Towards Higher Complexity in the RNA World: Hairpin Ribozyme Supported RNA Recombination

Robert Hieronymus and Sabine Müller*[a]

In the RNA world, the exchange of sequence patches between two RNAs is an intriguing evolutionary concept, allowing generation of new RNA molecules with novel functionality. Based on the hairpin ribozyme (HPR) with its unique cleavage-ligation properties, we here demonstrate RNA supported RNA recombination as a possible scenario for the emergence of larger RNA molecules with more complex functionality. A HPR variant designed for the purpose of recombination is capable of cleaving two different RNA molecules, one being a hammerhead ribozyme (HHR) and the other an aptamer (A), and to subsequently recombine and ligate the resulting fragments to a hammerhead ribozyme that is allosterically controlled (HHA) by a cognate ligand. Two such recombination processes involving aptamers for either theophylline or flavine mononucleotide (FMN) are demonstrated with yields of functional recombination product of up to 34%.

The RNA world[1,2] is a hypothetical period at the origin of life, where RNA has acted as the dominant biomolecule performing both tasks necessary for life: information storage and catalysis of chemical reactions. It will be never known to the end where and when the RNA world emerged or how it really took place, due to the vague evidence there is on earth today about the time around four billion years ago. Therefore, the design and studying of plausible model systems is a reasonable approach towards deeper understanding of processes at the origin of life.

We recently demonstrated such a model system that provides insights into possible early prebiotic RNA recombination.[3] Recombination is a significant source for generation of genetic variation within a population, and hence, empowers it to evolution. It is also a repair mechanism especially in asexual populations where deleterious mutations tend to accumulate and would lead to extinction if not counteracted.[4] Hence, it is plausible to hypothesize that a mechanism for recombination has also existed in the RNA world. Accordingly, RNA recombination has been studied in a RNA world context,[5–8] and it has been shown that recombination is useful for evolving new ribozyme functions from pre-existing sequences, something that may have played an important role early in evolution.[9] As shown by us and others, the hairpin ribozyme is a perfect candidate for the role of an RNA recombinase.[3,10] However, there is also evidence for ribozyme-independent spontaneous recombination in random RNA pools.[11,12]

The previously in our laboratory designed recombination system consists of a single hairpin ribozyme (HPR) variant that is capable of binding and cleaving two non-functional RNA molecules in separate reactions, followed by recombination and ligation of the resulting fragments to a new and fully functional RNA.[3] This generated RNA folds into a hammerhead ribozyme (HHR) in the presence of an added RNA substrate, and the following cleavage of this added substrate (HHRS) in trans verified the functionality of the recombination product. The recombination took place in a one-pot reaction with yields of up to 74%, thus being equally efficient as RNA recombination mediated by the much larger *azoarcus* group I ribozyme variant introduced previously by Lehman *et al.*.[5,6]

In the present work, we have addressed the question if we can further extend this recombination system to generate a more complex hammerhead ribozyme variant out of the HHR obtained in the first recombination.[3] Thus, the entire process would mimic an evolutionary reaction cascade that in a first recombination step processes non-functional RNAs to a functional RNA, such as a simple ribozyme, and then, in a second new recombination step, processes the simple ribozyme to a more complex RNA molecule, such as an aptazyme (Figure 1A). Aptazymes are catalytic RNA structures, where the ribozyme domain is linked via a communication module with an aptamer domain.[13,14] The aptamer binds a ligand with high affinity and selectivity. This binding event leads to refolding of the aptamer, and consequently, transmitted via the communication module, to refolding of the ribozyme domain. This way, ribozyme activity can be up or down regulated, depending on the specific design and the applied communication module. Looking to the RNA world, at some point such allosteric regulation of ribozymes would have been indispensable for an RNA organism to properly control its metabolism.

For the rational design approach of the new recombination system we again chose the hairpin ribozyme as the catalyst supporting recombination. The HPR fits well for the purpose of rational design, since it is small and has only a few conserved nucleotides, and, important for recombination, is both a good endonuclease and a ligase.[15] Given that the HHR is the

[a] R. Hieronymus, Prof. Dr. S. Müller
Institute for Biochemistry
University Greifswald
Felix-Hausdorff-Str. 4, 17487 Greifswald (Germany)
E-mail: smueller@uni-greifswald.de

Supporting information for this article is available on the WWW under https://doi.org/10.1002/syst.202100003

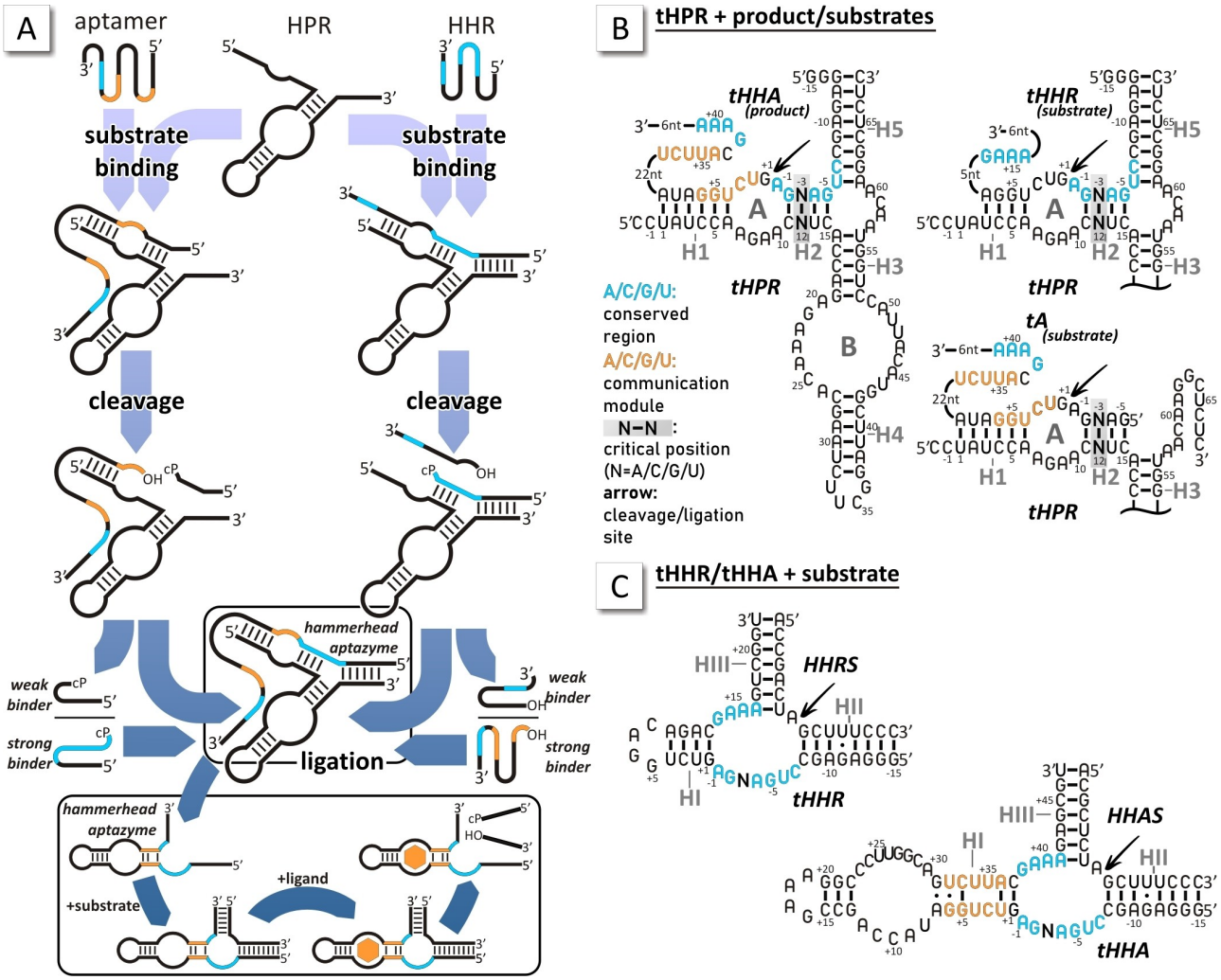


Figure 1. A) Schematic of the recombination system and functionality test of the product (box) in a one-pot reaction. cP: 2′,3′-cyclic phosphate. **B)** Sequences and secondary structures of the HPR variant for recombination in the theophylline system. tHPR bound to the recombination product tHHA and to the two initial substrates tHHR and tA. **C)** tHHR and tHHA variants with their respective substrate (HHRS/HHAS).

functional RNA product of the first recombination event,[9] we sought for hammerhead aptazymes (HHA) that could be generated from this HHR in a further recombination event. The reports on allosteric HHRs from the Breaker lab[13,16–19] were a rich pool of candidates for that endeavor. We first selected the theophylline regulated cm^{+}theo3[18] as an appropriate HHA to be produced by recombination and designed the individual ribozyme and substrate sequences (Figure 1B, C). The rational design process is described in detail in the supporting information (SI, S3.3).

Both initial substrates, aptamer and HHR (Figure 1A) carry different patches needed for HHA generation. The driving force for recombination as depicted in Figure 1A is strand displacement, i.e. discarding of the nonrequired sequence patches and binding of the two functional sequence patches by one HPR. Therefore, both substrates are designed to have strong and weak binding regions, up and down-stream from the cleavage site. Hence, after cleavage, the weaker bound fragments will be exchanged for the stronger binding ones (see also Figure 1A). This design is not only supportive to the recombination of fragments, but is also favorable in terms of hairpin ribozyme activity. Weakly bound substrates would be preferentially cleaved, while ligation would be the preferred reaction for stronger bound substrates.[15]

Our design approach resulted in RNA sequences for the hairpin ribozyme (tHPR), the hammerhead ribozyme (tHHR), the aptamer (tA) and the hammerhead aptazyme (tHHA) in the theophylline system (therefore denoted as "t") (Figure 1B). Furthermore, two substrates to be cleaved by tHHR (substrate HHRS) or by tHHA (substrate HHAS) were designed (Figure 1C).

The design has undergone some alterations up to the final sequences depicted in Figure 1B and C (described in detail in SI, S3.4). For example, our original designs, though satisfying in cleavage, never really showed ligation activity. Evaluation of the HPR-substrate complexes showed that the originally chosen sequence of helix 2 abolished ligation. Changing this sequence back to wild type resulted in a recombination yield of 33% (Figure S6).

The problem with helix 2 is, that upon substrate binding it integrates the conserved regions of tHHR and tHHA (blue sequence next to the cleavage site in Figure 1B), and thus the sequence is fixed. However, there is one exception, U-3 (designated as **N** in Figure 1B and C), being the only not fully conserved nucleobase in this region.[20] To investigate the sequence variability at this site, we prepared tHPR|tHHR/tA mutants A12U|U-3A, A12G|U-3C and A12C|U-3G and investigated their cleavage and ligation properties (Figure 2A). For this assay, we used the 3′-fragment of tA (tA-3′F), since pre-cleavage of tA is not relevant for ligation in this assay, thus making it simpler. To even further simplify, tA-3′F was additionally truncated at its 3′-terminus by 36 nt (tA-3′F[tr]), to ensure proper binding to tHPR, but avoid any disturbances that may be caused by the overhanging aptamer sequence (Figure 2A). The tHHR substrate was kept in full length and cleaved by tHPR to produce the 2′,3′-cyclic phosphate carrying 5′-fragment *in situ*, which then was directly ligated with tA-3′-F[tr], thus mimicking key steps of our final recombination scenario.

In separate experiments, we examined the cleavage of HHRS by tHHR, mutated at N-3 (Figure 2B), and the cleavage of full-length tA variants (varying at N-3) by the different tHPR mutants (Figure 2C).

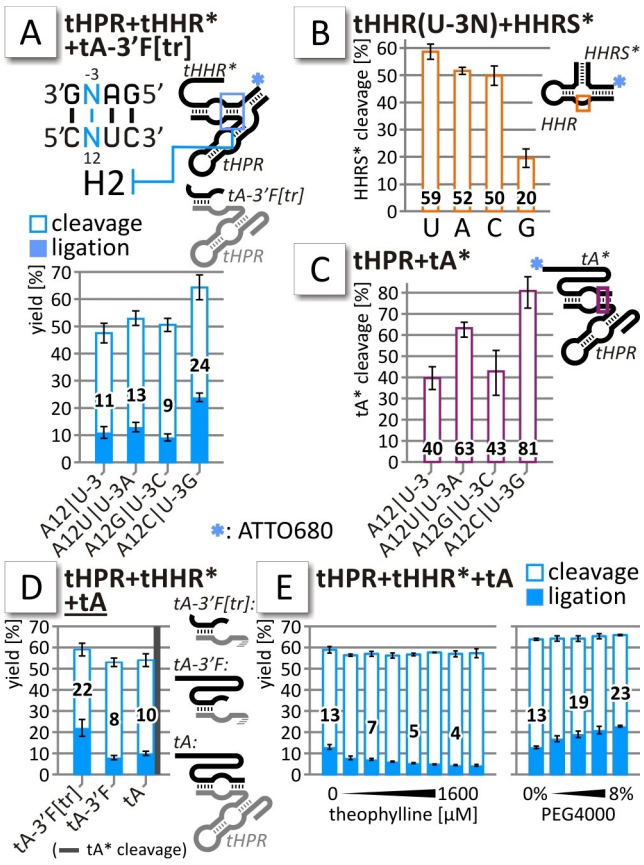


Figure 2. A) Mutation assay with tHPR(A12N) + tHHR(U-3N)* + tA-3′F[tr]. **B)** Cleavage assay tHHR(U-3N)* + HHRS. **C)** Cleavage assay tHPR(A12N) + tA(U-3N)*. **D)** Recombination with tHPR(A12C), tHHR(U-3G) and different variants of tA: truncated fragment (3′F[tr]), full-length fragment (3′F) and full-length. The cleavage of tA* (similar experiment) is additionally indicated by the grey column. **E)** Effect of theophylline and PEG on recombination.

Among the helix 2 variants, the mutation of A12C|U-3G appears to be the most promising, with 24% recombination yield (Figure 2A). All other mutants show clearly lower cleavage and ligation activity. Cleavage of HHRS by tHHR(U-3G) proceeds with the lowest yield among the four variants (20%, Figure 2B), and thus it is unlikely that aptazyme tHHA(U-3G) obtained from recombination, will reach activity above this level. However, this fact is well acceptable, since it is a downstream process of the recombination, required just for qualitative evaluation of the recombination products functionality. More importantly, tA(U-3G) is cleaved very efficiently in the presence of tHPR(A12C) (81%, Figure 2C). This way a sufficient amount of fragments is provided as substrates for recombination. This, together with the satisfying ligation yield shown in Figure 2A makes the A12C|U-3G mutant the favorable variant for our recombination system.

Integration of the full-length tA substrate in the recombination system was the most challenging step. Exchanging the truncated 3′-fragment of tA (tA-3′F[tr], 11 nt) for a non-truncated 3′-fragment (tA-3′F, 47 nt) caused a drop in ligation yield from 22% to 8%, which stayed on a similar level when the full-length tA(53 nt) was used (Figure 2D). Interestingly, cleavage of the full-length tA by tHPR appears to proceed efficiently (comp. Figure 2C and D). Yet, the 3′-overhang constituting the aptamer region seems to have a tremendous effect on substrate binding stability, and thus on recombination yield. Obviously, dissociation of the long structured 3′-fragment of tA upon cleavage is rather fast, thus supporting cleavage but being unfavorable for ligation. Furthermore, the addition of theophylline disturbs the system supplementary (Figure 2E), due to the fact that the ligand binding region of the aptamer sequence reaches into the HPR binding site (Figure 1B). Hence, we thought of stabilization of substrate binding to tHPR by addition of PEG4000 as a molecular crowding agent.[21] To our satisfaction, in the presence of 8% PEG4000, recombination was boosted to 23% (Figure 2E) and even reached 32% in the time-resolved assay (Figure 3B).

After having shown successful recombination, the assay still had to be completed by demonstrating the functionality of the recombination product. In the presence of theophylline, tHHA mediated cleavage of HHAS was clearly observed (Figure 3A lane 5). Neither tHPR alone nor tHHR together with tA was capable of initiating cleavage of HHAS (lane 6 and 7). Hence, recombination was absolutely required to deliver the fully functional aptazyme tHHA.

Cleavage of HHAS proceeds best at 35 °C, which however is an unfavorable condition for the recombination product tHHA, because it is back-cleaved by the hairpin ribozyme tHPR, still present in the recombination mix (as seen in Figure 3A lane 4 and 5). We therefore isolated the recombination product tHHA from recombination reaction at preparative scale and incubated it separately with the substrate HHAS in the absence or presence of theophylline (lane 10 and 11). For comparison, we carried out the same reaction with an *in vitro* transcribed version of tHHA(U-3G) (Figure 3A, lane 8 and 9).

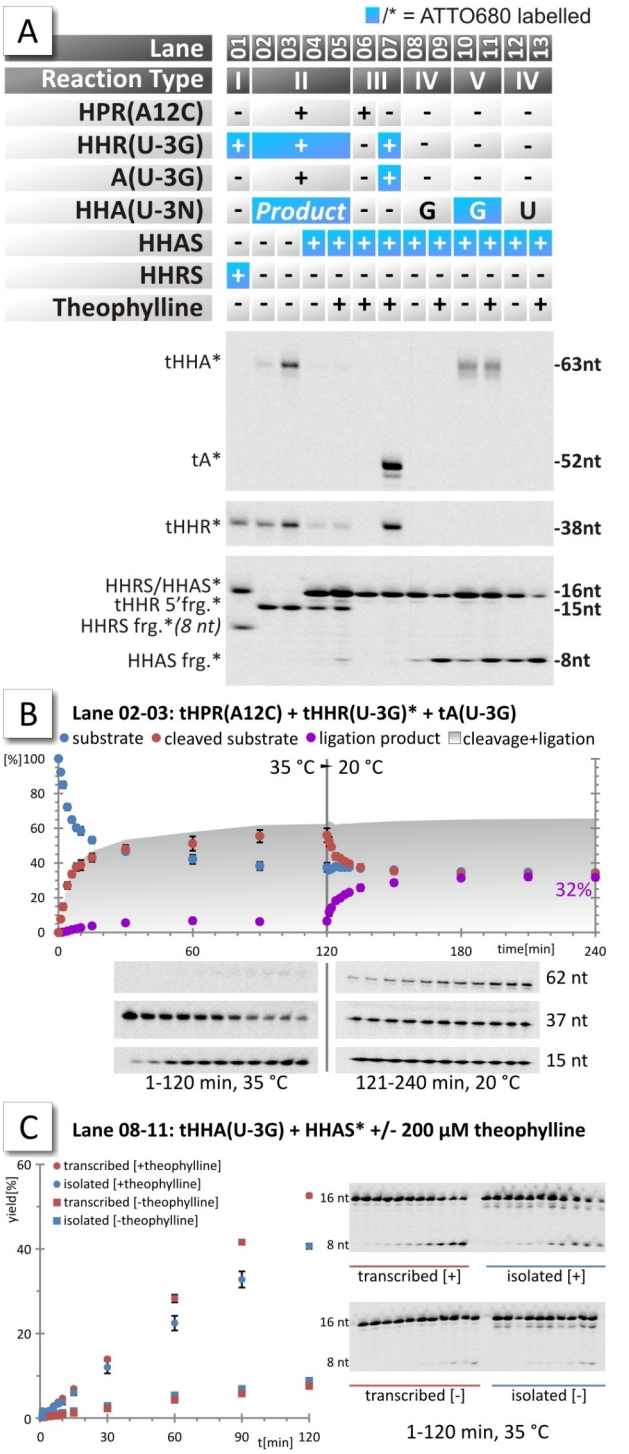


Figure 3. A) Denaturing PAGE of a collection of recombination and cleavage reactions analyzed via fluorescence detection on a LI–COR 4300 DNA Sequencer. **I:** Cleavage assay tHHR* + HHRS* (120 min at 35 °C). **II:** Recombination assay (lane 2: 120 min at 35 °C; lane 3: lane 2 + 120 min at 20 °C; lane 4/5: lane 3 + HHAS−/+ 200 µM theophylline + 120 min at 35 °C). **III:** Negative control for HHAS cleavage. **IV:** Cleavage assay tHHA (in vitro transcribed reference) + HHAS*. **V:** Cleavage assay tHHA* (isolated from recombination reaction) + HHAS*. **B)** Time course of recombination (lane 2–3 in A). **C)** Time course of HHAS* cleavage by the recombination product (lane 8–11 in A). Note: HHRS is 3′-labelled, HHAS 5′-labelled. Hence, the detectable 8 nt cleavage fragments carry a 5′-OH (HHRS frg.) or a 2′,3′-cP (HHAS frg.) and thus migrate differently through the gel. In addition, the 5′-dye is tethered via a phosphate linkage, which introduces a further negative charge.

The isolated recombination product shows similar, although somewhat weaker cleavage efficiency as the in vitro transcribed tHHA(U-3G) (Figure 3C), thus once more confirming the success of the recombination. The difference in the cleavage efficiency may be a result of the differently decorated 5′-termini of the in vitro transcribed and the isolated recombination product. The latter carries a 5′-terminal ATTO680 label resulting from the initial substrate tHHR for recombination (Figure 3A), which may hamper substrate binding and consequently make cleavage less efficient.

In order to demonstrate the universal character of hairpin ribozyme mediated RNA recombination, we designed another ensemble of RNA sequences for recombination. In particular, we exchanged the theophylline aptamer region with the flavine mononucleotide (FMN) aptamer to produce a FMN responsive hammerhead aptazyme by recombination. As a lead for this new HHA we used the FMN aptazyme with the "class VII inducing element" from the Breaker lab.[16] This FMN aptazyme shares the same position of the cleavage site as the theophylline responsive hammerhead aptazyme and differs only in the communication module and the aptamer sequence. Hence, we could use the sequence design of the previous system and just exchange the theophylline aptamer with the FMN aptamer sequence. This way, we designed and generated again a hairpin ribozyme as recombinase, a hammerhead ribozyme as one initial substrate, the FMN aptamer as the second initial substrate, and the FMN responsive hammerhead aptazyme as recombination product for the FMN system (denoted as "f"): fHPR, fHHR, fA and fHHA (Figure 4A). Substrates for functional testing remained the same as in the theophylline system: HHRS and HHAS.

In the new system, recombination yields of up to 29% after 120 min at 35 °C were achieved (Figure 4B), which demonstrates far more efficient recombination compared to the theophylline system, where a temperature shift and +120 min incubation were required (Figure 3B). This is probably due to the much shorter aptamer region and an additional G–C pair in helix 1. Even though adding only about 2 kcal/mol, this G–C pair might account for the much smaller effect observed when FMN (decrease of recombination yield from 29% to 24% at 1600 mM FMN; Figure 4B) or PEG4000 (increase from 29% to 34% at 8% PEG4000, Figure 4B and E), were added, due to the already more stable structure in the absence of FMN or PEG.

However, despite the satisfying recombination yields, we could not identify any cleavage of HHAS by the recombination product fHHA in the presence of FMN (Figure 4D, lane 17). Neither the isolated fHHA nor an in vitro transcribed variant showed activity (lane 19 and 21). This excluded any inhibitory interaction of fHHA with RNAs out of the recombination mix, but pointed to a problem with the chosen sequence of tHHA itself. While examining this further, we discovered that the U-3G mutation, taken over from the theophylline system, leads to an inactive form of fHHA, due to a strong inhibitory interaction of the catalytic core with the communication module (see SI, section S3.1). Accordingly, we tested other fHHA variants, namely: U-3C and U-3A. Both variants cleave the HHAS decently (lane 24–27), but these mutations transferred in the recombina-

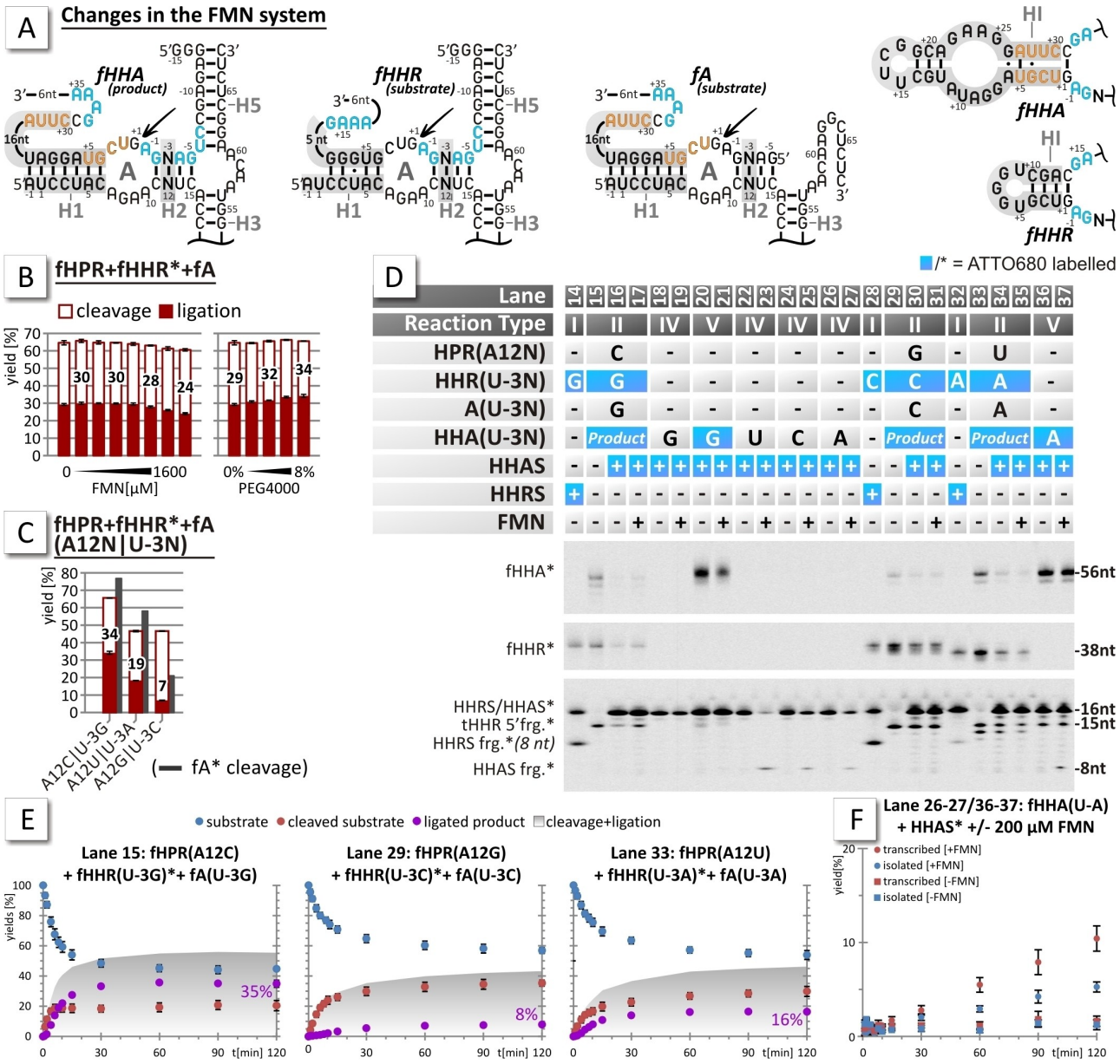


Figure 4. A) Secondary structures and sequences of RNAs of the FMN system (simplified; compare with Figure 1B and C). Changes in relation to the theophylline system are highlighted in grey. **B)** Recombination at different FMN or PEG4000 concentrations. **C)** Recombination yields for fHPR(A12N)+fHHR (U3N)+fA(U3N) (8% PEG4000; 0 µM FMN). Cleavage of fA* (similar experiment) is additionally indicated by the grey bars (full assay in the Supporting Information, S3.2). **D)** Denaturing PAGE of a collection of recombination and cleavage reactions analyzed via fluorescence detection on a LI–COR 4300 DNA Sequencer. **I:** Cleavage assay fHHR(U3N)*+HHRS* (120 min at 35 °C). **II:** Recombination assay with A12N|U-3N (lane 15/29/33: 120 min at 35 °C; lane 16/30/34 and 17/31/35: lane 15/29/33+HHAS+/−200 µM FMN+120 min at 35 °C). **IV:** Cleavage assay fHHA(U-3N) (in vitro transcribed references)+HHAS*. **V:** Cleavage assay fHHA(U-3N)* (isolated from recombination reaction)+HHAS*. **E)** Time course of recombination for different mutants. **F)** Time course of HHAS* cleavage by the recombination product fHHA(U-3A). Note: HHRS is 3′-labelled, HHAS 5′-labelled. Hence, the detectable 8 nt cleavage fragments carry a 5′-OH (HHRS frg.) or a 2′,3′-cP (HHAS frg.) and thus migrate differently through the gel. In addition, the 5′-dye is tethered via a phosphate linkage, which introduces a further negative charge.

tion assay, hamper recombination (lane 29 and 33). The mutation A12G|U-3C only yields 7% recombination product (Figure 4C), due to strong inhibition of tA cleavage (grey bars in Figure 4C or Figure S3). Based on the fact that the *in vitro* transcribed fHHA(U-3C) also did not show satisfying cleavage of HHAS (lane 25), this variant was not further considered. With mutant A12U|U-3A, 19% recombination yield was achieved (Figure 4C). The in vitro transcribed fHHA(U-3A) cleaved the substrate HHAS (lane 27), however, we could not detect any HHAS cleavage by the recombination product fHHA(U-3A) in the recombination mix (lane 35). Again, the stable helix 1 might be the reason for that, causing product inhibition after recombination, because of the interaction between fHPR and fHHA outcompeting binding of HHAS to fHHA for cleavage.

Nevertheless, incubation of the isolated recombination product fHHA(U-3A) with HHAS resulted in cleavage rates similar to the transcribed fHHA(U-3A) variant (Figure 4F). In conclusion, also the FMN system generates a functional RNA product, though it suffers more from product inhibition compared to the theophylline system. In future experimental designs, product inhibition might be counteracted by applying freeze-thaw or dry-wet cycling to promote strand exchange and thus to observe HHA activity already in the recombination mix.

In our previous work, we demonstrated a RNA mediated RNA recombination system that generates functional RNA (a simple hammerhead ribozyme) out of two non-functional RNAs, by the action of only a single hairpin ribozyme that catalyzes each of the involved reaction steps (cleavage and ligation).[3] Here, we have further extended this scenario by the rational engineering of novel recombination systems, which use already functional RNAs (ribozymes and aptamers) as substrates. The recombination of sequence patches from ribozymes and aptamers yields a complex RNA molecule with new/extended functionality. We successfully demonstrated the functionality of such systems by merging HHRs with theophylline or FMN aptamers, resulting in the generation of HHAs, the activities of which are now allosterically regulated in response to the cognate ligand. The evolutionary generation of such aptazymes would be indispensable for a more progressed RNA world, where a more advanced RNA organism needed to regulate its metabolism. This study also once more demonstrates the versatility of the hairpin ribozyme, which owing to its unique cleavage/ligation characteristics is a superior tool for rational design of RNA supported RNA processing reactions for modelling RNA world scenarios,[22,23] and with potential application in molecular biology and medicine.[24–26]

Conflict of Interest

The authors declare no conflict of interest.

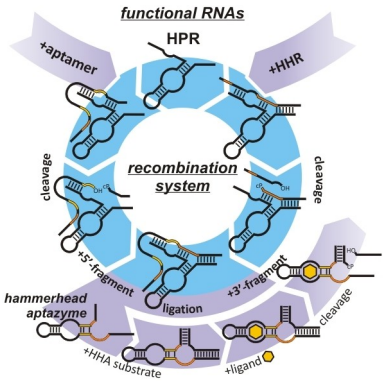
Keywords: aptazymes · ribozymes · RNA · RNA engineering · RNA recombination

[1] W. Gilbert, *Nature* **1986**, *319*, 618–618.
[2] A. Pressman, C. Blanco, I. A. Chen, *Curr. Biol.* **2015**, *25*, R953–R963.
[3] R. Hieronymus, S. P. Godehard, D. Balke, S. Müller, *Chem. Commun.* **2016**, *52*, 4365–4368.
[4] H. J. Muller, *Am. Nat.* **1966**, *100*, 493–517.
[5] C. A. Riley, N. Lehman, *Chem. Biol.* **2003**, *10*, 1233–1243.
[6] E. J. Hayden, C. A. Riley, A. S. Burton, N. Lehman, *RNA* **2005**, *11*, 1678–1687.
[7] E. J. Hayden, N. Lehman, *Chem. Biol.* **2006**, *13*, 909–918.
[8] N. Lehman, *J. Mol. Evol.* **2003**, *56*, 770–777.
[9] Q. S. Wang, P. J. Unrau, *RNA* **2005**, *11*, 404–411.
[10] H. Mutschler, A. Wochner, P. Holliger, *Nat. Chem.* **2015**, *7*, 502–508.
[11] H. Mutschler, A. I. Taylor, B. T. Porebski, A. Lightowlers, G. Houlihan, M. Abramov, P. Herdewijn, P. Holliger, *eLife* **2018**, *7*, e43022.
[12] B. A. Smail, B. E. Clifton, R. Mizuuchi, N. Lehman, *RNA* **2019**, *25*, 453–464.
[13] J. Tang, R. R. Breaker, *Chem. Biol.* **1997**, *4*, 453–459.
[14] M. Famulok, J. S. Hartig, G. Mayer, *Chem. Rev.* **2007**, *107*, 3715–3743.
[15] L. A. Hegg, M. J. Fedor, *Biochemistry* **1995**, *34*, 15813–15828.
[16] G. A. Soukup, R. R. Breaker, *Proc. Natl. Acad. Sci.* **1999**, *96*, 3584–3589.
[17] G. A. Soukup, R. R. Breaker, *Curr. Opin. Struct. Biol.* **2000**, *10*, 318–325.
[18] G. A. Soukup, G. A. M. Emilsson, R. R. Breaker, *J. Mol. Biol.* **2000**, *298*, 623–632.
[19] K. H. Link, L. Guo, T. D. Ames, L. Yen, R. C. Mulligan, R. R. Breaker, *Biol. Chem.* **2007**, *388*, 779–786.
[20] D. E. Ruffner, G. D. Stormo, O. C. Uhlenbeck, *Biochemistry* **1990**, *29*, 10695–10702.
[21] B. P. Paudel, D. Rueda, *J. Am. Chem. Soc.* **2014**, *136*, 16700–16703.
[22] S. Müller, *Ann. N. Y. Acad. Sci.* **2015**, *1341*, 54–60.
[23] R. Hieronymus, S. Müller, *Ann. N. Y. Acad. Sci.* **2019**, *1447*, 135–143.
[24] S. Vauléon, S. A. Ivanov, S. Gwiazda, S. Müller, *ChemBioChem* **2005**, *6*, 2158–2162.
[25] D. Balke, C. Wichert, B. Appel, S. Müller, *Appl. Microbiol. Biotechnol.* **2014**, *98*, 3389–3399.
[26] D. Balke, A. Becker, S. Müller, *Org. Biomol. Chem.* **2016**, *14*, 6729–6737.

Manuscript received: January 12, 2021
Accepted manuscript online: February 11, 2021
Version of record online: ■■, ■■■■

COMMUNICATIONS

RNA world: Recombination may have played a central role in the RNA world for the generation of RNA molecules with novel functions. A small catalytic RNA derived from the hairpin ribozyme acts as RNA recombinase supporting the extension of ribozyme catalyzed RNA cleavage to an allosterically controlled process. Aptazymes responding to theophylline or FMN as cognate ligands are produced by ribozyme supported recombination.



*R. Hieronymus, Prof. Dr. S. Müller**