doi:10.2533/chimia.2023.235 Chimia 77 (2023) 235–241

Group II Introns: Highly Structured yet Dynamic

Esra Ahunbay[§]*, Susann Zelger-Paulus, and Roland K. O. Sigel*

[§]SCS-DSM Award for best poster in Chemical Biology

Abstract: RNA splicing, the removal of introns and ligation of exons, is a crucial process during mRNA maturation. Group II introns are large ribozymes that self-catalyze their splicing, as well as their transposition. They are living fossils of spliceosomal introns and eukaryotic retroelements. The yeast mitochondrial *Sc*.ai5γ is the first identified and best-studied self-splicing group II intron. A combination of biochemical, biophysical, and computational tools enables its catalytic properties, structure, and dynamics to be studied, while also serving to develop new therapeutic and biotechnological tools. We survey the history of group II intron studies paralleling the trends in RNA methodology with *Sc*.ai5γ in the spotlight.

Keywords: Group II intron · RNA catalysis · RNA folding · RNA labeling · RNA splicing

Esra Ahunbay obtained her BSc in Chemistry at Bogazici University, Turkey, and her MSc in Molecular and Biological Chemistry at École Polytechnique Fédérale de Lausanne (EPFL). At EPFL, she developed carbon nanotube-based fluorescent biosensors under the supervision of Prof. Ardemis Boghossian. Then, she investigated the epigenetic mechanisms in cellular reprogramming as a guest researcher in the group of Prof. Tamer Onder at Koç University. She is currently pursuing a PhD in the smFRET subgroup headed by Dr. Susann Zelger-Paulus in the group of Prof. Roland Sigel at the University of Zurich. Esra labels self-splicing RNAs and studies their dynamic behavior using single-molecule FRET.

1. Introduction to Group II Intron Splicing

"Please write down the amino acid sequence of the protein to be expressed from the gene given the following nucleotide sequence" is a classical biology exam question. The task is to transcribe the DNA sequence into RNA, group into three-nucleotide units (codons), identify the start codon (AUG), and finally match each codon with the corresponding amino acid to decode the protein of interest. In reality, nature is more complex. Genes are often discontinuous and split. The "discovery of split genes" by Richard Roberts and Phillip Sharp was awarded the Nobel Prize in Physiology or Medicine in 1993. Protein-coding sequences (exons, expressed regions) are interrupted by non-coding sequences (introns, intervening or intragenic regions), which is the rule and not the exception. RNA splicing is the process that removes the introns and strings together the exons during mRNA maturation so that the correct protein is expressed.

The spliceosome, a multi-megadalton RNA-protein complex, is in charge of intron splicing in eukaryotic nuclei. Most introns in archaeal and eukaryotic tRNAs are spliced by protein enzymes. At the other end of the spectrum are introns that undergo self-splicing without protein support. These types of introns can be further classified into three subtypes (group I, II, and III), differing mainly in size, structure, and splicing mechanism. Among these, group II introns stand out as the progenitor of the spliceosome.[1,2] Spliceosomes and group II introns share the exact splicing mechanism and key regulatory elements. Group II introns are exceptionally long (400–1000 nt) RNAs and are abundant in housekeeping genes that require tight regulation and efficient splicing. They are found in archaea, bacteria, viruses, plants, and organelles of lower eukaryotes,[2] and have been recently identified in annelids.[3] Their absence in human cells makes them potential antimicrobial drug targets.[4,5]

Back in 1980, using electron microscopy, Arnberg *et al*. reported the excised circular introns in yeast mitochondria,[6] which hitherto was an observation exclusive to the spliceosomal introns. In 1986, van der Veen *et al.* and Peebles *et al.* both showed that a *Saccharomyces cerevisiae* (*S. cerevisiae, Sc*.) mitochondrial group II intron self-splices *in vitro* to form these so-called lariat structures.[7,74] This first identified group II intron was *Sc*.ai5γ (or ai5c). Group II introns were later subclassified as A, B, C, D, E, F, mitochondrial-like, and chloroplast-like,[1,2] with *Sc*.ai5γ belonging to the subclass of group IIB introns. *Sc*.ai5γ is the seventh intron of the *cox1* (or *oxi3*) gene, encoding for cytochrome c oxidase subunit 1 (COX1). COX1 is essential to cell survival as it plays a key role in the mitochondrial respiratory chain by driving oxidative phosphorylation.

Group II introns are versatile RNA that are the ancestor of the spliceosome and eukaryotic retroelements. They are capable of autocatalytic splicing and retromobility. Despite little primary sequence conservation, secondary elements of group II intron render its characteristic architecture of six helical domains (D1-6).[2] Before the intronic RNA becomes catalytically active, it must fold into a specific three-dimensional (3D) structure. The formation of so-called tertiary contacts (intradomain, interdomain, and intron-exon interactions) is crucial in this process. They fold in an exceptionally distinct manner to form their highly structured multidomain scaffold. Once the RNA is correctly folded, the intron self-splices by catalyzing the cleavage and ligation of the exonic regions. This review surveys how advances in RNA methodologies, as well as changing perspectives and interests have shaped

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

the way group II introns were studied. The *Sc*.ai5γ group II intron persisted at the heart of these studies.

2. Self-catalyzed Splicing

Protein enzymes might today dominate catalysis in biological systems, yet in the early world, RNA served as the principal biocatalyst.[8,9] RNA enzymology emerged as a new field with the discoveries of group I introns (by Thomas Cech)[10,11] and RNase P (by Sidney Altman).[12] The Nobel Prize in Chemistry in 1989 honored their "discovery of catalytic properties of RNA". Ribozymes, RNA enzymes, are remnants of the "RNA world". At the time of coining the term "ribozyme",[13] the *Tetrahymena thermophila* self-splicing group I intron.[10,11] was a unique observation. Cech went for this general classification in the anticipation that there must be other catalytic RNAs. Indeed, there were plenty in all kingdoms of life, and in different sizes. Classified among the 11 subclasses of small self-cleaving ribozymes (<200 nt) are glucosamine-6-phosphate synthase (*glmS*), hammerhead, hairpin, hepatitis delta virus (HDV), HDV-like, pistol, twister ribozymes, and so forth.[14] Group I and II introns, spliceosome, and RNase P constitute the class of large phosphoryl transfer ribozymes (>200 nt).[15] For example, the *Sc*.ai5γ group II intron, excluding its exons, is 887-nt long (~290 kDa). The ribosome is a peptidyl transfer ribozyme.[16]

Group II intron and spliceosomal intron splicing is mechanistically related.[17] RNA splicing is fully RNA-catalyzed with a two-metal ion center in both cases. In the first step of the so-called group II intron branching splicing pathway (Fig. 1), the 2'-OH of an internal adenosine initiates the S_N2 nucleophilic attack on the 5'-splice site phosphate. In the second step, the cleaved 5'-exon attacks the 3'-splice site, ligating the exons. This two-step transesterification mechanism releases the intron as a lariat; a branched RNA with a 2'-5' phosphodiester bond. Instead of an internal nucleotide, a water molecule or a hydroxide ion can act as a nucleophile (Fig. 1). However, this hydrolytic splicing reaction is not fully reversible and yields a linear intron. Whether the hydrolytic splicing path is an artifact under *in vitro* conditions or naturally exists is under debate.[18] Following the branching pathway is essential for retromobility: The excised lariats of the mobile group II introns can reinsert into DNA by retrohoming (into a homologous site) or retrotransposition (into an ectopic site), again fully catalyzed by RNA.[19,20] This is possible as both steps of the branching pathway are reversible. Mobile group II introns are reengineered as gene targeting vectors (targetrons) to insert sequences into the genome,[21] thus serving as biotechnological tools to edit genomes.

Piccirilli and coworkers investigated the two-metal ion catalysis at the active site of group II introns (Fig 1).[22] To identify the structural features that make group II introns self-catalytic, Jarrell *et al.* divided the *Sc*.ai5γ into two pieces.[23] Although not covalently connected, the bipartite group II intron retained catalytic activity, exhibiting modularity. Domain 5 acts enzymatically on eD123 (5'-exon and domains 1, 2, and 3) in *trans*, cleaving its substrate, the 5'-exon (Fig. 2). Many others followed suit and applied variations of this multipartite approach to *Sc*.ai5γ,[24,25] as well as to *Pli.LSUI2*, the mitochondrial rRNA group II intron of brown alga *Pylaiella littoralis* (*P. littoralis*),[26] which is currently the only crystallized group IIB intron.[27] Limiting one of the components over the other enabled studying single- and multiple-turnover kinetics,[24] demonstrating its enzymatic power.

Domains 1 (the scaffold), 3 (the catalytic effector), and 5 (the catalytic core) are sufficient to sustain the catalytic activity of self-cleavage. D135 (Fig. 2),[28] a minimal construct composed of these three domains, successfully cleaves its substrate in *trans*. The definition of substrate differs among studies. Here, the substrate refers to an RNA-oligomer, which contains the last 17-nt of

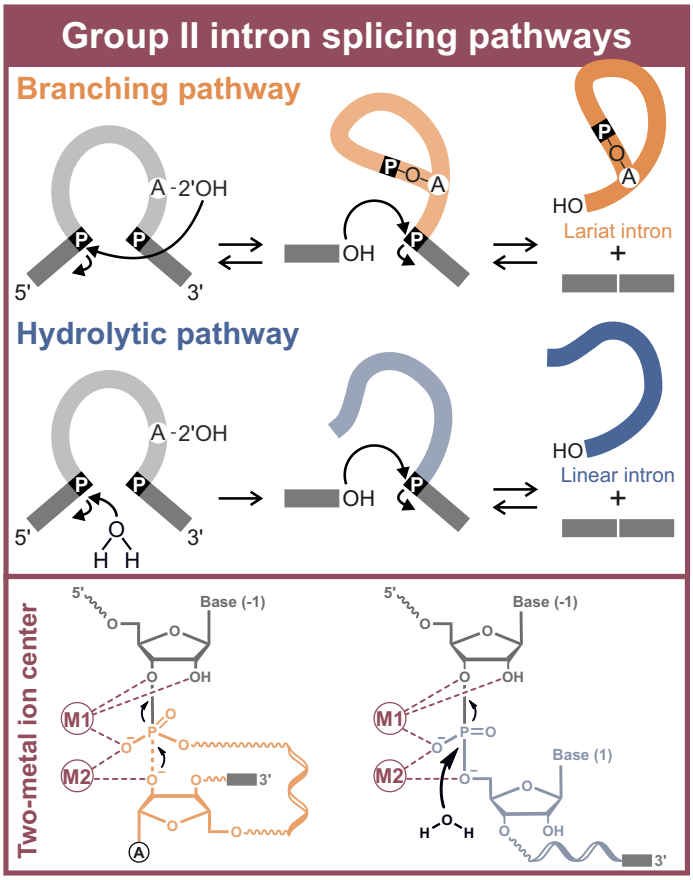

Fig. 1. Group II intron self-splicing *via* branching and hydrolytic pathways (top panel). The bottom panel shows the two catalytic metal ions (M1 and M2) during the transition state (left) and the nucleophilic attack (right) of the first step of both pathways. RNA structure and intron-exon recognition sites are omitted for clarity.

the 5'-exon and the first 7-nt of the intron. As such the 5'-splice site is formed so that the 5'-exon cleavage can be followed.[24] The 5'-exon lengths, shortened in these substrates, are sufficient to cover the intron-binding sites 1 and 2 (IBS1/2) that are complementary to their corresponding exon-binding sites (EBSs). The tenacious grip between EBS-IBS ensures splice site recognition. Another use of *trans* substrate was a short ligation product of 5'- and 3'-exons to follow the spliced exon reopening (SER) reaction.[26]

Two rounds of transesterification define RNA splicing: first 5'-exon cleavage, then exon ligation. The minimal constructs that lack D6 and the 3'-exon (that has the 1-nt long IBS3) can only go through the first. D6 contains the bulged adenosine residue that initiates the 5'-splice site attack and becomes the branch point. In its absence, hydrolysis takes place instead of transesterification (Fig. 1).

Historically, denaturing polyacrylamide gel electrophoresis (PAGE) is the biochemical method of choice in studies of ribozyme catalysis. For this, ^{32}P-labeled phosphate is incorporated into the phosphate backbone (either throughout the RNA by co-transcriptional body-labeling or post-transcriptionally at the 5'- or 3'-ends). The radioactively labeled intron is then subjected to *in vitro* splicing conditions followed by size separation on a gel. Nucleotide analog interference mapping (NAIM) using minimal constructs[29] contributed to defining the structural features for self-catalysis. NAIM drew attention to the interdomain interactions, especially between D1 and D5, that are essential for catalysis, moreover for folding. The efforts then focused on visualizing the catalysis and the conformational changes involved,[30,31] as detailed in the next section.

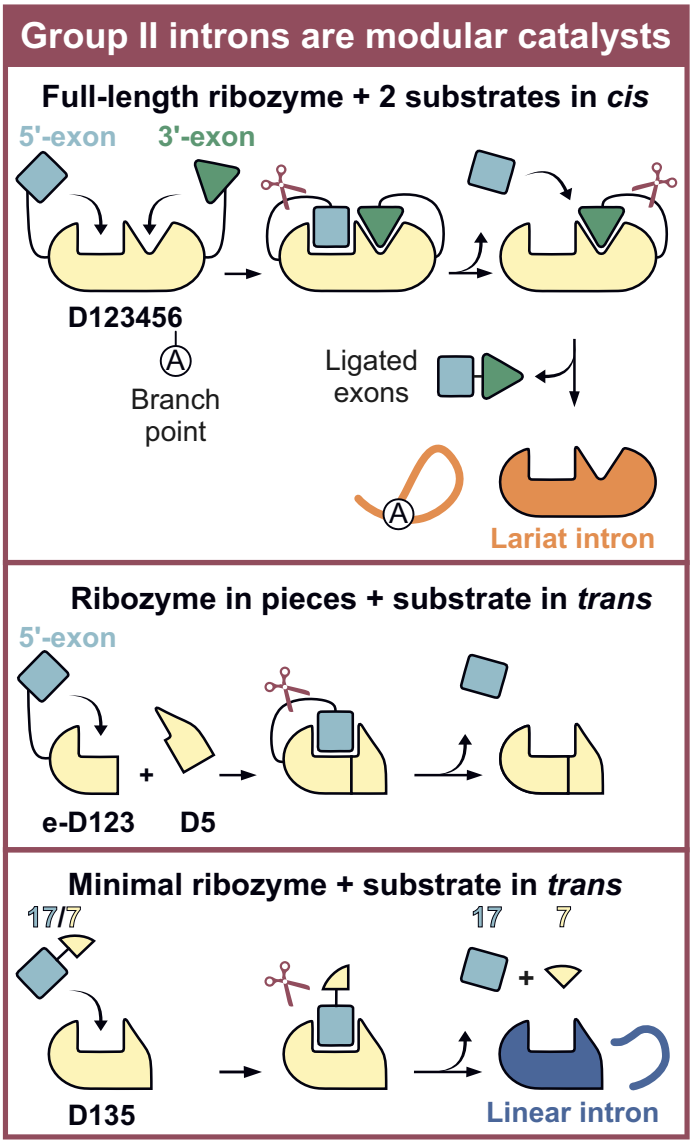

Fig. 2. Studying the RNA splicing through different *Sc*.ai5γ group II intron construct designs. Top; Full-length *Sc*.ai5γ group II intron flanked between its exons *cis*-splices in two steps that ligate its two substrates, the flanking exons. Lariat intron (orange) is released with the adenosine at D6 as the branch point, (bottom) while in its absence, only a linear intron (dark blue) can be released. Group II introns are ribozymes that can reassemble *in vitro* from individually transcribed pieces, with preserved catalytic properties. They act on their substrates (the exons) both in *cis* and *trans*. Middle; D5 hydrolytically cleaves the 5'-exon off of e-D123 (comprised of the 5'-exon and first three domains). Bottom; *Sc*.ai5γ group II intron derivative D135 (composed of domains 1, 3, and 5) is minimally active and *trans*-cleaves its substrate, 17/7 (17-nt of 5'-exon and the first 7-nt of the intron). Scission sites are represented with scissors.

3. Group II Intron Structure and Folding

Group II introns are highly complex molecules that can assume different folding states following a hierarchical pathway. The folding pathway contains unstable intermediates with fast transitions,[32,33] which renders the RNA structure highly dynamic and structure determination very challenging. DI, the largest domain, folds first and serves as a scaffold to which the other domains gradually dock.[34] The interaction of the intron with its flanking exons is crucial and defines the splice sites, which ensures splicing fidelity.[35,36] Correct folding is therefore a prerequisite for proper splicing and thus producing the correct protein.

Sc.ai5γ is catalytically active under specific *in vitro* conditions, *i.e.* high ionic strength (ideally 500 mM KCl and 100 mM $MgCl_2$ at pH 6.9–7.5) and elevated temperatures (42 °C).[37] During folding, metal ions stabilize the RNA structure by charge compensation.[38] In a simplified picture (Fig. 3), the secondary structure forms in the presence of monovalent cations (such as K^+), while the tertiary contacts form in the presence of divalent cations (such as Mg^{2+}). *Sc*.ai5γ splicing in the presence of K^+ ions goes through the hydrolytic pathway, while ammonium (NH_4^+, similar in size to K^+) enigmatically promotes the branching pathway. Daniels *et al*. identified through kinetic analysis that these two reactions compete with each other.[37] Two divalent metal ions (M1 and M2) are directly involved in catalysis (Fig. 1). High Mg^{2+}-ion dependence of *Sc*.ai5γ folding to the active state was investigated by footprinting experiments upon the addition of the substrate (17/7) in *trans*.[39] Solvent-accessible regions that get chemically modified (or cleaved, depending on the reagent), and structured regions involved in secondary and tertiary interactions that are protected can be monitored. This way, local (but not global) changes at the nucleotide level display the folding progress. More detailed pictures of group II intron sections were obtained by nuclear magnetic resonance (NMR), revealing structures of individual domains,[40,41] 5'-splice site,[42] and three-way junction[43] with great detail. However, group II introns as a whole are too large for structure determination by NMR. The high ionic strength requirement of *Sc*.ai5γ poses additional difficulties in NMR.

The dynamics of group II introns do not make them first-choice candidates for X-ray crystallography. With tremendous efforts, Toor *et al.* reported the first crystal structure of a group II intron, *Oceanobacillus iheyensis* bacterial group IIC intron *Oi*.I1.[44] The authors have tested numerous different constructs to achieve packing and crystallizability. Impressively, all these constructs were catalytically active owing to the previous knowledge gained in group II intron catalysis. The structure was then solved and refined by modeling the sequence information.[45] Echoing this work, Marcia and Pyle extended the *Oi*.I1 crystal depository to various stages of catalysis, showing the metal-ion selectivity at the functional four-metal center.[30,31] However, the catalytic core for the branched RNA was not visible. The electron density of D6 was missing[44] or the used constructs lacked D6[30] as its flexibility substantially increased the dynamics, thus decreasing the crystal stability. Robart *et al*. reported the crystal structure of the *Pli.LSUI2* group IIB intron, importantly as a lariat.[27] The technical aspects of crystallographic structure determination of group II introns are discussed elsewhere.[46]

Despite numerous attempts, the crystallization of *Sc*.ai5γ has not been fruitful. Hence the intron has been considered uncrystallizable. However, as the historically most intensively studied model, its tertiary fold provokes great interest. Somarowthu *et al*. used the previously obtained *Oi*.I1 crystal structure as a basis to build a homology model of *Sc*.ai5γ after the first step of splicing, with the cleaved 5'-exon remaining bound (Fig. 3).[47] They added the *Sc*.ai5γ structural motifs obtained *via* NMR, footprinting, and experimentally validated the *de novo* model with hydroxyl radical footprinting and selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE). For more detailed information, the reader is referred to the applications of mutational profiling (SHAPE-MaP) on group II intron.[48]

Compared to the *in vitro* splicing requirements, *Sc*.ai5γ encounters different conditions in its natural environment, the yeast mitochondria, with significantly lower ionic strength (100 mM K^+, 1 mM Mg^{2+} at pH 7.4), an optimal yeast growth temperature of 30 °C, and cellular crowding. Fedorova *et al*. investigated the folding under near-physiological conditions.[49] Dimethyl sulfate (DMS) structural footprinting and native PAGE compared the compaction of D1356 construct at different temperatures; 30 °C versus 42 °C. Intriguingly, pre-folded *Sc*.ai5γ (in the presence of 100 mM Mg^{2+}) cannot maintain its native fold upon dilution. The

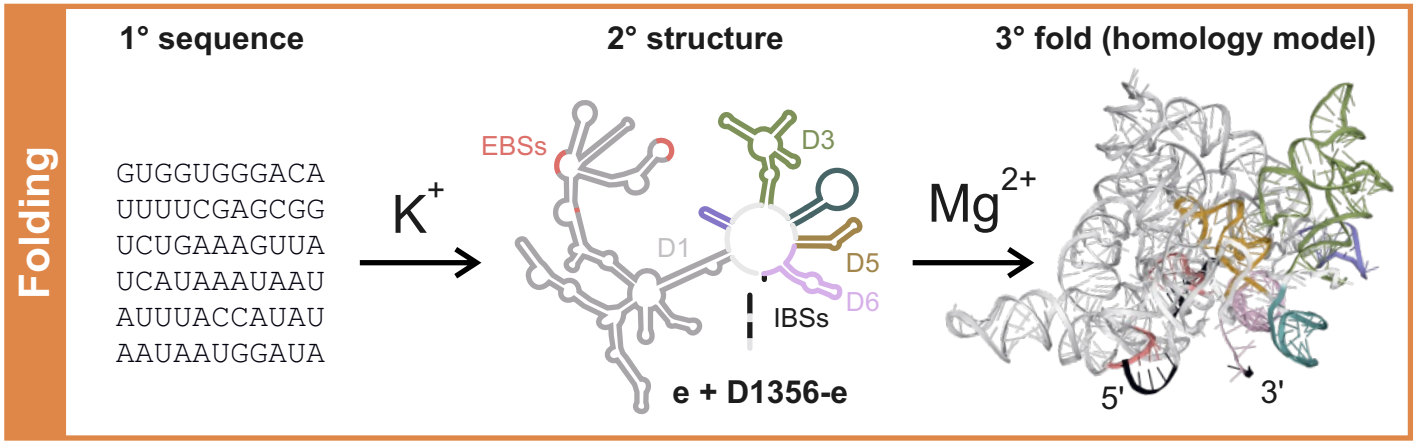

Fig. 3. Folding of the *Sc*.ai5γ group IIB intron in the presence of metal ions towards the secondary and tertiary structure. Secondary structural elements form in the presence of monovalent cations (such as K^+), while Mg^{2+} is essential for long-range tertiary contacts that build the 3D structure. Homology modeling predicts the 3D fold of the *Sc*.ai5γ group IIB intron after the first hydrolytic step. D1 (dark gray) serves as a scaffold where D5 (brown) can dock near the 5'-splice site. Note the base pairing between the exon-binding sites (EBSs, pink) and the intron-binding sites (IBSs, black). Structural features are color coordinated. The homology model is rendered using the file nar-02186-r-2013-File010.pdb.

authors hypothesize that chaperone proteins are required to stabilize the RNA fold under cellular ionic strength.

The *Sc*.ai5γ folding has been associated with Mss116 (or Mss116p) DEAD-box protein that acts as a general RNA chaperone for yeast group I and II introns.[50] Pyle and colleagues have monitored the protein-facilitated folding to reveal that Mss116 stimulates the compaction of D1 and accelerates folding to the native state without directly stabilizing the native state.[51,52]

4. Folding Dynamics with Single-molecule Resolution

The folding of group II introns involves many well-ordered conformational changes, which finally lead to the catalytically active form. Their highly structured framework and dynamic nature complicate their investigation. Traditional methods such as footprinting, NMR, and X-ray crystallography have laid the foundation of group II intron structural studies. However, static structural models lack information on transition states, transition kinetics or the order into which individual states interconvert. Single-molecule resolution is a game-changer. Two methods promise to resolve the group II intron heterogeneity and dynamics; single-particle cryogenic electron microscopy (cryo-EM) and single-molecule Förster resonance energy transfer (smFRET). Cryo-EM, in the hands of Nagai, Lührmann, Stark, Shi, and their coworkers, has contributed tremendously to the structural elucidation of the spliceosomal complexes.[53] The spliceosome might have the upper hand here in terms of protein stabilization, as protein-free ribozymes usually suffer from low resolution in cryo-EM (like crystallography). Nonetheless, cryo-EM has already captured snapshots of group I[54,55] and II introns.[56–58] Single-particle analysis determines the structures of individual states and intermediates, which can then be used to predict the transitions among them. On the other hand, as the sample is frozen, the structures obtained are static.

In contrast, FRET adds dynamic information. FRET is a non-radiative energy transfer between two differently colored fluorophores with spectral overlap. The energy transfer efficiency depends on the distance between the fluorophores (sensitive in the range of 2–10 nm). In other words, it reports the distance changes between the fluorophores when tracked over time. At the single-molecule level, dynamic information obtained includes kinetic connectivity and exchange rates.[59] An important prerequisite for FRET, as well as a limitation compared to the abovementioned methods, is the site-specific labeling of the RNA of interest with fluorophores that report on structural distances and their changes.

4.1 *Group II Intron Labeling*

Group II intron fluorescent labeling, which we recently summarized,[60] is not straightforward. Different RNAs (in terms of size or catalytic activity) call for different labeling strategies (Fig. 4). The labeling positions of choice must fulfill the following criteria: (i) being accessible for chemical or enzymatic reactions, (ii) non-disturbing to preserve RNA function, and (iii) being informative to report on RNA dynamics. Steiner *et al.* initiated the single-molecule group II intron folding studies. The only way, at that time, to label such a large and complex RNA internally was the hybridization technique.[33] The minimal ribozyme (D135) served as a starting point in construct design. Loops on D1 (d2b hairpin) and truncated D4 were incorporated to accommodate the dye-carrying complementary DNA probes, constructing the so-called D135-L14 derivative. Later, Schmitz *et al.* shortened these labeling platforms (D135-shL14) using peptide nucleic acid (PNA) instead of DNA oligonucleotides as hybridizing agents.[61,62] RNA-PNA heteroduplexes have a higher affinity than RNA-DNA. This allows decreasing the 18-nt DNA to 10-nt PNA with even higher melting temperatures (*T*m), hence increasing stability and minimizing the modifications on the RNA of interest. The following year, Egloff *et al.* reported covalent labeling of L4 through reactive group transfer.[63] However, the method suffered from low efficiencies as the secondary structures obstruct internal labeling. RNA termini can be targeted when internal positions are buried and difficult to reach. Ahunbay *et al.* have recently developed a method to efficiently dual-end label large ribozymes for ensemble and single-molecule fluorescence and showed its utility on the wild-type *Sc*.ai5γ group II intron flanked between exons.[60] This strategy targets the two chemically different features of virtually any RNA molecule, the 5'-phosphate and the 3'-ribose. The RNA ends are first activated and then bioconjugated with a FRET pair of fluorescent dyes.

Other fluorescence studies of group II introns used labeled smaller sections of the intron, substrates or splicing products that circumvented to label the ribozyme. For example, in 2020, Steffen *et al.* dissected the EBS1-IBS1 interaction with smFRET, labeling a short intronic sequence with a donor and the cognate exonic sequence with an acceptor dye.[64] In this way, the stabilization effect of Mg^{2+} on the intron-exon recognition of RNA-RNA and RNA-DNA was demonstrated. The same year, Smathers and Robart used a dual-fluorescently labeled DNA target to study the intron-encoded protein (IEP)-dependent splicing and retrohoming of bacterial group IIC intron *Ta.it*.I1 from *Thermoanerobacter italicus*.[65]

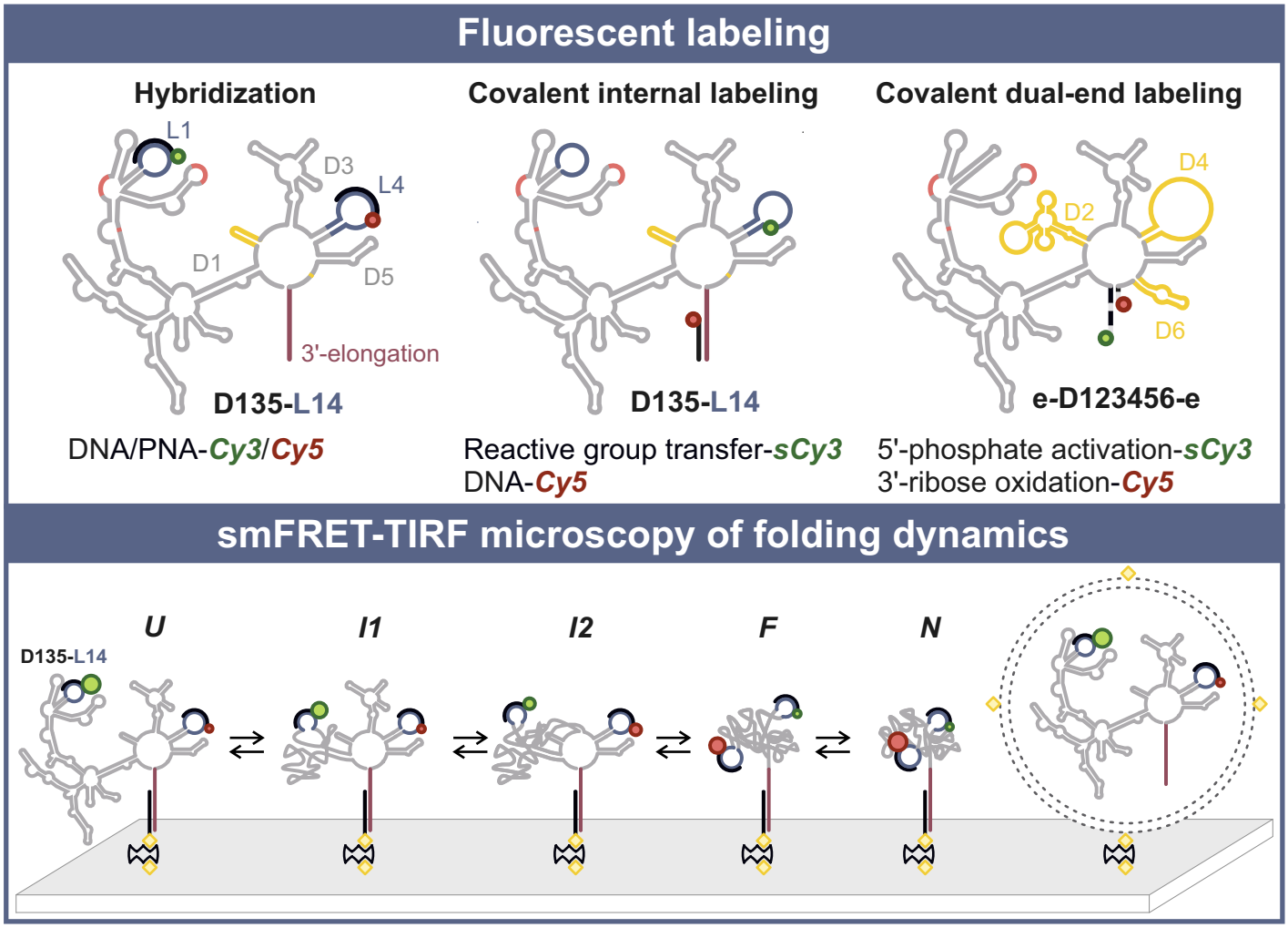

Fig. 4. Group II intron folding dynamics brought to light by single-molecule fluorescence. The top panel shows different site-specific fluorescent labeling strategies of *Sc*.ai5γ group II intron (gray) constructs. By *hybridization* minimally active D135-L14 is labeled through loops on D1 and D4 (blue) that are introduced as labeling platforms, where fluorescent DNA/PNA probes can hybridize. *Covalent internal labeling* of L4 is achieved by functionalization with the donor fluorophore that is covalently attached through reactive group transfer, while the acceptor fluorophore is introduced onto the elongated 3'-end (dark pink) through hybridization. In *covalent dual-end labeling*, the 5'- and 3'-ends of the full-length RNA with flanking exons (e-D123456-e) are chemically activated to attach two different fluorophores covalently. Modifications (on D135) and their original counterparts (on full-length intron) are depicted in yellow. Exon-binding sites (EBS, pink) and intron-binding sites (IBS, black) are complementary. Bottom panel: The folding pathway of D135-L14 (unfolded state, *U* ↔ intermediates, *I1* and *I2* ↔ folded state, *F* ↔ native state, *N*) as observed by single-molecule FRET is shown. The RNA is labeled through hybridization. TIRF microscopy requires surface immobilization. The intron is directly tethered on its elongated 3'-end that hybridizes to a biotinylated DNA oligonucleotide. Alternatively, the intron is encapsulated into a biotinylated phospholipid vesicle (not to scale) that is surface tethered.

Molecular beacons found an elegant use in probing group II intron splicing of pathogen *Histoplasma capsulatum*.[66] In this example, the molecular beacon, which is complementary to the exon ligation product, forms a stem-loop bringing the attached fluorophore near its quencher. Once the exons ligate, hybridization opens the molecular beacon, which restores the fluorescence of the quenched fluorophore.

4.2 *One Group II Intron at a Time*

One way to study RNA folding in single-molecule fluorescence is by total internal reflection fluorescence (TIRF) microscopy. TIRF generates an evanescent field that excites only the molecules that are very close (<200 nm) to the surface. This way, the background signal is drastically reduced, which allows observing single molecules over time. TIRF microscopy requires surface immobilization of the molecule of interest (Fig. 4). A standard procedure to surface-immobilize group II intron derivatives is the attachment *via* the elongated 3'-end through the hybridization of a biotinylated ssDNA, which in turn binds to a streptavidin-coated surface.[33,67–69] Surface immobilization always bears the risk that the corresponding molecule is hindered in its folding and function. Encapsulating group II introns in biotinylated phospholipid vesicles (~200 nm diameter)[62] abolishes the concern of impeding structural reorganization. The vesicle provides a closed environment where the RNA freely diffuses without restraints while keeping all splicing products within the evanescent field.

In the first smFRET studies on the group II intron, Steiner *et al.* showed, by dwell time analysis, that the transitions between FRET states gained speed with increasing ionic strength at room temperature.[33] The authors also observed a stabilization effect of the 17/7 substrate added in *trans* on the native state. The nature of divalent cations is decisive: Ca^{2+} (larger than Mg^{2+}) inhibits activity.[70] A follow-up smFRET study showed the folding dynamics in response to Ca^{2+}, where two distinct subpopulations emerged.[71] Furthermore, smFRET studies have also demonstrated that the cofactor Mss116,[67] as well as crowding agents such as polyethylene glycol (10% PEG 8000, w/v),[69,72] stabilize the native fold

of the group II intron. In turn, the metal ion concentration and temperature can be lowered closer to physiological conditions.

5. Conclusion

After four decades of experiments, group II intron studies withstand the test of time, and they certainly offer more to unravel. This review revolves around the *Sc*.ai5γ group IIB intron of yeast mitochondria and the studies to uncover its structure–function relationship. The large size and high structure of the group II introns have been the major obstacle, which compelled novel and multidisciplinary approaches to tackle the RNA. Conventional methods (chemical and enzymatic footprinting, NMR, and X-ray crystallography) provided a structural understanding of catalysis and folding. Dramatic structural rearrangements are best captured with single-molecule methods, such as cryo-EM and smFRET. The predictive power of molecular dynamics (MD) corroborates with cryo-EM[73] and smFRET.[64] The future promises the visualization of group II intron dynamics.

Acknowledgements

EA is grateful to the Swiss Chemical Society and DSM Nutritional Products AG for the generous award of 'best poster presentation in Chemical Biology'. Financial support from the Swiss National Science Foundation [200020_192153, to SZP and RKOS], the University of Zurich, the UZH Forschungskredit [FK-20-081 to EA], and the Graduate School of Chemical and Molecular Sciences Zurich (CMSZH) is gratefully acknowledged.

Received: February 4, 2023

[1] S. Zimmerly, C. Semper, *Mob. DNA* **2015**, *6*, 7, https://doi.org/10.1186/s13100-015-0037-5.
[2] A. M. Pyle, *Annu. Rev. Biophys*. **2016**, *45*, 183, https://doi.org/10.1146/annurev-biophys-062215-011149.
[3] Y. Vallès, K. M. Halanych, J. L. Boore, *PLoS ONE* **2008**, *3*, e1488, https://doi.org/10.1371/journal.pone.0001488.
[4] O. Fedorova, G. E. Jagdmann, R. L. Adams, L. Yuan, M. C. van Zandt, A. M. Pyle, *Nat. Chem. Biol*. **2018**, *14*, 1073, https://doi.org/10.1038/s41589-018-0142-0.
[5] T. Liu, A. M. Pyle, *Nucleic Acids Res*. **2021**, *49*, 12422, https://doi.org/10.1093/nar/gkab1077.
[6] A. C. Arnberg, G. J. van Ommen, L. A. Grivell, E. F. van Bruggen, P. Borst, *Cell* **1980**, *19*, 313, https://doi.org/10.1016/0092-8674(80)90505-x.
[7] R. van der Veen, A. C. Arnberg, G. van der Horst, L. Bonen, H. F. Tabak, L. A. Grivell, *Cell* **1986**, *44*, 225, https://doi.org/10.1016/0092-8674(86)90756-7.
[8] M. P. Robertson, G. F. Joyce, *Cold Spring Harb. Perspect. Biol*. **2012**, *4*, https://doi.org/10.1101/cshperspect.a003608.
[9] 'The RNA world: The nature of modern RNA suggests a prebiotic RNA World', Eds. R. F. Gesteland, T. R. Cech, J. F. Atkins, Cold Spring Harbor Laboratory, New York, **2006**.
[10] T. R. Cech, A. J. Zaug, P. J. Grabowski, *Cell* **1981**, *27*, 487, https://doi.org/10.1016/0092-8674(81)90390-1.
[11] K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, T. R. Cech, *Cell* **1982**, *31*, 147, https://doi.org/10.1016/0092-8674(82)90414-7.
[12] 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.
[13] T. R. Cech, *Biosci. Rep*. **1990**, *10*, 239, https://doi.org/10.1007/BF01117241.
[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] R. Cojocaru, P. J. Unrau in 'Ribozymes', Eds. S. Müller, B. Masquida, W. Winkler, Wiley-VCH, **2021**.
[16] T. A. Steitz, P. B. Moore, *Trends Biochem. Sci*. **2003**, *28*, 411, https://doi.org/10.1016/S0968-0004(03)00169-5.
[17] C. M. Smathers, A. R. Robart, *Biochim. Biophys. Acta Gene Regul. Mech.* **2019**, *1862*, 194390, https://doi.org/10.1016/j.bbagrm.2019.06.001.
[18] M. Podar, V. T. Chu, A. M. Pyle, P. S. Perlman, *Nature* **1998**, *391*, 915, https://doi.org/10.1038/36142.
[19] A. M. Lambowitz, S. Zimmerly, *Annu. Rev. Genet.* **2004**, *38*, 1, https://doi.org/10.1146/annurev.genet.38.072902.091600.
[20] A. M. Lambowitz, S. Zimmerly, *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, a003616, https://doi.org/10.1101/cshperspect.a003616.
[21] P. J. Enyeart, G. Mohr, A. D. Ellington, A. M. Lambowitz, *Mob. DNA* **2014**, *5*, 2, https://doi.org/10.1186/1759-8753-5-2.
[22] E. J. Sontheimer, P. M. Gordon, J. A. Piccirilli, *Genes. Dev*. **1999**, *13*, 1729, https://doi.org/10.1101/gad.13.13.1729.
[23] K. A. Jarrell, R. C. Dietrich, P. S. Perlman, *Mol. Cell Biol*. **1988**, *8*, 2361, https://doi.org/10.1128/mcb.8.6.2361-2366.1988.
[24] W. J. Michels, A. M. Pyle, *Biochemistry* **1995**, *34*, 2965, https://doi.org/10.1021/bi00009a028.
[25] L. J. Su, C. Waldsich, A. M. Pyle, *Nucleic Acids Res*. **2005**, *33*, 6674, https://doi.org/10.1093/nar/gki973.
[26] O. H. Gumbs, R. A. Padgett, K. T. Dayie, *RNA* **2006**, *12*, 1693, https://doi.org/10.1261/rna.137006.
[27] A. R. Robart, R. T. Chan, J. K. Peters, K. R. Rajashankar, N. Toor, *Nature* **2014**, *514*, 193, https://doi.org/10.1038/nature13790.
[28] J. Swisher, C. M. Duarte, L. J. Su, A. M. Pyle, *EMBO J*. **2001**, *20*, 2051, https://doi.org/10.1093/emboj/20.8.2051.
[29] M. Boudvillain, A. M. Pyle, *EMBO J.* **1998**, *17*, 7091, https://doi.org/10.1093/emboj/17.23.7091.
[30] M. Marcia, A. M. Pyle, *Cell* **2012**, *151*, 497, https://doi.org/10.1016/j.cell.2012.09.033.
[31] M. Marcia, S. Somarowthu, A. M. Pyle, *Mob. DNA* **2013**, *4*, 14, https://doi.org/10.1186/1759-8753-4-14.
[32] C. Waldsich, A. M. Pyle, *J. Mol. Biol*. **2008**, *375*, 572, https://doi.org/10.1016/j.jmb.2007.10.052.
[33] M. Steiner, K. S. Karunatilaka, R. K. O. Sigel, D. Rueda, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 13853, https://doi.org/10.1073/pnas.0804034105.
[34] A. M. Pyle, O. Fedorova, C. Waldsich, *Trends Biochem. Sci*. **2007**, *32*, 138, https://doi.org/10.1016/j.tibs.2007.01.005.
[35] M. W. Müller, R. J. Schweyen, C. Schmelzer, *Nucleic Acids Res*. **1988**, *16*, 7383, https://doi.org/10.1093/nar/16.15.7383.
[36] M. Costa, F. Michel, E. Westhof, *EMBO J*. **2000**, *19*, 5007, https://doi.org/10.1093/emboj/19.18.5007.
[37] D. L. Daniels, W. J. Michels, A. M. Pyle, *J. Mol. Biol*. **1996**, *256*, 31, https://doi.org/10.1006/jmbi.1996.0066.
[38] B. Fazliji, C. Ferreira Rodrigues, H. Wang, R. K. O. Sigel in 'Reference Module in Chemistry, Molecular Sciences and Chemical Engineering', Eds. J. Reedijk, K. Poepelmeier, Elsevier, **2022**.
[39] J. F. Swisher, L. J. Su, M. Brenowitz, V. E. Anderson, A. M. Pyle, *J. Mol. Biol*. **2002**, *315*, 297, https://doi.org/10.1006/jmbi.2001.5233.
[40] M. C. Erat, O. Zerbe, T. Fox, R. K. O. Sigel, *ChemBioChem* **2007**, *8*, 306, https://doi.org/10.1002/cbic.200600459.
[41] M. Furler, B. Knobloch, R. K. O. Sigel, *Inorganica Chim. Acta* **2009**, *362*, 771, https://doi.org/10.1016/j.ica.2008.03.095.
[42] D. Kruschel, M. Skilandat, R. K. O. Sigel, *RNA* **2014**, *20*, 295, https://doi.org/10.1261/rna.041137.113.
[43] S. Bartova, M. Pechlaner, D. Donghi, R. K. O. Sigel, *J. Biol. Inorg. Chem*. **2016**, *21*, 319, https://doi.org/10.1007/s00775-016-1341-3.
[44] N. Toor, K. S. Keating, S. D. Taylor, A. M. Pyle, *Science* **2008**, *320*, 77, https://doi.org/10.1126/science.1153803.
[45] N. Toor, K. S. Keating, O. Fedorova, K. Rajashankar, J. Wang, A. M. Pyle, *RNA* **2010**, *16*, 57, https://doi.org/10.1261/rna.1844010.
[46] T. Wiryaman, N. Toor, *Methods* **2017**, *125*, 10, https://doi.org/10.1016/j.ymeth.2017.06.020.
[47] S. Somarowthu, M. Legiewicz, K. S. Keating, A. M. Pyle, *Nucleic Acids Res.* **2014**, *42*, 1947, https://doi.org/10.1093/nar/gkt1051.
[48] T. Wiryaman, N. Toor in 'Ribozymes, Methods Mol. Biol.', Eds. R. J. Scarborough, A. Gatignol, **2021**.
[49] O. Fedorova, C. Waldsich, A. M. Pyle, *J. Mol. Biol*. **2007**, *366*, 1099, https://doi.org/10.1016/j.jmb.2006.12.003.
[50] H.-R. Huang, C. E. Rowe, S. Mohr, Y. Jiang, A. M. Lambowitz, P. S. Perlman, *Proc. Natl. Acad. Sci. U.S.A*. **2005**, *102*, 163, https://doi.org/10.1073/pnas.0407896101.
[51] O. Fedorova, A. Solem, A. M. Pyle, *J. Mol. Biol*. **2010**, *397*, 799, https://doi.org/10.1016/j.jmb.2010.02.001.
[52] O. Fedorova, A. M. Pyle, *J. Mol. Biol*. **2012**, *422*, 347, https://doi.org/10.1016/j.jmb.2012.05.037.
[53] S. M. Fica, K. Nagai, *Nat. Struct. Mol. Biol*. **2017**, *24*, 791, https://doi.org/10.1038/nsmb.3463.
[54] Z. Su, K. Zhang, K. Kappel, S. Li, M. Z. Palo, G. D. Pintilie, R. Rangan, B. Luo, Y. Wei, R. Das, W. Chiu, *Nature* **2021**, *596*, 603, https://doi.org/10.1038/s41586-021-03803-w.
[55] X. Zhang, S. Li, G. Pintilie, M. Z. Palo, K. Zhang, *Nucleic Acids Res*. **2023**, https://doi.org/10.1093/nar/gkac1268.
[56] G. Qu, P. S. Kaushal, J. Wang, H. Shigematsu, C. L. Piazza, R. K. Agrawal, M. Belfort, H.-W. Wang, *Nat. Struct. Mol. Biol*. **2016**, *23*, 549, https://doi.org/10.1038/nsmb.3220.
[57] N. Liu, X. Dong, C. Hu, J. Zeng, J. Wang, J. Wang, H.-W. Wang, M. Belfort, *Nucleic Acids Res*. **2020**, *48*, 11185, https://doi.org/10.1093/nar/gkaa773.
[58] K. Chung, L. Xu, P. Chai, J. Peng, S. C. Devarkar, A. M. Pyle, *Science* **2022**, *378*, 627, https://doi.org/10.1126/science.abq2844.
[59] M. Götz, A. Barth, S. S.-R. Bohr, R. Börner, J. Chen, T. Cordes, D. A. Erie, C. Gebhardt, M. C. A. S. Hadzic, G. L. Hamilton, N. S. Hatzakis, T. Hugel, L. Kisley, D. C. Lamb, C. de Lannoy, C. Mahn, D. Dunukara, D. de Ridder, H. Sanabria, J. Schimpf, C. A. M. Seidel, R. K. O. Sigel, M. B. Sletfjerding,

J. Thomsen, L. Vollmar, S. Wanninger, K. R. Weninger, P. Xu, S. Schmid, *Nat. Commun*. **2022**, *13*, 5402, https://doi.org/10.1038/s41467-022-33023-3.
[60] E. Ahunbay, F. D. Steffen, S. Zelger-Paulus, R. K. O. Sigel in 'DNAzymes, Methods Mol. Biol.', Eds. G. Steger, H. Rosenbach, I. Span, Humana, New York, NY, **2022**, p. 191–204.
[61] A. G. Schmitz, S. Zelger-Paulus, G. Gasser, R. K. O. Sigel, *ChemBioChem* **2015**, *16*, 1302, https://doi.org/10.1002/cbic.201500180.
[62] S. Zelger-Paulus, M. C. A. S. Hadzic, R. K. O. Sigel, R. Börner in 'RNA Spectroscopy, Methods Mol. Biol.', Eds. V. Arluison, F. Wien, Humana, New York, NY, **2020**.
[63] D. Egloff, I. A. Oleinich, M. Zhao, S. L. B. König, R. K. O. Sigel, E. Freisinger, *ACS Chem. Biol*. **2016**, *11*, 2558, https://doi.org/10.1021/acschembio.6b00343.
[64] F. D. Steffen, M. Khier, D. Kowerko, R. A. Cunha, R. Börner, R. K. O. Sigel, *Nat. Commun*. **2020**, *11*, 104, https://doi.org/10.1038/s41467-019-13683-4.
[65] C. M. Smathers, A. R. Robart, *RNA* **2020**, *26*, 664, https://doi.org/10.1261/rna.075044.120.
[66] Q. Q. Omran, O. Fedorova, T. Liu, A. M. Pyle, *Nucleic Acids Res*. **2022**, https://doi.org/10.1093/nar/gkac242.
[67] K. S. Karunatilaka, A. Solem, A. M. Pyle, D. Rueda, *Nature* **2010**, *467*, 935, https://doi.org/10.1038/nature09422.
[68] L. Cardo, K.S. Karunatilaka, D. Rueda, R.K. O. Sigel in 'Ribozymes, Methods Mol. Biol.', Ed. J. S. Hartig, Humana, New York, NY, **2012**.
[69] B. P. Paudel, E. Fiorini, R. Börner, R. K. O. Sigel, D. S. Rueda, *Proc. Natl. Acad. Sci. U.S.A*. **2018**, *115*, 11917, https://doi.org/10.1073/pnas.1806685115.
[70] M. C. Erat, R. K. O. Sigel, *J. Biol. Inorg. Chem*. **2008**, *13*, 1025, https://doi.org/10.1007/s00775-008-0390-7.
[71] M. Steiner, D. Rueda, R. K. O. Sigel, *Angew. Chem. Int. Ed. Engl*. **2009**, *48*, 9739, https://doi.org/10.1002/anie.200903809.
[72] E. Fiorini, R. Börner, R. K. O. Sigel, *CHIMIA* **2015**, *69*, 207, https://doi.org/10.2533/chimia.2015.207.
[73] M. Marcia, J. Manigrasso, M. de Vivo, *J. Chem. Inf. Model*. **2021**, *61*, 2511; https://doi.org/10.1021/acs.jcim.1c00572.
[74] C. L. Peebles, P. S. Perlman, K. L. Mecklenberg, M. L. Petrillo, J. H. Tabor, K. A. Jarrell, H.-L. Cheng, *Cell* **1986**, *44*, 213, https://doi.org/10.1016/0092-8674(86)90755-5.

License and Terms

The definitive version of this article is the electronic one that can be found at https://doi.org/10.2533/chimia.2023.235