be the more stable structure with regard to binding enthalpies. When the truncated 3′-intron segment (3′IN[tr]) is substituted with the full-length 3′-intron (3′IN) (Figure 4C), no side product (in this case ribozyme strand 3′HPR) can be formed due to the fact that the synthesised 3′-intron (3′IN) lacks the 2′,3′-cyclic phosphate required for ligation with the 3′-exon (3′EX) (Figure 5A and B, lane/column C). However, one can estimate the effect of binding competition, which results in decrease of ligation yield from 50% to 38% (Figure 5A and B, lanes/columns A and C). With both non-truncated introns present (Figure 4D), both effects seen for

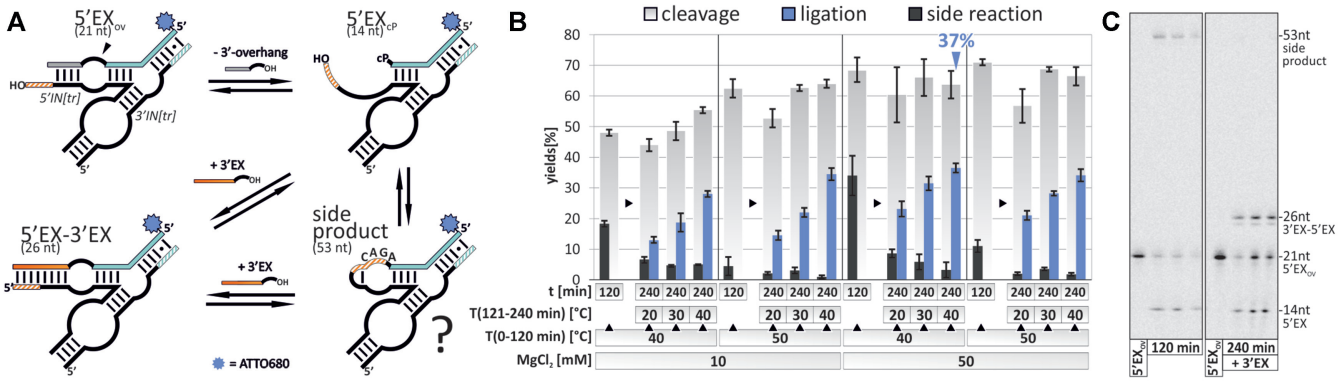

Figure 3. (**A**) Scheme of reaction of the extended 5′-exon (5′EX$_{OV}$, ATTO680 labelled) with truncated intron fragments 3′IN[tr] and 5′IN[tr]. (**B**) Yields of reactions shown in (**A**) at different $MgCl_2$ concentrations and temperatures. The 3′-exon (3′EX) was added to the reaction mix 120 min after reaction start. The mixture was shortly denatured and then incubated for another 120 min at different temperatures. Analyses were performed in triplicate measurement. (**C**) Denaturing PAGE (12%) of the reaction at 50 mM $MgCl_2$, 40°C (0–120 min) → x°C (121–240 min). Visualisation by fluorescence imaging (ATTO680 labelled RNA) on a LI-COR4300 DNA Sequencer (a supporting gel for size determination of the side product is provided in SI, Supplementary Figure S1).

the schemes depicted in Figure 4B and C combine to cause a reduction of ligation yield to only 16% (Figure 5A and B, lane/column D).

We then introduced the full-length RNA species into the system, combining one full-length hairpin ribozyme strand (5′HPR or 3′HPR) with the fragmented counterpart (5′EX + 5′IN or 3′EX + 3′IN) (Figure 4E and F). Ribozyme strand 5′HPR is cleaved to only 36% (Figure 5A and B, lane/column E), corresponding to the observed high yield for the reverse reaction being ligation of the 5′-exon (5′EX$_{cP}$) with the 5′-intron segment (5′IN) to ribozyme strand 5′HPR, as observed in the reaction scheme before (Figures 4B and 5A and B, lanes/columns B). Thus, for 5′HPR as initial substrate, equilibrium is shifted towards ligation. Cleavage of ribozyme strand 3′HPR proceeds more efficiently (68%), while cleavage of the extended 5′-exon (5′EX$_{OV}$) is less efficient (41%, Figure 5A and B, lane/column F) compared to the scenario shown in Figure 4D, where the 3′-exon (3′EX) and the 3′-intron segment (3′IN) are present instead of 3′HPR (65%, Figure 5A and B, lane/column D). This drop in 5′EX$_{OV}$ cleavage is very likely caused by the formation of the side product resulting from ligation of intron termini to the 86mer product 5′IN-3′IN as shown in Figures 4F and 5A (lane F), which suppresses exon hybridisation and ligation. As discussed above, in competition to exon binding, hybridisation of the intron ends to the exon binding site is enthalpically disfavoured, but has a lower entropic cost. Therefore, as an intrinsic property of the system, intron end ligation is always competing with exon-exon ligation and cannot be suppressed completely.

Self-splicing hairpin ribozyme

In the full-length system (Figure 4G), a slight increase of 5′HPR cleavage and a decrease of intron end ligation (Figure 5A and B, lane/column G) is observed. Comparing scenarios depicted in Figure 4 E–G, the yield of ligated exons remains stable at approximately 8% (Figure 5B, columns E–G). Using higher temperatures (50 and 60°C) for the first 120 min of reaction, 5′HPR cleavage increased somewhat (SI, Supplementary Figure S2), but the generated 2′,3′-cP is less stable at those conditions. This is mirrored in decreased yields for both, exon ligation (to 5′EX-3′EX) and intron ligation (5′IN-3′IN). In addition, we tested 8% PEG4000 as a molecular crowding agent (23) with the purpose of improving hybridisation and structural stability, and consequently ligation. However, PEG was found to have only a small effect, although the yield of exon ligation increased from 8% to 11% at 40°C (SI, Supplementary Figure S2). The effect of PEG is generally stronger at higher temperatures due to counteraction of the higher fragment dissociation rates, which is particularly visible in the splicing assay with fragmented RNA species (SI, Supplementary Figure S2).

Prove of correct exon–exon ligation by sequencing

To verify the result of the self-splicing reaction, we decided to isolate the product RNA from a preparative reaction and to confirm the identity of the splice product by sequencing. Thus, a preparative splice reaction was conducted with 2 nmol of each ribozyme strand 5′HPR and 3′HPR. Reaction products were separated via denaturing PAGE and the band at the height of a 25 nt reference RNA, assigned to the ligated exons 5′EX-3′EX was isolated and eluted from the gel (SI, Supplementary Figure S6). The isolated RNA was reverse transcribed utilizing an extended RT primer that carries the complementary sequence of 3′-EX (no overlap with 5′EX) and on its 5′end the sequence of the M13 reversed sequencing primer (M13R, Figure 5C). The generated cDNA was transferred to a PCR reaction with another extended primer carrying the sequence of 5′EX (no overlap with 3′EX), and on the 5′-end the sequence of the M13 forward sequencing primer (M13F). Thus, if splicing occurred as predicted, the generated dsDNA would consist of the 5′EX-3′EX sequence flanked by the two M13 primers (Figure 5C). For the analysis of such a short DNA by Sanger sequencing, we used SEQplus modified M13 forward and reverse primer, which facilitate the defined detection of the first base straight away. Both the forward and reverse sequencing resulted in sequences correspond-

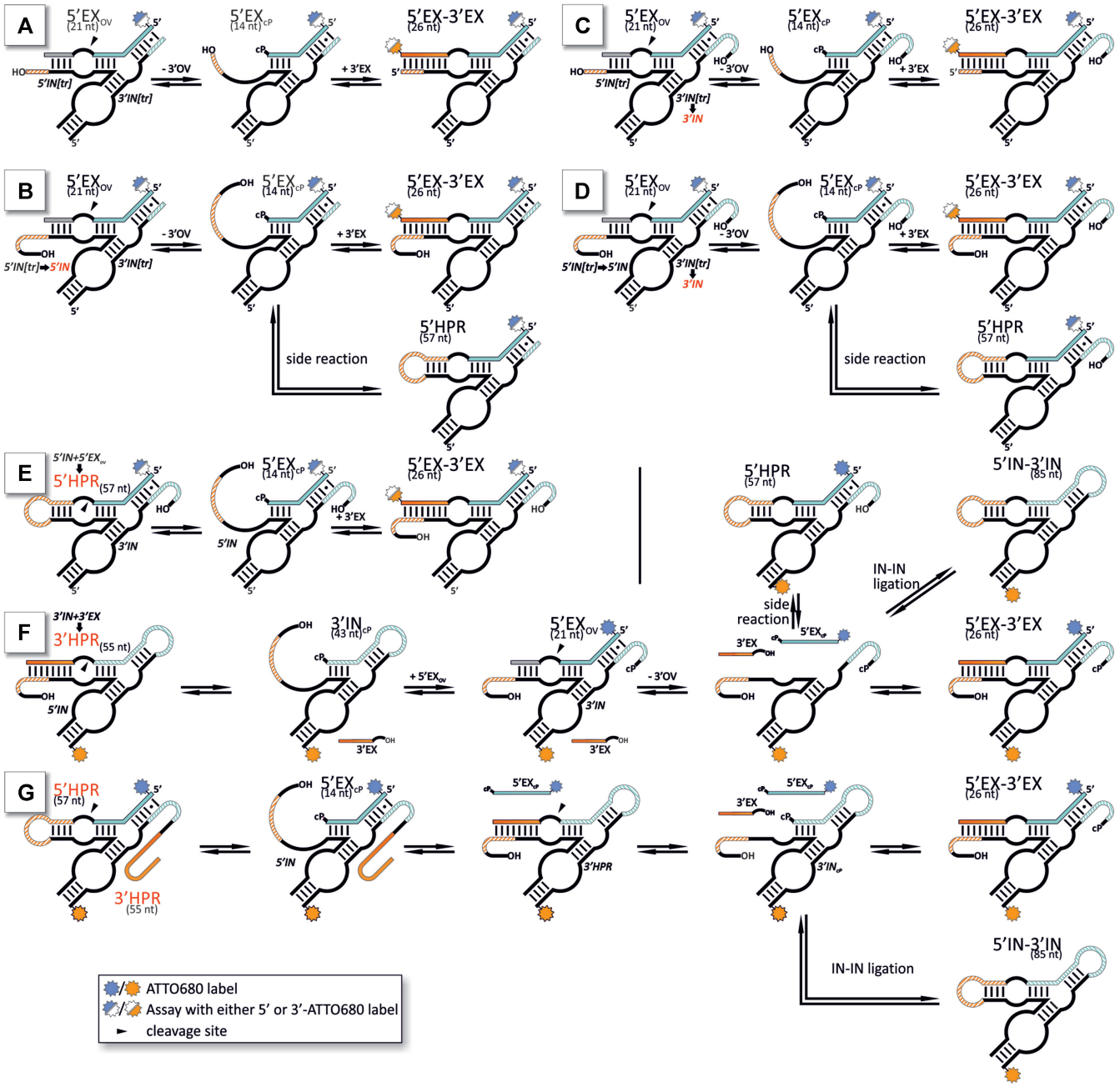

Figure 4. Splice assays with combinations of different truncated and full-length RNA strands. (**A**) 5′EX$_{OV}$ + 3′EX + 5′IN[tr] + 3′IN[tr]. (**B**) as A, but 5′IN[tr] → 5′IN. (**C**) as **A**, but 3′IN[tr] → 3′IN. (**D**) as A, but 5′IN[tr] + 3′IN[tr] → 5′IN + 3′IN. (**E**) as D, but 5′IN + 5′EX → 5′HPR. (**F**) as D, but 3′IN + 3′EX → 3′HPR. (**G**) 3′HPR + 5′HPR. Reaction conditions: 50 mM Tris/HCl (pH 7.5), 50 mM $MgCl_2$, 50 nM of each RNA species, 240 min, 40°C. Reactions were carried out with ATTO680 labelled 5′EX$_{OV}$ or 5′HPR and repeated with ATTO680 labelled 3′EX or 3′HPR as control. Yields of reactions shown in A to G and a corresponding gel image are provided in Figure 5A and B.

ing to the expected 5′EX-3′EX construct (Figure 5C). Although the primers used for cDNA synthesis and PCR contain the sequences of the 5′ and 3′ exon, the achieved cDNA synthesis and amplification together with the result of the sequence analysis confirm the successful self-splicing and exon-exon ligation of the designed HPR system in the expected manner. Taken together, self-splicing of the designed hairpin ribozyme variant has been unambiguously demonstrated.

Computer-aided approach for rational design of an alternative self-splicing hairpin ribozyme

Further improvement of ligation yield by redesign of the sequence and screening of reaction conditions may be possible. To this end, we decided to not continue with design by hand, but to verify our approach by using a computer-aided algorithm for the prediction of self-splicing hairpin ribozyme sequences. NUPACK Design (24,25) is a web soft-

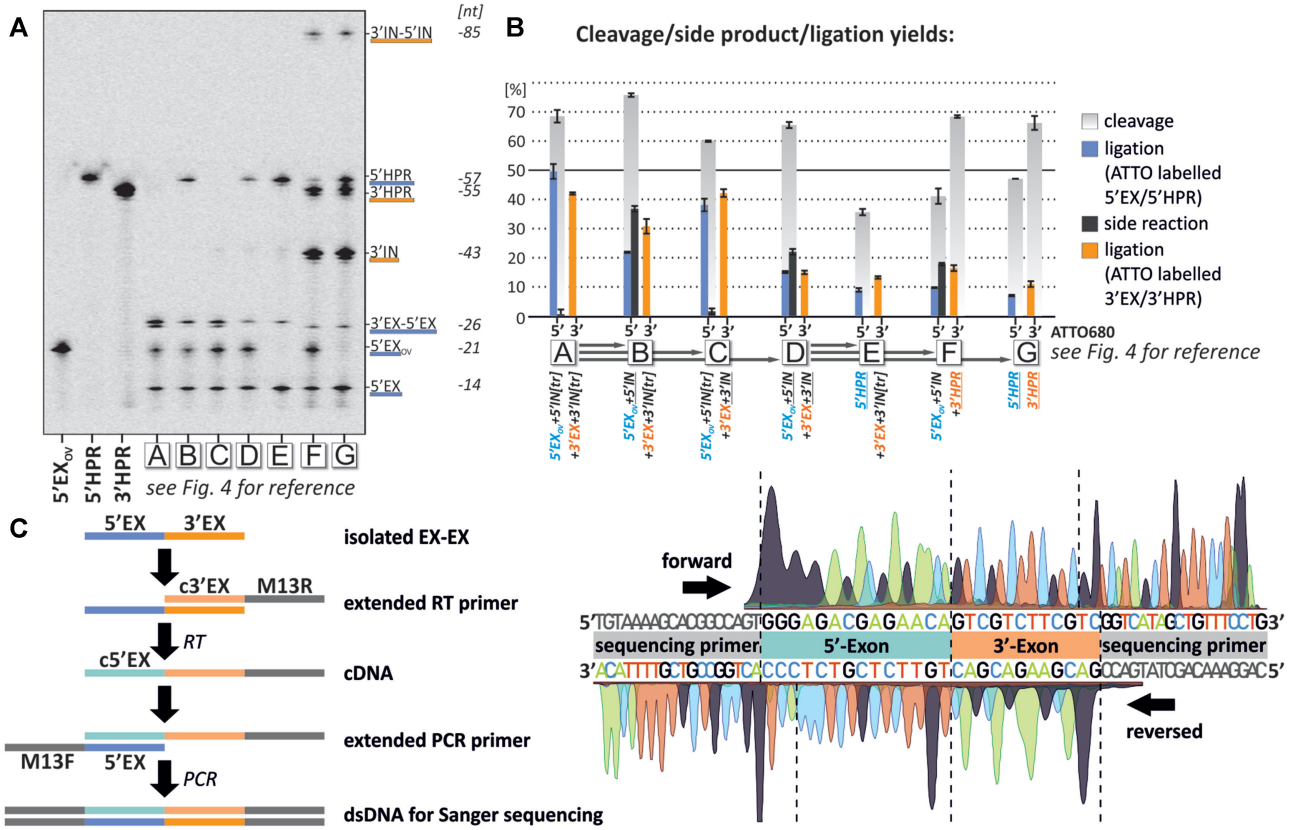

Figure 5. Analysis of splice reactions. (**A**) Denaturing PAGE (8%) of the splice assays shown in Figure 4A–G. Visualization by fluorescence imaging (ATTO680 labelled 5′EX/5′HPR/3′HPR) on a LI-COR4300 DNA Sequencer. (**B**) Reaction yields corresponding to assays shown in Figure 4A–G. 5′-bar (blue): data obtained with ATTO680 labelled 5′-EX_{OV} or 5′HPR. 3′-bar (orange): data obtained with ATTO680 labelled 3′EX or 3′IN or 3′HPR. Analyses were performed in triplicate measurement. (**C**) Sequencing analysis of the isolated splice product 5′EX-3′EX. RNA corresponding to the assigned ligated exons was isolated from a preparative splice reaction and reverse transcribed into cDNA using an extended RT primer with a binding site for a sequencing primer. To integrate another primer binding site on the opposite end, a PCR reaction was executed, generating a dsDNA composed of the 5′EX-3′EX sequence in the centre flanked by binding sites for sequencing primers. Specific (SEQplus) forward and reversed M13 sequencing primers were used to allow for analysis of the rather short DNA by Sanger sequencing.

ware application that, among other things, enables the computation of oligonucleotide sequences that are forced to fold into more than one defined secondary structure. We created a short script to compute a series of RNA sequences that match our defined rules for a self-splicing hairpin ribozyme (SI, S4, Supplementary Figure S4). We introduced the nomenclature SH-x (Self-splicing Hairpin ribozymes, x = number of individual variant). The hairpin ribozyme variant designed by hand and discussed above was included into the nomenclature and given the name SH-I. Accordingly, the five computed variants were named SH-II to SH-VI (SI, S5, Supplementary Figure S5). Two of the computed hairpin ribozyme structures were synthesised and experimentally tested, one of the two (SH-II) is exemplary presented here, the other one (SH-V) turned out being incapable of self-splicing (SI, S5, Supplementary Figure S5). The selection of SH-II and SH-V out of the five computed variants resulted from a short survey of the suggested secondary structures, focussing on the relation of properly folded structures for exon-exon ligation to predicted falsely-folded structures.

The novel self-splicing hairpin ribozyme SH-II (Figure 6A) has only 28% sequence similarity to the system designed by hand described above (SH-I). To challenge the algorithm, we reduced the length of helix 1, 4 and 5 by one base pair each. As for the splice assay discussed above, we used truncated introns (5′/3′IN-II[tr]) to evaluate the performance of the system in first experiments. The 5′-exon again was extended to carry a sequence overhang (5′EX-II_{Ov}) to allow cleavage and generation of the 3′-terminal cP *in situ*. We tested different temperatures for cleavage of 5′EX-II_{Ov} in the absence of the 3′-exon (3′EX-II). At 40°C, we observed the highest cleavage yield (70% Figure 6B). Interestingly, with 10°C, the optimal temperature for exon-exon ligation in SH-II is much lower than expected (SI, Supplementary Figure S3), in comparison to SH-I. The lower G–C content at the exon binding sites likely cause weaker binding of exons for ligation, requiring compensation with lower reaction temperatures. The highest ligation yields for 5′EX-II-3′EX-II were 20% in the absence of and 24% in the presence of PEG4000 (Figure 6B), which is only half of the activity of the SH-I system. Interestingly, this difference is neglected in the full-length system of SH-II, where exon-exon ligation proceeds with the same efficiency (8% without and 11% with PEG, Figure 6B) as observed with SH-I (Figure 5B, column C).

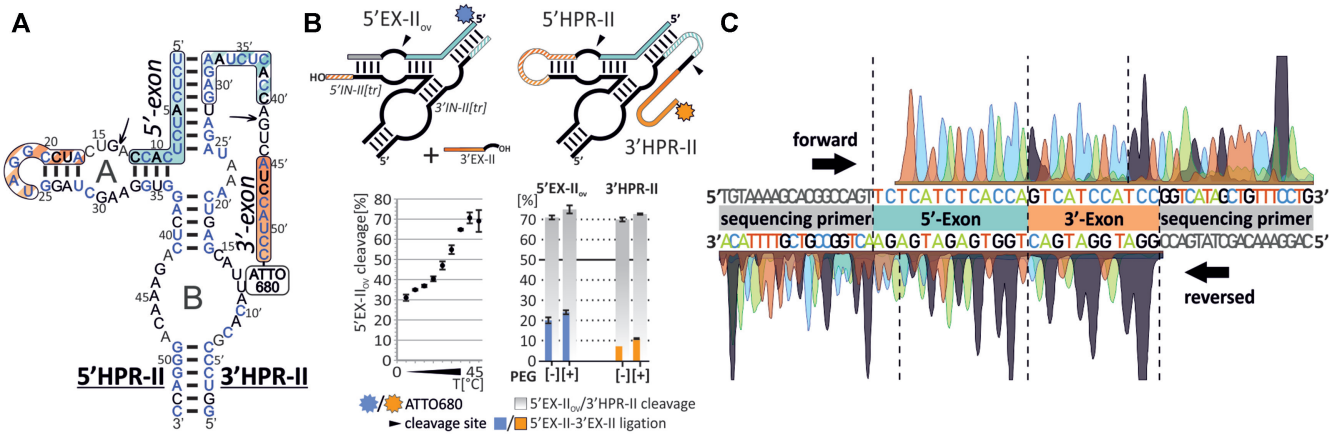

Figure 6. (**A**) Self-splicing hairpin ribozyme SH-II. Sequence computed by NUPACK. Blue bases indicate a substitution in comparison to SH-I (Figure 2). (**B**) Top: Schemes for splice assays (5′EX-II$_{OV}$ with 3′-overhang to generate the 3′-terminal cP *in situ*). Bottom left: cleavage of 5′EX-II$_{OV}$ (in the absence of 3′EX-II) at different temperatures. Bottom right: splice reactions with the exon fragments and the truncated introns in comparison to the full-length ribozyme strands 5′HPR-II and 3′HPR-II (full assay in SI, Supplementary Figure S3). Analyses were performed in triplicate measurement. (**C**) Sequencing data obtained for the isolated splice product (Sanger sequencing of reverse transcribed 5′EX-II-3′EX-II, see SI, Supplementary Figure S6).

To again prove the correct sequence of the splice product 5′EX-II-3′EX-II, we repeated the procedure for sequence analysis that was used for SH-I (Figure 5C), but this time with extension primers that were specific for the new splice product. Here, as well, both sequencing directions yielded in the correct sequences (Figure 6C), thus confirming successful self-splicing of SH-II.

CONCLUSION

In this study, we have presented two hairpin ribozyme variants capable of self-splicing, engineered using different design approaches. The designed hairpin ribozyme structures cut themselves out of an RNA molecule and support ligation of the resulting truncated terminal fragments. This shows that splice-like reactions can be supported by small RNAs with cleavage and ligation activity. Such splice-like reactions may have played a role in the early RNA world prior to evolution of more elaborate RNA scaffolds, as they would produce new phenotypes by restructuring segments of a given gene.

The self-cleaving hairpin ribozyme SH-I, which was designed by hand, shows exon–exon ligation yields of up to 51% in its core design with the exon fragments and truncated introns. With extension of the introns up to the full-length RNAs, exon exon ligation strongly competes with reverse reaction to the educts and intron ligation. The fully assembled hairpin ribozyme accomplishes the entire self-splicing reaction with yields up to 11% in the presence of 8% PEG4000. For design of SH-II as an alternative self-splicing hairpin ribozyme, a computer-aided approach based on the NUPACK software package, was used. We predefined the same basic ruleset as for SH-I, but challenged the algorithm by reducing the base pairs in the substrate binding regions. This system showed in its full-length variant exon-exon ligation yields up to 11%, similar to SH-I. The obtained yields for ligated exons may not be particularly high, however, with regard to the RNA world may be sufficient to demonstrate RNA processing dynamics and to allow sequence variability to occur. The two successful operating systems demonstrate the multiplicity that the variable sequence of the hairpin ribozyme provides for rational design. Two hairpin ribozyme variants of highly different sequence, predicted by hand (SH-I) or by computer-aided design (SH-II), catalyse the same cascade of cleavage and ligation reactions with similar results.

Apart from the implications for RNA processing in early life, (self-)splicing mediated by a small ribozyme structure may have potential for biotechnological application, i.e. for detection of a specific RNA sequence required to complement the splicing competent structure, or for intracellular generation of a circular RNA (the spliced out intron). For the latter, cells might be transfected with a plasmid encoding the self-splicing RNA, which upon transcription would perform self-splicing and intron circularization, as discussed previously (26–28).

In general, the work presented here, demonstrates once more that the hairpin ribozyme with its intrinsic cleavage-ligation characteristics is ideally suited for the design of RNA catalysts supporting important RNA processing reactions. Thus, the self-splicing hairpin ribozyme variants shown herein add nicely to previously introduced hairpin ribozymes mediating RNA repair, recombination, circularization and oligomerization (12,29).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

A scholarship from the *Landesgraduiertenförderung der Universität Greifswald*, awarded to Jikang Zhu is gratefully acknowledged.

FUNDING

The open access publication charge for this paper has been waived by Oxford University Press – *NAR* Editorial Board

members are entitled to one free paper per year in recognition of their work on behalf of the journal.
Conflict of interest statement. None declared.

REFERENCES

1. Smathers,C.M. and Robart,A.R. (2019) The mechanism of splicing as told by group II introns: ancestors of the spliceosome. *Biochim. Biophys. Acta (BBA) - Gene Regul. Mech.*, **1862**, 194390.
2. Penny,D., Hoeppner,M.P., Poole,A.M. and Jeffares,D.C. (2009) An overview of the introns-first theory. *J Mol Evol*, **69**, 527–540.
3. Koonin,E.V., Senkevich,T.G. and Dolja,V.V. (2006) The ancient virus world and evolution of cells. *Biol. Direct.*, **1**, 29.
4. Black,D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem.*, **72**, 291–336.
5. Castle,J.C., Zhang,C., Shah,J.K., Kulkarni,A.V., Kalsotra,A., Cooper,T.A. and Johnson,J.M. (2008) Expression of 24,426 human alternative splicing events and predicted cis regulation in 48 tissues and cell lines. *Nat. Genet.*, **40**, 1416–1425.
6. Fedorova,L. and Fedorov,A. (2003) Introns in gene evolution. *Genetica*, **118**, 123–131.
7. Patthy,L. (1999) Genome evolution and the evolution of exon-shuffling — a review. *Gene*, **238**, 103–114.
8. Woodson,S.A. and Cech,T.R. (1989) Reverse self-splicing of the tetrahymena group I intron: implication for the directionality of splicing and for intron transposition. *Cell*, **57**, 335–345.
9. Rearick,D., Prakash,A., McSweeny,A., Shepard,S.S., Fedorova,L. and Fedorov,A. (2011) Critical association of ncRNA with introns. *Nucleic Acids Res*, **39**, 2357–2366.
10. Brehm,S.L. and Cech,T.R. (1983) The fate of an intervening sequence RNA: excision and cyclization of the Tetrahymena ribosomal RNA intervening sequence in vivo. *Biochemistry*, **22**, 2390–2397.
11. Hieronymus,R., Godehard,S.P., Balke,D. and Müller,S. (2016) Hairpin ribozyme mediated RNA recombination. *Chem. Commun.*, **52**, 4365–4368.
12. Hieronymus,R. and Müller,S. (2021) Towards higher complexity in the RNA world – hairpin ribozyme supported RNA recombination. *ChemSystemsChem*, **1447**, 135–143.
13. Riley,C.A. and Lehman,N. (2003) Generalized RNA-directed recombination of RNA. *Chemistry & Biology*, **10**, 1233–1243.
14. Hayden,E.J., Riley,C.A., Burton,A.S. and Lehman,N. (2005) RNA-directed construction of structurally complex and active ligase ribozymes through recombination. *RNA*, **11**, 1678–1687.
15. Jimenez,R.M., Polanco,J.A. and Lupták,A. (2015) Chemistry and biology of self-cleaving ribozymes. *Trends Biochem. Sci.*, **40**, 648–661.
16. Strobel,S.A. and Cochrane,J.C. (2007) RNA catalysis: ribozymes, ribosomes, and riboswitches. *Curr. Opin. Chem. Biol.*, **11**, 636–643.
17. Mörl,M., Lizano,E., Willkomm,D.K. and Hartmann,R.K. (2008) Production of RNAs with Homogeneous 5′ and 3′ Ends. In *Handbook of RNA Biochemistry*. John Wiley & Sons, Ltd, pp. 22–35.
18. Zarzosa-Alvarez,A.L., Sandoval-Cabrera,A., Torres-Huerta,A.L. and Bermudez-Cruz,R.M. (2010) Electroeluting DNA fragments. *J. Vis. Exp.*, https://doi.org/10.3791/2136.
19. Hegg,L.A. and Fedor,M.J. (1995) Kinetics and thermodynamics of intermolecular catalysis by hairpin ribozymes. *Biochemistry*, **34**, 15813–15828.
20. Berzal-Herranz,A., Joseph,S., Chowrira,B. m., Butcher,S. e. and Burke,J. m. (1993) Essential nucleotide sequences and secondary structure elements of the hairpin ribozyme. *EMBO J.*, **12**, 2567–2573.
21. Reuter,J.S. and Mathews,D.H. (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics*, **11**, 129.
22. Pérez-Ruiz,M., Barroso-delJesus,A. and Berzal-Herranz,A. (1999) Specificity of the Hairpin Ribozyme Sequence requirements surrounding the cleavage site. *J. Biol. Chem.*, **274**, 29376–29380.
23. Paudel,B.P. and Rueda,D. (2014) Molecular crowding accelerates ribozyme docking and catalysis. *J. Am. Chem. Soc.*, **136**, 16700–16703.
24. Wolfe,B.R., Porubsky,N.J., Zadeh,J.N., Dirks,R.M. and Pierce,N.A. (2017) Constrained multistate sequence design for nucleic acid reaction pathway engineering. *J. Am. Chem. Soc.*, **139**, 3134–3144.
25. Zadeh,J.N., Steenberg,C.D., Bois,J.S., Wolfe,B.R., Pierce,M.B., Khan,A.R., Dirks,R.M. and Pierce,N.A. (2011) NUPACK: analysis and design of nucleic acid systems. *J. Comput. Chem.*, **32**, 170–173.
26. Petkovic,S., Badelt,S., Block,S., Flamm,C., Delcea,M., Hofacker,I. and Müller,S. (2015) Sequence-controlled RNA self-processing: computational design, biochemical analysis, and visualization by AFM. RNA, **21**, 1249–1260.
27. Petkovic,S. and Müller,S. (2013) RNA self-processing: formation of cyclic species and concatemers from a small engineered RNA. *FEBS Letters*, **587**, 2435–2440.
28. Pieper,S., Vauléon,S. and Müller,S. (2007) RNA self-processing towards changed topology and sequence oligomerization. *Biol. Chem.*, **388**, 743–746.
29. Hieronymus,R. and Müller,S. (2019) Engineering of hairpin ribozyme variants for RNA recombination and splicing. *Ann. N. Y. Acad. Sci.*, **1447**, 135–143.