doi:10.2533/chimia.2024.200 Chimia 78 (2024) 200–204

From Enigma to Revelation: Unravelling Biological Functions of Ubiquitous Small Ribozymes

Kasimir Kienbeck§*, Lukas Malfertheiner, Susann Zelger-Paulus, Silke Johannsen, Christian von Mering, and Roland K. O. Sigel*

§SCS-Metrohm Award for best oral presentation in Medicinal Chemistry

Abstract: RNA, widely recognized as an information-carrier molecule, is capable of catalyzing essential biological processes through ribozymes. Despite their ubiquity, specific functions in a biological context and phenotypes based on the ribozymes' activity are often unknown. Here, we present the discovery of a subgroup of minimal HDV-like ribozymes, which reside 3' to viral tRNAs and appear to cleave the 3'-trailers of viral premature tRNA transcripts. This proposed tRNA-processing function is unprecedented for any ribozyme, thus, we designate this subgroup as *theta* ribozymes. Most *theta* ribozymes were identified in *Caudoviricetes* bacteriophages, the main constituent (>90%) of the mammalian gut virome. Intriguingly, our findings further suggest the involvement of *theta* ribozymes in the transition of certain bacteriophages between distinct genetic codes, thus possibly contri-buting to the phage lysis trigger. Our discovery expands the limited repertoire of biological functions attributed to HDV-like ribozymes and provides insights into the fascinating world of RNA catalysis.

Keywords: Gut virome · RNA catalysis · tRNA maturation · Small ribozymes

Kasimir Kienbeck started his career in Chemical Biology at age fourteen by attending the Secondary College for Chemical Technology, HBLVA Rosensteingasse, in Vienna, Austria. His interest in understanding life at a molecular level made him pursue a BSc in Molecular Biotechnology at the University of Applied Sciences, FH Campus Wien, followed by an MSc in Molecular Microbiology and Immunobiology at the University of Vienna. After five years of working as a chemist at a small startup, Kasimir decided to start pursuing a PhD at the University of Zurich, combining the topics of his bachelor's (bacteriophages) and master's theses (RNA).

§Parts of this article have been published in open access form in 'Identification of HDV-Like *Theta* Ribozymes Involved in tRNA-based Recoding of Gut Bacteriophages', K. Kienbeck, L. Malfertheiner, S. Zelger-Paulus, S. Johannsen, C. von Mering, R. K. O. Sigel, *Nature Commun.* **2024**, 15.[45]

1. Introduction to Catalytic RNAs

RNA has gained public awareness in recent years due to its use in mRNA vaccines and thus its significant role in combating the Covid-19 pandemic. However, only a minimal fraction of the human transcriptome codes for proteins (coding RNA). Interestingly, the proportion of non-coding RNA (ncRNA) increases proportionally to the complexity of organisms: In prokaryotes, 13% of the genome is transcribed to ncRNA, whereas this percentage rises to 98% in multicellular organisms.[1] Investigation of what was long thought to be 'junk' has not only led to Nobel Prize-winning gene therapy tools[2] but has also changed our view on the catalysis of enzymatic reactions in cells and the origin of life as we know it.[3–5] While certain classes are thoroughly investigated, our understanding of the entirety of ncRNA functions remains incomplete.[6,7] RNA enzymes, in short, ribozymes, are ncRNA sequences that fold into specific secondary and tertiary structures, allowing them to catalyze chemical reactions. Thus, RNA is the only naturally occurring molecule capable of both storing genetic information and exhibiting catalytic activity. Ribozymes have been identified in all domains of life and are essential participants in life-sustaining mechanisms such as peptide bond formation of proteins,[8] transfer-RNA (tRNA) maturation,[9] and mRNA splicing.[10]

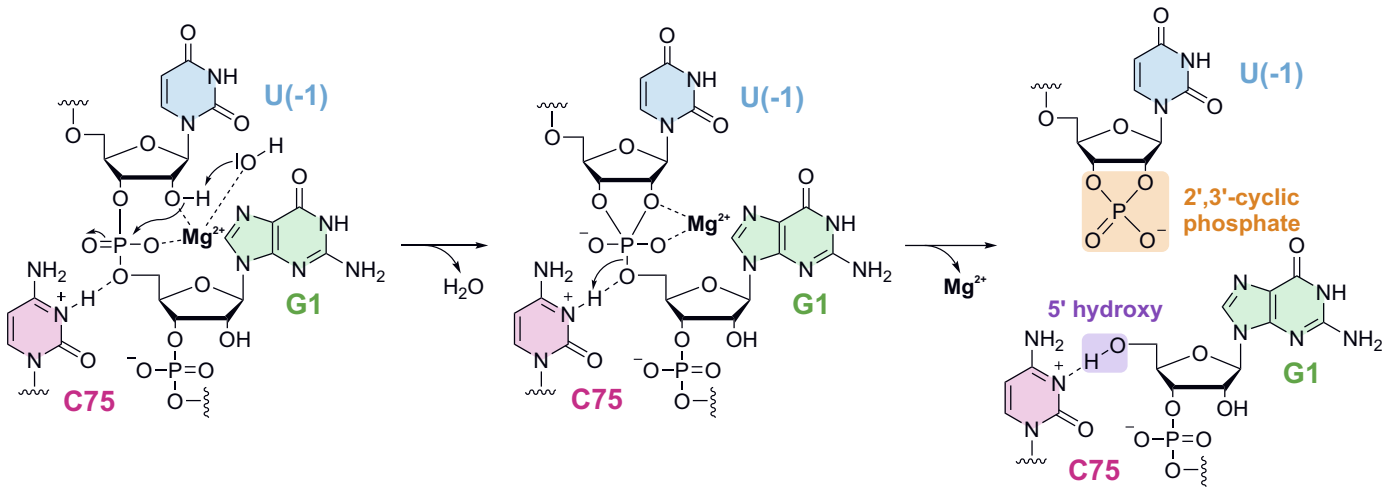
1.1 *Small Self-cleaving Ribozymes*

The classes of small self-cleaving ribozymes are characterized by their concise sequences (typically shorter than 200 nucleotides (nt)) and are restricted to self-scission and/or -ligation. These ubiquitous ribozymes are astoundingly diverse in their sequences, three-dimensional structures, and biological functions.[11–15] Despite this diversity, they all catalyze the same chemical reaction: Internal transesterification of their own phosphate backbone. The 3'-phosphate is attacked by the adjacent 2'-oxygen of the ribose, resulting in a 5'- hydroxy and a 2',3'-cyclic phosphate group (Scheme 1).[16] This reaction also occurs uncatalyzed in RNA molecules, giving them their intrinsic instability. However, the tertiary architecture of the ribozyme fold together with further activators accelerates the reaction by several orders of magnitude and restricts it to a single site in the primary sequence (site-specific).

1.2 *HDV-like Ribozymes*

Among the, roughly, dozen classes of small self-cleaving ribozymes, this article focuses on a ribozyme class that was first discovered in the Hepatitis Delta Virus (HDV, Fig. 1a).[17] Since its discovery in the 1980s, ribozymes with similar secondary structure and cleavage mechanism have been discovered abundantly

*Correspondence: K. Kienbeck, E-mail: kasimir.kienbeck@chem.uzh.ch; Prof. R. K. O. Sigel, E-mail: roland.sigel@chem.uzh.ch
Department of Chemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich



Scheme 1. Mechanism of small ribozyme self-scission as hypothesized for the Hepatitis Delta Virus (HDV) ribozyme. Acid-base catalysis of the internal transesterification in the HDV ribozyme active site: The highly conserved cytosine residue C75 in its protonated form acts together with a partially hydrated Mg^{2+} being the activator for the nucleophile as the general acid and general base, respectively. Adapted from Golden *et al*.[16]

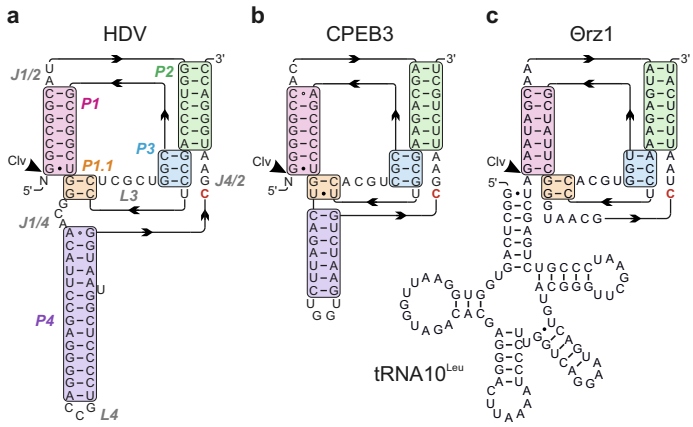


Fig. 1. Secondary structures of HDV-like ribozyme examples. **a**, The genomic HDV ribozyme.[17] **b**, The human CPEB3 ribozyme.[18] **c**, The most common theta ribozyme (Θrz) found in annotated and metagenomic databases linked to its most commonly associated tRNA.[45] Domains are indicated in **a** and double helices are color-coded. The catalytic cytosine residue is marked in dark red. The cleavage site (Clv) is indicated by an arrow.

in all domains of life, including humans (CPEB3 ribozyme, Fig. 1b).[18] Despite the abundance of these so-called HDV-like ribozymes, only a handful of biological functions have been attributed to them to date.[17,19–23]

1.3 *Structure-function Prediction*

When trying to assign a biological function to a ribozyme, the first place to look is its genetic context, as genetic loci in proximity are usually biologically connected. Currently known HDV-like examples were discovered mostly in non-informative genetic loci, such as within non-long terminal repeats[20,21] and coding regions or in metagenomic sequencing raw reads.[24] The latter complicates things further, since raw reads usually span only ~150 to 200 nt, aggravating the prediction of coding regions and thus genomic contexts. Apart from that, the identification of novel ribozymes in the first place is difficult since their primary sequences are poorly conserved. Instead, their complex tertiary structural motifs enable their catalytic activities. For example, the HDV-like motif comprises two intertwined pseudoknots, *i.e.* structures that result from the loop of a hairpin base pairing with a complementary, single-stranded region outside of that hairpin (Fig. 1). Such complex secondary interactions are currently impossible to predict reliably using available software. Furthermore, even if novel potentially active RNA motifs are discovered, a catalytic activity must be assigned, which can be extremely challenging. For this reason, self-cleaving ribozymes are among the most studied ribozymes since methods to characterize their site-specific activity are well-established. However, if a specific structural motif of a ribozyme class is known, it is possible to perform searches using this pre-defined motif and software such as RNArobo.[25] The search results comprise all sequences that are theoretically capable of adapting this pre-defined secondary structure. This approach can be used to discover genetic locations of already known ribozyme motifs within newly available sequence databases, thus serving as a supporting method for identifying potential biological functions of these ribozymes.

2. Results

2.1 *The Discovery of Theta Ribozymes*

We were intrigued by minimal versions of HDV-like ribozymes, which lack one of the five double helical domains (P4), but exhibit an unchanged catalytic signature (Fig. 1c). These minimal variants were discovered 10 years ago by a motif search,[24] yet their biological functions remained elusive. We used an adapted search motif here on more recently available sequencing data. This led to the observation that minimal HDV-like ribozymes often occur with their cleavage site directly adjacent to tRNA genes in bacteriophages associated with the human gut. Due to this association, we termed these specific HDV-like ribozymes *theta* ribozymes (Θrz). Premature tRNAs require the scission of their respective 5'- and 3'-trailers for their maturation. While the 5'-trailer is processed solely by RNase P,[26] 3'-trailer processing requires complex enzyme machineries.[27] We interpreted the association of a Θrz with a tRNA as a clear indication that Θrzs can take over the task of site-specific 3'-trailer scission, thus reducing the genetic space required to encode for the machinery to perform this function. This was the first time such a biological function was proposed for any ribozyme.

2.2 *The Metagenomic Hunt for Θrzs*

We wanted to know whether our findings were rare or widespread, so we expanded our search to larger databases, including metagenomic sequencing data. First, we optimized the search motif to reduce the number of false-positive hits, which we defined as putatively inactive motifs (Fig. 2a and b). In four optimization

steps utilizing annotated viral databases,[28–37] we achieved a final search motif yielding a very low false positive rate and a high association with tRNAs (Fig. 2c and d). Using this final motif, we searched both annotated databases[28–37] as well as publicly available metagenomic raw reads within the MicrobeAtlas project mostly based on MAPseq[38] (microbeatlas.org) and discovered 1'753 unique Θrz sequences adjacent to 5'810 unique tRNAs heavily overrepresented in sequences associated with the mammalian gut. This large number was astounding considering that we searched raw reads (~150-200 nt) approximately the same length as a Θrz linked to a tRNA (~130-150 nt). Furthermore, previous, similar searches yielded orders of magnitude fewer ribozymes.[24]

2.3 *Θrzs are Active in vitro*

Having discovered a large number of ribozymes with a proposed novel function, we wanted to know if they perform HDV-like self-scission *in vitro*. For this reason, we measured the apparent self-cleavage rate constant (k_{obs}) of four different Θrz/tRNA pairs at different Mg^{2+} concentrations and pH values. As expected for HDV-like ribozymes, the self-cleavage rates increase with rising Mg^{2+} concentration (Fig. 2e) and have their optima at neutral pH (Fig. 2f). This is because the internal transesterification mechanism of HDV-like ribozymes relies on an essential cytosine residue in the J4/2 junction (C75 in the HDV ribozyme; Fig. 2a). This residue shows a perturbed pK_a and together with a nearby essential Mg^{2+} ion enables acid-base catalysis at ambient pH, activating the nucleophile as well as stabilizing the transition state and leaving group (Scheme 1). We confirmed the identity and necessity of the catalytic cytosine residue by mutating it to uracil, which completely abolished the self-scission activity (data not shown), thus also validating our approach of estimating false positives in the motif search (Fig. 2b). This is consistent with recent studies showing similar inactivation with no rescue mutation.[39,40]

2.4 *Suppressor tRNA-associated Θrzs Reveal Stop-codon Recoding*

To better understand the possible biological relevance of Θrzs, we investigated the identity of their associated tRNAs. Surprisingly, one fifth of our hits is associated with a suppressor tRNA ($tRNA^{Sup}$), making this the second-most commonly associated tRNA type after $tRNA^{Met}$. Furthermore, 99.7% of the associated $tRNA^{Sup}$ contain an anticodon to the amber stop-codon (UAG), strongly suggesting that this stop-codon may be reassigned to incorporate an amino acid instead of terminating translation, *i.e.* the bacteriophage is recoded. Based on recent findings by Borges *et al.*[31] that ~2-6% of human and animal gut phages are recoded, we used a similar approach to predict the genetic codes of the annotated bacteriophage genomes used in our motif searches. This method relies on the fact that recoded genomes have fragmented genes when predicted using the standard genetic code instead of high coding density when using an alternative code. Remarkably, we found a very high positive correlation between the predicted recoding of genomes and the occurrence of tRNA-associated Θrzs (59.0%). Notably, if a genome contains a $tRNA^{Sup}$-associated Θrz, we observe recoding in 96.4% of those phages. In all of these cases, the amber stop-codon is recoded to glutamine, which correlates with $tRNA^{Sup}$ isotype predictions (69.4% show a bacterial Gln isotype).

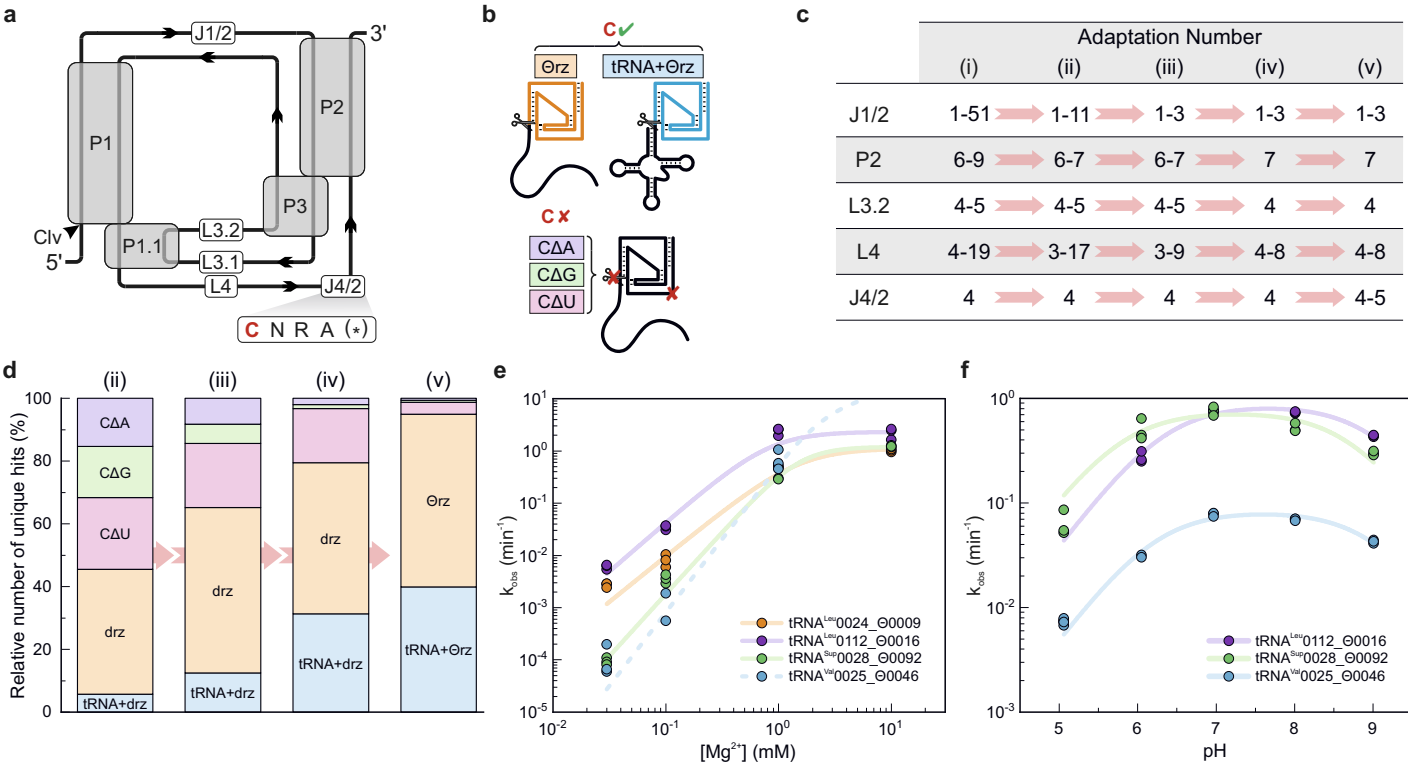


Fig. 2. Search motif adaptation and *in vitro* confirmations. **a**, Representation of the *theta* ribozyme (Θrz) search motif with domain and linker designations. The catalytic cytosine residue in J4/2 is marked in dark red. The cleavage site (Clv) is indicated by an arrow. **b**, Four different search motifs were used, an active one with the catalytic cytosine intact (blue and orange) as well as false positive motifs where this residue is replaced by A, G, or U (purple, green, and magenta, respectively). **c**, Adaptations to the length of each modified domain and linker region are indicated by light red arrows. (i) corresponds to the minimal motif from Riccitelli *et al.*[24] Nucleotide identity restrictions were also introduced at certain positions but are not shown. **d**, Hits obtained by using each of the four search motifs colored as in **b**. Each bar is labeled to the corresponding adaptation number (ii) – (v) shown in **c**. Adaptations are indicated by light red arrows. HDV-like ribozymes: drz. **e**, Calculated apparent kinetic rate constants (k_{obs}) of four Θrz/tRNA-pairs at varying Mg^{2+} concentrations. The dashed line represents an estimation due to too fast cleavage of one pair. **f**, Calculated k_{obs} of three Θrz/tRNA-pairs at varying pH values. Adapted from Kienbeck *et al.*[45]

3. Conclusions

Our results demonstrate a strong correlation between Θrz occurrence and recoding of bacteriophages. Recent research has shown that phage lysis- and structural genes are overrepresented among the recoded genes,[31] thus leading us to our hypothesis that Θrzs may play a role in the code switch of certain bacteriophages (Fig. 3). This transition to an alternative genetic code may ultimately contribute to the lysis trigger, a still poorly understood event. Errors in the precise timing of this trigger can be detrimental to the phage, which is used by some bacteria as a defense mechanism: By initiating lysis before phage particles are fully assembled, the phage efficiency is severely compromised.[41–43]

Our results suggest that 3'-trailer cleavage by Θrzs is not a rare phenomenon but rather a well-established mechanism among certain *Caudoviricetes* phages, the predominant double-stranded DNA virus class in the human gut virome.[44] We hypothesize that the reduced genomic space required for processing tRNA 3'-trailers by Θrzs contributes to phage fitness and may offer additional regulatory opportunities. This assumption still holds true if we consider that one Θrz must be present at every tRNA, as opposed to a single enzyme complex capable of processing multiple tRNA 3'-trailers, due to the minimal size of Θrzs (mostly <60 nt). However, how Θrzs are activated and/or regulated *in vivo* is still unclear. A reasonable hypothesis is that once transcribed, Θrzs remain in an 'always on' state, thus drastically increasing $tRNA^{Sup}$ concentrations in the host cell. Once a critical threshold is surpassed, this concentration effect alone could overwhelm host-encoded termination factors and lead to the code switch and subsequent production of structural and lysis genes. This opens up avenues for regulation at the transcriptional level, not necessarily of the ribozyme itself.

Our results highlight the importance and abundance of small ribozymes and introduce a subset of tRNA-processing HDV-like ribozymes, which may enable bacteriophages to manipulate their bacterial hosts in mammalian gut microbiomes.

Acknowledgements

Kasimir is grateful to the Swiss Chemical Society and Metrohm AG for the generous award of 'Best Oral Presentation in Chemical Biology'. Financial support from the Swiss National Science Foundation [200020_192153, to SZP and RKOS and 310030_192569, to LM and CvM], the University of Zurich [RKOS, CvM, and Candoc Grant FK-23-101 to KK], and the Graduate School of Chemical and Molecular Sciences Zurich (CMSZH) is gratefully acknowledged.

Received: January 30, 2024

[1] G. Liu, J. S. Mattick, R. J. Taft, *Cell Cycle* **2013**, *12*, 2061, https://doi.org/10.4161/cc.25134.
[2] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, *Science* **2012**, *337*, 816, https://doi.org/10.1126/science.1225829.
[3] W. Gilbert, *Nature* **1986**, *319*, 618, https://doi.org/10.1038/319618a0.
[4] M. Neveu, H.-J. Kim, S. A. Benner, *Astrobiology* **2013**, *13*, 391, https://doi.org/10.1089/ast.2012.0868.
[5] A. Pressman, C. Blanco, I. A. Chen, *Curr. Biol.* **2015**, *25*, R953, https://doi.org/10.1016/j.cub.2015.06.016.
[6] T. R. Cech, J. A. Steitz, *Cell* **2014**, *157*, 77, https://doi.org/10.1016/j.cell.2014.03.008.
[7] L. R. Ganser, M. L. Kelly, D. Herschlag, H. M. Al-Hashimi, *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 474, https://doi.org/10.1038/s41580-019-0136-0.
[8] P. Nissen, J. Hansen, N. Ban, P. B. Moore, T. A. Steitz, *Science* **2000**, *289*, 920, https://doi.org/10.1126/science.289.5481.920.
[9] C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace, S. Altman, *Cell* **1983**, *35*, 849, https://doi.org/10.1016/0092-8674(83)90117-4.
[10] M. E. Wilkinson, C. Charenton, K. Nagai, *Annu. Rev. Biochem.* **2020**, *89*, 359, https://doi.org/10.1146/annurev-biochem-091719-064225.
[11] P. C. Bevilacqua, R. Yajima, *Curr. Opin. Chem. Biol.* **2006**, *10*, 455, https://doi.org/10.1016/j.cbpa.2006.08.014.
[12] M. Egger, R. Bereiter, S. Mair, R. Micura, *Angew. Chem. Int. Ed* **2022**, *61*, e202207590, https://doi.org/10.1002/anie.202207590.
[13] C. E. Weinberg, Z. Weinberg, C. Hammann, *Nucleic Acids Res.* **2019**, *47*, 9480, https://doi.org/10.1093/nar/gkz737.
[14] H. Peng, B. Latifi, S. Müller, A. Lupták, I. A. Chen, *RSC Chem. Biol.* **2021**, *2*, 1370, https://doi.org/10.1039/d0cb00207k.
[15] A. Ren, R. Micura, D. J. Patel, *Curr. Opin. Chem. Biol.* **2017**, *41*, 71, https://doi.org/10.1016/j.cbpa.2017.09.017.

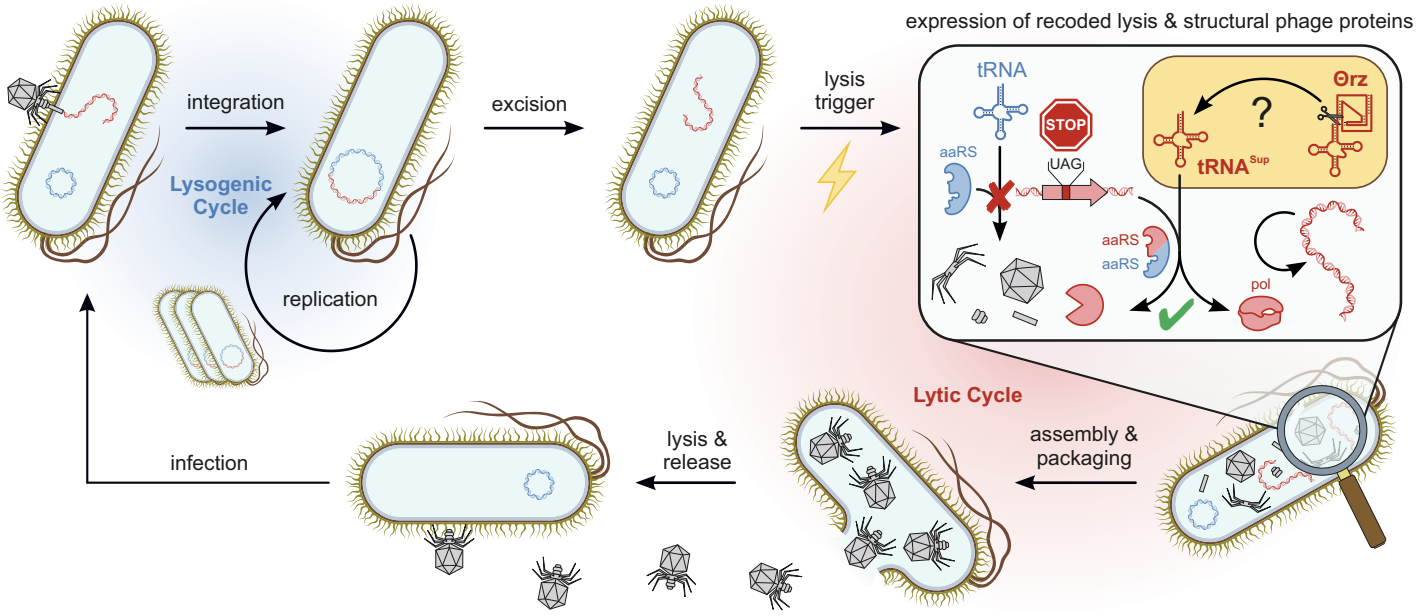


Fig. 3. Putative phage infection cycles involving *theta* ribozymes (Θrz). Infection of a bacterium of the *Bacteroidota* or *Bacillota* phylum by a recoded *Caudoviricetes* bacteriophage. The phage can either enter the lysogenic cycle, where the host replicates the integrated phage genome, or a not fully understood lysis trigger can lead to the lytic cycle. We propose that suppressor tRNA ($tRNA^{Sup}$)-associated Θrzs play a role in shifting the equilibrium of translation factors into the recoded context, thus enabling the expression of phage lysis and structural genes. These genes could otherwise not be expressed by the host machinery due to premature stop-codons resulting in gene fragmentation. Once the phage particles self-assemble, the host cell is lysed and the phage particles are released, leading to a new infection cycle. Aminoacyl synthetase: aaRS. Polymerase: pol. Colors: Host DNA/proteins: blue; phage DNA/RNA/proteins: red.

[16] B. L. Golden, *Biochemistry* **2011**, *50*, 9424, https://doi.org/10.1021/bi201157t.
[17] L. Sharmeen, M. Y. P. Kuo, G. Dinter-Gottlieb, J. Taylor, *J. Virol.* **1988**, *62*, 2674. https://doi.org/10.1128/JVI.62.8.2674-2679.1988.
[18] K. Salehi-Ashtiani, A. Lupták, A. Litovchick, J. W. Szostak, *Science* **2006**, *313*, 1788, https://doi.org/10.1126/science.1129308.
[19] C. Vogler, K. Spalek, A. Aerni, P. Demougin, A. Müller, K.-D. Huynh, A. Papassotiropoul, D. J.-F. de Quervain, *Front. Behav. Neurosci.* **2009**, *3*, 1, https://doi.org/10.3389/neuro.04.010.2009.
[20] D. J. Ruminski, C.-H. T. Webb, N. J. Riccitelli, A. Lupták, *J. Biol. Chem.* **2011**, *286*, 41286, https://doi.org/10.1074/jbc.M111.297283.
[21] D. G. Eickbush, T. H. Eickbush, *Mol. Cell. Biol.* **2010**, *30*, 3142, https://doi.org/10.1128/MCB.00300-10.
[22] F. J. Sánchez-Luque, M. C. López, F. Macias, C. Alonso, M. C. Thomas, *Nucleic Acids Res.* **2011**, *39*, 8065, https://doi.org/10.1093/nar/gkr478.
[23] J. L. Jakubczak, W. D. Burke, T. H. Eickbush, *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 3295, https://doi.org/10.1073/pnas.88.8.3295.
[24] N. J. Riccitelli, E. Delwart, A. Lupták, *Biochemistry* **2014**, *53*, 1616, https://doi.org/10.1021/bi401717w.
[25] L. Rampášek, R. M. Jimenez, A. Lupták, T. Vinař, B. Brejová, *BMC Bioinformatics* **2016**, *17*, 216, https://doi.org/10.1186/s12859-016-1074-x.
[26] J. C. Ellis, J. W. Brown, *RNA Biol.* **2009**, *6*, 362, https://doi.org/10.4161/rna.6.4.9241.
[27] S. Schiffer, S. Rösch, A. Marchfelder, *EMBO J.* **2002**, *21*, 2769, https://doi.org/10.1093/emboj/21.11.2769.
[28] M. J. Tisza, C. B. Buck, *Proc. Natl. Acad. Sci. U.S.A* **2021**, *118*, e2023202118, https://doi.org/10.1073/pnas.2023202118.
[29] L. F. Camarillo-Guerrero, A. Almeida, G. Rangel-Pineros, R. D. Finn, T. D. Lawley, *Cell* **2021**, *184*, 1098, https://doi.org/10.1016/j.cell.2021.01.029.
[30] S. Nayfach, D. Páez-Espino, L. Call, S. J. Low, H. Sberro, N. N. Ivanova, A. D. Proal, M. A. Fischbach, A. S. Bhatt, P. Hugenholtz, N. C. Kyrpides, *Nat. Microbiol.* **2021**, *6*, 960, https://doi.org/10.1038/s41564-021-00928-6.
[31] A. L. Borges, Y. C. Lou, R. Sachdeva, B. Al-Shayeb, P. I. Penev, A. L. Jaffe, S. Lei, J. M. Santini, J. F. Banfield, *Nat. Microbiol.* **2022**, *7*, 918, https://doi.org/10.1038/s41564-022-01128-6.
[32] S. Li, R. Guo, Y. Zhang, P. Li, F. Chen, X. Wang, J. Li, Z. Jie, Q. Lv, H. Jin, G. Wang, Q. Yan, *iScience* **2022**, *25*, 104418, https://doi.org/10.1016/j.isci.2022.104418.
[33] A. C. Gregory, A. A. Zayed, N. Conceição-Neto, B. Temperton, B. Bolduc, A. Alberti, M. Ardyna, K. Arkhipova, M. Carmichael, C. Cruaud, C. Dimier, G. Domínguez-Huerta, J. Ferland, S. Kandels, Y. Liu, C. Marec, S. Pesant, M. Picheral, S. Pisarev, J. Poulain, J.-É. Tremblay, D. Vik, Tara Oceans Coordinators, M. Babin, C. Bowler, A. I. Culley, C. de Vargas, B. E. Dutilh, D. Iudicone, L. Karp-Boss, S. Roux, S. Sunagawa, P. Wincker, M. B. Sullivan, *Cell* **2019**, *177*, 1109, https://doi.org/10.1016/j.cell.2019.03.040.
[34] M. Stano, G. Beke, L. Klucar, *Database J. Biol. Databases Curation* **2016**, baw162, https://doi.org/10.1093/database/baw162.
[35] S. Roux, D. Páez-Espino, I.-M. A. Chen, K. Palaniappan, A. Ratner, K. Chu, T. B. K. Reddy, S. Nayfach, F. Schulz, L. Call, R. Y. Neches, T. Woyke, N. N. Ivanova, E. A. Eloe-Fadrosh, N. C. Kyrpides, *Nucleic Acids Res.* **2021**, *49*, D764, https://doi.org/10.1093/nar/gkaa946.
[36] A. Fullam, I. Letunic, T. S. B. Schmidt, Q. R. Ducarmon, N. Karcher, S. Khedkar, M. Kuhn, M. Larralde, O. M. Maistrenko, L. Malfertheiner, A. Milanese, J. F. M. Rodrigues, C. Sanchis-López, C. Schudoma, D. Szklarczyk, S. Sunagawa, G. Zeller, J. Huerta-Cepas, C. von Mering, P. Bork, D. R. Mende, *Nucleic Acids Res.* **2023**, *51*, D760, https://doi.org/10.1093/nar/gkac1078.
[37] N. A. O'Leary, M. W. Wright, J. R. Brister, S. Ciufo, D. Haddad, R. McVeigh, B. Rajput, B. Robbertse, B. Smith-White, D. Ako-Adjei, A. Astashyn, A. Badretdin, Y. Bao, O. Blinkova, V. Brover, V. Chetvernin, J. Choi, E. Cox, O. Ermolaeva, C. M. Farrell, T. Goldfarb, T. Gupta, D. Haft, E. Hatcher, W. Hlavina, V. S. Joardar, V. K. Kodali, W. Li, D. Maglott, P. Masterson, K. M. McGarvey, M. R. Murphy, K. O'Neill, S. Pujar, S. H. Rangwala, D. Rausch, L. D. Riddick, C. Schoch, A. Shkeda, S. S. Storz, H. Sun, F. Thibaud-Nissen, I. Tolstoy, R. E. Tully, A. R. Vatsan, C. Wallin, D. Webb, W. Wu, M. J. Landrum, A. Kimchi, T. Tatusova, M. DiCuccio, P. Kitts, T. D. Murphy, K. D. Pruitt, *Nucleic Acids Res.* **2016**, *44*, D733, https://doi.org/10.1093/nar/gkv1189.
[38] J. F. Matias Rodrigues, T. S. B. Schmidt, J. Tackmann, C. von Mering, *Bioinformatics* **2017**, *33*, 3808, https://doi.org/10.1093/bioinformatics/btx517.
[39] A. Ke, K. Zhou, F. Ding, J. H. D. Cate, J. A. Doudna, *Nature* **2004**, *429*, 201, https://doi.org/10.1038/nature02522.
[40] J. M. Roberts, J. D. Beck, T. B. Pollock, D. P. Bendixsen, E. J. Hayden, *eLife* **2023**, *12*, e80360, https://doi.org/10.7554/eLife.80360.
[41] E. Durmaz, T. R. Klaenhammer, *J. Bacteriol.* **2007**, *189*, 1417, https://doi.org/10.1128/JB.00904-06.
[42] S. G. Hays, K. D. Seed, *eLife* **2020**, *9*, e53200, https://doi.org/10.7554/eLife.53200.
[43] R. Johnson-Boaz, C.-Y. Chang, R. Young, *Mol. Microbiol.* **1994**, *13*, 495, https://doi.org/10.1111/j.1365-2958.1994.tb00444.x.
[44] G. Liang, F. D. Bushman, *Nat. Rev. Microbiol.* **2021**, *19*, 514, https://doi.org/10.1038/s41579-021-00536-5.
[45] K. Kienbeck, L. Malfertheiner, S. Zelger-Paulus, S. Johannsen, C. von Mering, R. K. O. Sigel, *Nat. Commun.* **2024**, *15*, https://doi.org/10.1038/s41467-024-45653-w.

The definitive version of this article is the electronic one that can be found at https://doi.org/10.2533/chimia.2024.200