


 Cite this: *RSC Adv.*, 2022, 12, 5398

# Influence of sequence variation on the RNA cleavage activity of Zn<sup>2+</sup>-dimethyl-dppz-PNA-based artificial enzymes†

 Olivia Luige, <sup>a</sup> Kristina Karalé, <sup>ab</sup> Partha Pratim Bose, <sup>a</sup> Martin Bollmark,<sup>b</sup> Ulf Tedebark,<sup>b</sup> Merita Murtola<sup>a</sup> and Roger Strömberg <sup>\*a</sup>

The development of Zn<sup>2+</sup>-dependent dimethyl-dppz-PNA conjugates (PNAzymes) as efficient site-specific artificial ribonucleases enables rapid sequence-specific degradation of clinically relevant RNA target sequences, but the significance of the RNA/PNAzyme sequence and structural demands for the identification of novel RNA targets are not fully understood. In the present study, we investigated the influence of sequence variation in the recognition arms of the RNA/PNAzyme complex on the RNA cleavage activity of the artificial enzymes. The base pairs closing the 3-nucleotide bulge region on both sides of the bulge as well as the neighbouring nucleobases were shown to significantly influence the RNA cleavage activity. Elongation of the RNA/PNAzyme complex was shown to be tolerated, although potentially prohibitive for catalytic turnover. The specificity of PNAzyme action was clearly demonstrated by the significantly reduced or absent cleavage activity in complexes containing mismatches. Further investigation into 2- and 4-nucleotide RNA bulges indicated that formation of 3-nucleotide bulges in the target RNA gives the optimal cleavage rates, while some potential off-target cleavage of formed 4-nucleotide bulges of select sequences should be considered.

 Received 12th November 2021  
 Accepted 30th January 2022

 DOI: 10.1039/d1ra08319h  
[rsc.li/rsc-advances](http://rsc.li/rsc-advances)

## Introduction

Nucleic acid manipulation is important in molecular biology<sup>1</sup> as well as for therapeutic interventions in the clinic.<sup>2,3</sup> The development of efficient artificial enzymes capable of sequence-specific cleavage of RNA has been a long-standing goal of nucleic acid chemistry.<sup>4–8</sup> Such artificial ribonucleases have the potential to achieve degradation of disease-related RNA targets without the assistance of endogenous enzymes, thereby offering an alternative to gapmer antisense oligonucleotides (ASOs) and small interfering RNAs (siRNAs),<sup>3,9</sup> both of which are used in the clinic today.

Artificial ribonucleases of various designs have been reported, including both metal-free<sup>10–13</sup> and metal cofactor-dependent systems.<sup>14–22</sup> The self-reliance of metal independent systems is attractive,<sup>10</sup> although higher RNA cleavage rates have been accomplished with metal-dependent artificial nucleases.<sup>16,18,19</sup> Zn<sup>2+</sup> is an attractive cofactor due to its relative biocompatibility, and has been used in artificial ribonucleases

developed with both 2'-*O*-methyl RNA<sup>14,20,22</sup> and PNA backbones.<sup>17,21</sup> However, Zn<sup>2+</sup>-dependent artificial ribonucleases have typically shown inferior RNA cleavage kinetics and lower site-specificity<sup>14,17,20–22</sup> compared to Cu<sup>2+</sup>-dependent peptide nucleic acid–neocuproine conjugates (Cu<sup>2+</sup>-neocuproine PNAzymes).<sup>16,18,19</sup> Nonetheless, the recently reported novel Zn<sup>2+</sup>-dependent dimethyl-dipyridophenazine (dppz)-based PNAzymes are a breakthrough, as they accomplish rapid RNA cleavage with down to 16 min half-lives at pH 7.0 and 10 min at pH 7.4.<sup>15</sup> In addition, these artificial enzymes can be tailored to disease-related RNA targets.<sup>15</sup> These PNAzymes are designed to be partially complementary to the RNA target sequences in order to force the formation of a bulge in the RNA adjacent to the dimethyl-dppz “molecular scissors”, thus taking advantage of the inherent reactivity of phosphodiester linkages within single-stranded RNA bulges.<sup>23,24</sup> Fast site-specific cleavage was shown for 3-nucleotide bulge-forming RNAs, including clinically relevant RNA target sequences (*i.e.* a malaria parasite *Plasmodium falciparum* mRNA model target and a SARS-CoV-2 genomic RNA model target).<sup>15</sup> While this significant development has successfully demonstrated rapid RNA cleavage of clinically relevant RNA models, the extent to which the target sequence can be modified while still maintaining efficient cleavage activity is still not fully understood. A thorough understanding of the intricacies of the sequence–activity relationships in the RNA/PNAzyme complexes can shed light on various aspects of PNAzyme action, while also being critical for

<sup>a</sup>Department of Biosciences and Nutrition, Karolinska Institutet, Neo, 141 83 Huddinge, Sweden. E-mail: roger.stromberg@ki.se

<sup>b</sup>RISE, Department of Chemical Process and Pharmaceutical Development, Forskargatan 18, 15136 Södertälje, Sweden

† Electronic supplementary information (ESI) available: PNAzyme purity analysis, IEX-HPLC chromatograms and ESI-MS data for the presented RNA cleavage as well as CD spectra of RNA/PNAzyme complexes. See DOI: 10.1039/d1ra08319h



the effective identification of novel RNA targets and for understanding possible off-target effects. The present study takes a particular interest in the bulge-closing base pairs on both sides of the bulge, as well as the sequence and length of the shorter recognition arm of the RNA/PNAzyme complex. Moreover, in addition to the cleavage of 3-nucleotide RNA bulges, which have been studied in detail,<sup>15</sup> we demonstrate the cleavage of 2-nucleotide RNA bulges and study the sequence-dependence in 4-nucleotide bulges.

## Results and discussion

### Sequence–activity relationships in the long recognition arm

The reported RNA/PNAzyme complexes (Fig. 1) contain a CG base pair closing the bulge in the longer hybridised stem, *i.e.* the long recognition arm.<sup>15</sup> In such systems, the AAA bulge is cleaved site-specifically with an approximately 1 h half-life (Fig. 1c).<sup>15</sup> Extension of the long recognition arm by one base pair gives essentially the same cleavage rate (Fig. 2, RNA 2 and PNAzyme II). In order to further investigate the role of the bulge-closing base pair in the long recognition arm, we constructed novel systems where variations were introduced to the RNA 2/PNAzyme II sequence (Fig. 2). Four PNA-dimethyl-dppz conjugates were compared (PNAzymes II–V, Fig. 2), where AAA bulges are formed in the corresponding RNA targets (RNAs 2–5, Fig. 2), while the bulge-closing base pair is varied in each complex.

The cleavage of RNA 2, where the bulge was closed with a CG pair, gave the highest cleavage rate while the reversed GC bulge-closing base pair (RNA 3/PNAzyme III) showed slightly lower activity. Rather strikingly, the cleavage activity suffered a major loss when the AAA bulge was closed with a UA (RNA 5/PNAzyme V) or AT base pair (RNA 4/PNAzyme IV), where the RNA cleavage half-lives were estimated to be in the 4 and 11 h range, respectively (Fig. 2). MS analysis confirmed that the cleavage site was between the second and third nucleotide in the bulge

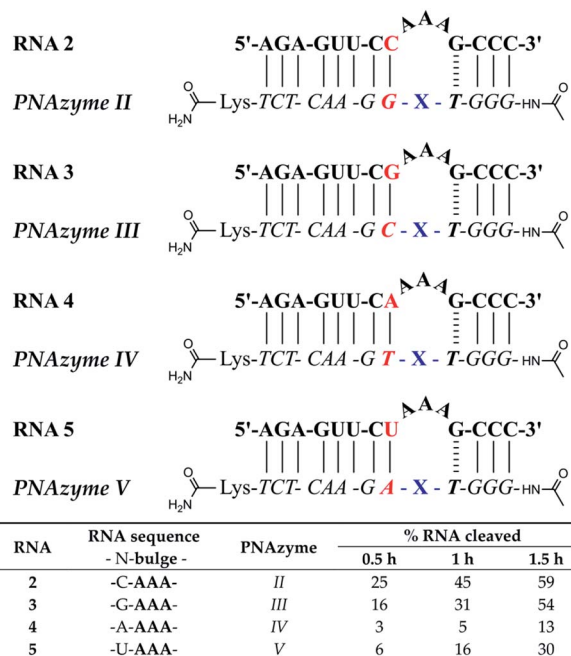


Fig. 2 Schematic representations of RNA/PNAzyme complexes where the bulge-closing base pair in the long recognition arm varies, followed by the extent of RNA cleavage observed in each complex after incubation of the RNA/PNAzyme complex (4  $\mu$ M) in the presence of  $Zn^{2+}$  (100  $\mu$ M) at 37  $^{\circ}$ C, pH 7. The structure of “X” (the “molecular scissors”) is shown in Fig. 1b.

(5'-AA/A-3'), while the least efficiently cleaved RNA 4 also showed significant cleavage between the third bulge nucleotide and the GT wobble (5'-AAA/G-3') (see ESI-S3† for IE-HPLC chromatograms and MS data).

The overall dependence of RNA cleavage activity on the bulge-closing base pairs (most to least favourable CG > GC > UA > AT) suggests that the strength of hydrogen bonding is key. Moreover, there seems to be an additional benefit if the pyrimidine is in the RNA and the purine in the PNA strand. Herein as well as in prior reports,<sup>15</sup> the nucleobase preceding the bulge-closing base pair has been either of the two pyrimidine bases in the RNA and the complementary purine in the PNA strand. As such, it should be carefully considered in the selection of novel RNA targets in future investigations whether instead of UC or CC (pyrimidine-C), perhaps AG or GG (purine-G) sequence would be tolerated equally well in the bulge-closing region of the long recognition arm. It should be noted that the activity of more complex triplex-forming  $Cu^{2+}$ -neocuproine PNAzymes has indeed been higher with purine-purine (AG) rather than purine-pyrimidine (AC) sequence in the RNA target in the same bulge-closing region.<sup>18</sup>

The key role of the bulge-closing base pair suggests it has a profound influence on the structure of the bulge. Indeed, bulges are known to be diverse structural motifs with nucleotide arrangements defined by the competing interactions of the bulge nucleotides and the surrounding hybridised base pairs.<sup>25</sup> The bulge nucleobases can point out into the solvent, fold back onto adjacent hydrophobic surfaces, or they may take part in

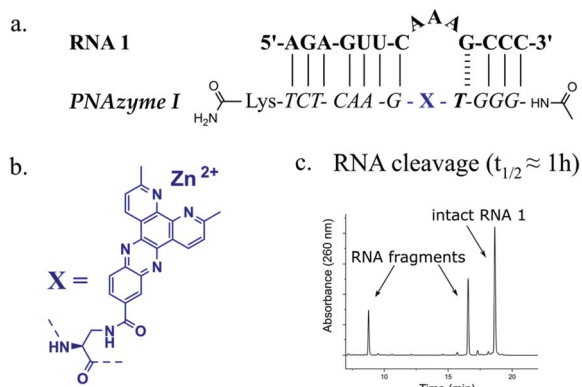


Fig. 1 (a) Schematic representation of the previously reported RNA target 1 that forms a 3-nucleotide AAA bulge when in complex with PNAzyme I (b) structure of the dimethyl-dipyridophenazine (dppz) “molecular scissors” moiety which was conjugated to PNA to create  $Zn^{2+}$  dimethyl-dppz PNAzymes (c) previously reported (by Luige *et al.*)<sup>15</sup> IEX-HPLC chromatogram showing the extent of cleavage of RNA 1 by  $Zn^{2+}$ -dependent PNAzyme I after 50 min of incubation at 37  $^{\circ}$ C, pH 7 in the presence of 100  $\mu$ M  $Zn^{2+}$ .

the continuous stacking of the surrounding double strands, favouring kinking of the helix.<sup>25</sup> As such, the bulge-closing base pair is clearly changing the structural arrangement, which is also evident in the comparison of the circular dichroism (CD) spectra of these four complexes (see ESI-S4†).

Generally, PNAzyme-promoted cleavage of RNA has been shown to exhibit characteristic bulge-sequence dependence.<sup>15–17,19</sup> For RNA/Zn<sup>2+</sup>-dimethyl-dppz PNAzyme complexes where the recognition arms have the sequence shown for RNA 1/PNAzyme I in Fig. 1, changing the RNA bulge sequence from AAA (Fig. 1) to AUA, GUA or UUA enhances the cleavage rate, corresponding to a shift in the half-life from 1 h to 16 min.<sup>15</sup> In order to determine whether similar sequence dependence could occur in the less efficient systems where the bulge is closed with an AT or UA base pair in the long recognition arm, we investigated RNA targets with other bulge sequences in these systems. Interestingly, complexes with the AT pair did not exhibit the characteristic bulge sequence-dependence to a significant extent. Some variation was observed in the extent of cleavage (with adenosine being favoured over uridine in the third position), but the activity remained very low for all five bulge sequences (RNAs 4, 6–9, Fig. 3 and ESI-S3†). Nonetheless, complexes with the UA bulge-closing base pair displayed some bulge sequence-dependence (RNAs 5 and 10–12, Fig. 3 and ESI-S3†). The fastest rate was observed for the AUA bulge (RNA 11), where the cleavage took place at a single site (5'-AU/A-3', see ESI-S3†) with an estimated half-life in the 1.5 h range, which is an improvement from the approximately 4 h half-life reported above for the AAA bulge closed with the same base pair.

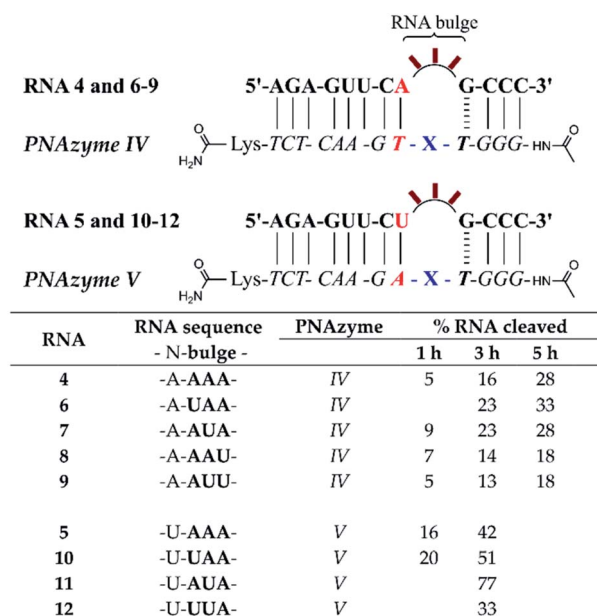


Fig. 3 Schematic representations of RNA/PNAzyme complexes where the 3-nucleotide RNA bulge sequence is varied in complexes with either an AT or UA bulge-closing base pair in the long recognition arm, followed by the extent of RNA cleavage observed in each complex after incubation of the RNA/PNAzyme complex (4  $\mu$ M) in the presence of Zn<sup>2+</sup> (100  $\mu$ M) at 37 °C, pH 7. The structure of "X" (the "molecular scissors") is shown in Fig. 1b.

However, the overall cleavage activity of the systems where 3-nucleotide bulges were closed with AT or UA base pairs was significantly lower than the previously reported 16 min half-lives for the systems with a CG bulge-closing base pair.<sup>15</sup>

To further probe the activity of complexes with the weaker bulge-closing base pairs, we investigated whether faster cleavage rates could be accomplished by targeting smaller 2-nucleotide bulges (Fig. 4). Notably, a significantly improved cleavage rate was indeed observed for the AT bulge-closing base pair when the RNA formed a UU bulge (RNA 15). In this case the cleavage was site-specific (5'-UU/G-3', see ESI-S3†) with an estimated half-life in the 1.5 h range.

Comparison of the cleavage of the 2-nucleotide UA bulges closed with different base pairs showed that the RNA/PNAzyme complexes which contained the CG and GC bulge-closing base pairs (RNAs 18 and 19) once again outperformed those with the AT and UA pairs (RNAs 13 and 16, Fig. 4), although their activity was significantly higher with 3-nucleotide bulges. In the case of all 2-nucleotide UA bulge-forming systems regardless of the bulge-closing base pair, cleavage fragments originating from two cleavage sites (5'-U/A-3' and 5'-UA/G-3') were detected (see ESI-S3†). These results suggest that 3-nucleotide RNA bulges closed with CG or GC base pairs in the long recognition arm are

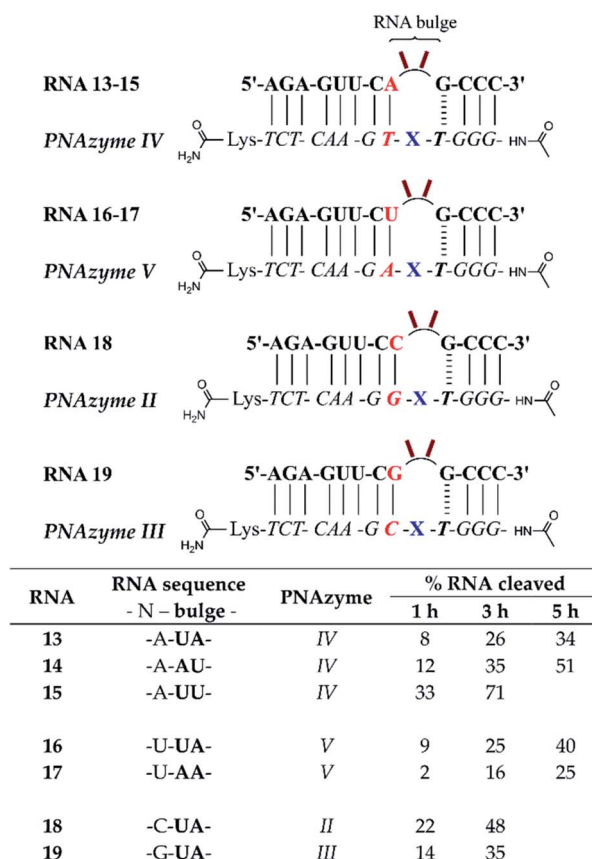


Fig. 4 Schematic representations of RNA/PNAzyme complexes where 2-nucleotide RNA bulges are formed, followed by the extent of RNA cleavage observed in each complex after incubation of the RNA/PNAzyme complex (4  $\mu$ M) in the presence of Zn<sup>2+</sup> (100  $\mu$ M) at 37 °C, pH 7. The structure of "X" (the "molecular scissors") is shown in Fig. 1b.

the optimal targets since these systems give the highest overall cleavage rates. Moreover, the sequence dependence in 2-nucleotide bulges closed with a CG or GC base pair could be a subject for future investigations.

### Sequence–activity relationships in the short recognition arm

The short recognition arm of the RNA/PNAzyme complex comprises a GT wobble, followed by three CG Watson–Crick

base pairs. The replacement of the GT wobble with a GC Watson–Crick base pair in Cu<sup>2+</sup>-neocuproine PNAzymes decreased the RNA cleavage activity substantially.<sup>19</sup> Here, we synthesised a Zn<sup>2+</sup>-dimethyl dppz PNAzyme designed to form a GC base pair instead of the GT wobble with RNA 13 (Fig. 5). The replacement of the GT wobble with a GC base pair (RNA 13/PNAzyme VI) significantly lowered the RNA cleavage rate (Fig. 5). A profound structural change was also evident in the comparison of the CD spectra of these complexes (ESI-S4†).

The position next to the bulge-closing base pair in the short recognition arm was then modified to include a GC instead of the CG base pair next to the GT wobble (RNA 20/PNAzyme VII). Interestingly, this change in the position next to the bulge-closing base pair in the short recognition arm led to a slightly lower cleavage rate. As such, the GC sequence in RNA 13 closing the bulge in the short recognition arm was favoured over GG in RNA 20. It remains unclear whether the strength of hydrogen bonding in the position next to the bulge-closing base pair is critical, or perhaps GU (purine–pyrimidine) would be tolerated equally well, hence allowing the possibility of a UA base pair in the RNA/PNAzyme complex next to the bulge-closing GT wobble.

Moreover, a further change was then introduced by replacing the GT wobble with an AT bulge-closing base pair, which was then followed by a GC base pair (RNA 6/PNAzyme VII). A further decrease in the RNA cleavage rate was observed. The cleavage still took place at a single site between the second and third nucleotide in the bulge (5'-AU/A-3'; see ESI-S3†).

In summary, the bulge-closing base pair in the short recognition arm had an influential role on the RNA cleavage activity. The GT wobble was clearly preferred, while the AT base pair was shown to be detrimental, yet tolerated more than a GC base pair.

The significant effects of altering the sequence in the immediate vicinity of the bulge led us to investigate the influence of the terminal base pair in the short recognition arm. We synthesised PNAzymes VIII–X (Fig. 5), which preserve the first three base pairs in the short recognition arm but form different terminal base pairs with RNAs 21–23. In summary the alteration of the terminal base gave almost the same cleavage rates as with the CG pair in RNA13/PNAzyme I, which perhaps is not unexpected since terminal base pairs are known to be more flexible.

### Elongation of the RNA/PNAzyme complex and catalytic turnover

Apart from altering the sequence of the recognition arms in the RNA/PNAzyme complex, elongation of the complex is another alteration that could potentially have implications for the cleavage efficacy. Additionally, if extension of the sequence is tolerated, then this may aid in structural studies (*e.g.* crystallography with a non-cleavable modification at the cleavage site) since the complex stability will be higher. The previously studied complexes (*e.g.* RNA 1/PNAzyme I) consist of two recognition arms comprising 7 Watson–Crick base pairs on the one side and 1 GT wobble followed by 3 Watson–Crick base pairs on the other side. The elongated complexes studied herein

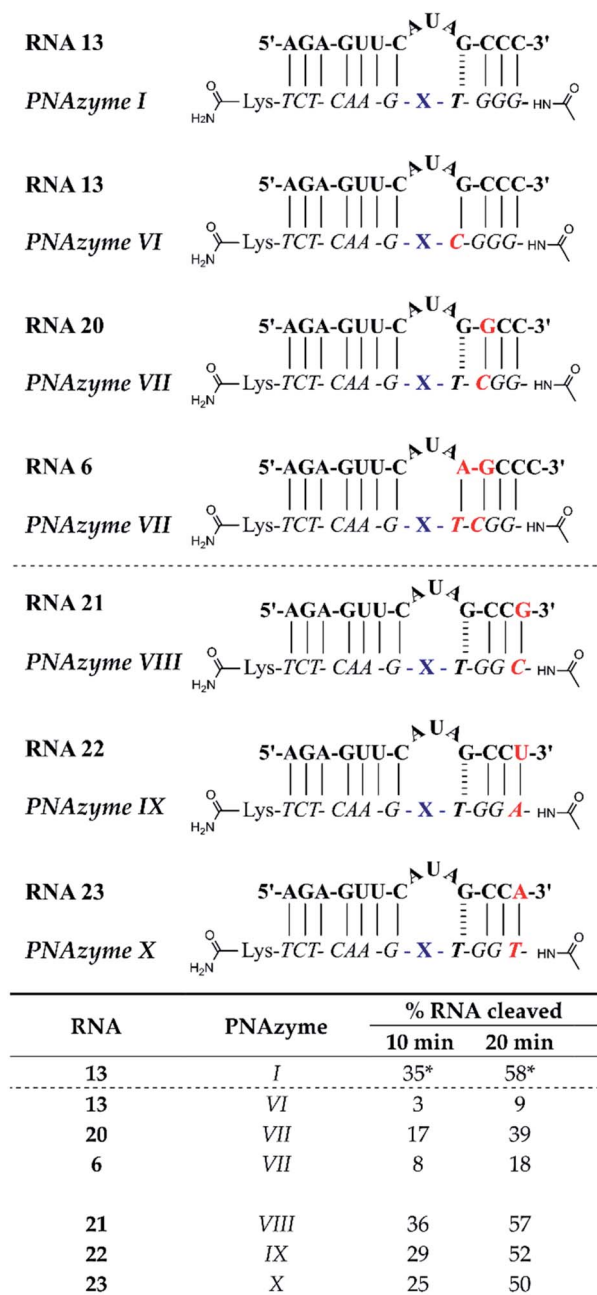


Fig. 5 Schematic representations of RNA/PNAzyme complexes where variations are introduced in the short recognition arm, followed by the extent of RNA cleavage observed in each complex after incubation of the RNA/PNAzyme complex (4  $\mu$ M) in the presence of Zn<sup>2+</sup> (100  $\mu$ M) at 37 °C, pH 7. The structure of "X" (the "molecular scissors") is shown in Fig. 1b. \*The extent of cleavage of RNA 13 by PNAzyme I has been previously reported by Luige *et al.*<sup>15</sup>

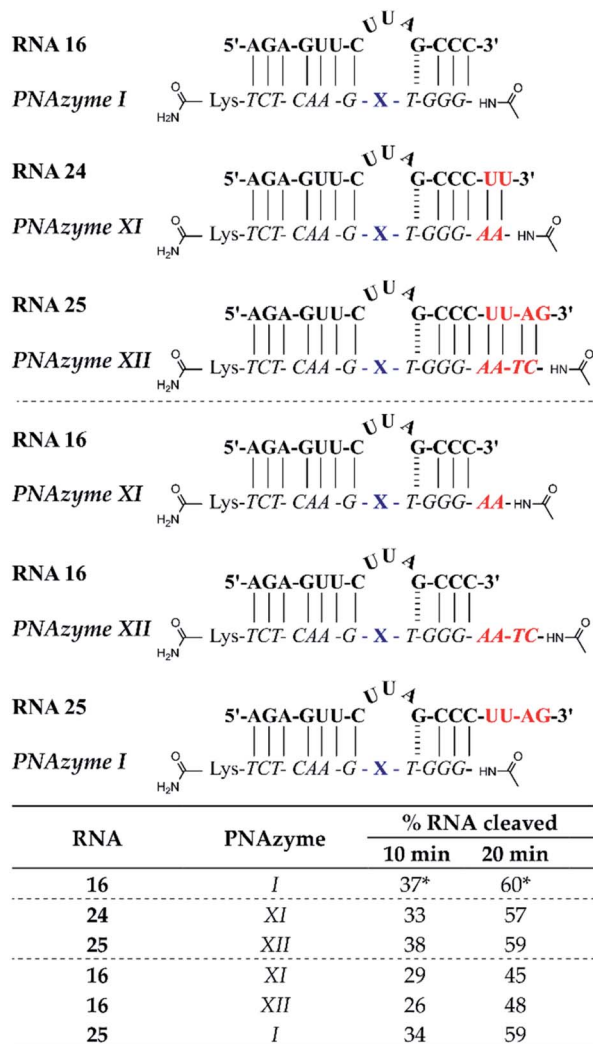


Fig. 6 Schematic representations of elongated RNA/PNAzyme complexes where the standard short recognition arm (comprising 3 Watson–Crick and 1 wobble base-pair) is extended by 2 and 4 Watson–Crick base pairs, as well as the extension of only the PNA strand which creates an overhang with the standard RNA target, followed by the extent of RNA cleavage observed in each complex after incubation of the RNA/PNAzyme complex (4  $\mu\text{M}$ ) in the presence of  $\text{Zn}^{2+}$  (100  $\mu\text{M}$ ) at 37  $^{\circ}\text{C}$ , pH 7. The structure of “X” (the “molecular scissors”) is shown in Fig. 1b. \*The extent of cleavage of RNA 16 by PNAzyme I has been previously reported by Luige *et al.*<sup>15</sup>

contained elongated short recognition arms. The RNA 24/PNAzyme XI and RNA 25/PNAzyme XII complexes have two and four additional Watson–Crick base pairs at the RNA 3' terminus. Elongation of the complex in the short recognition arm led to virtually the same extent of cleavage as the shorter system in the presence of equimolar PNAzyme (Fig. 6). Before this experiment we reasoned that the base pairs possibly need to “breathe” in order for the cleavage to take place, which could explain the critical role of the weak GT wobble closing the bulge. The fact that the extended stem systems are as efficient suggests that this is not needed. Moreover, elongation of the complex may be beneficial, as it can increase the specificity of PNAzyme binding.

In addition, we investigated the cleavage of complexes where either the PNAzyme or the RNA target sequence had an overhang (Fig. 6). The elongated PNAzyme overhang had a minor detrimental effect on the cleavage rate of RNA 16 both in the case of PNAzyme XI and XII. The significance of these results is especially relevant in the context of possible off-target sequences, where the recognition arm is incomplete, thus resulting in a PNAzyme overhang. Although it is likely that such off-targets would contain other mismatches that would potentially eliminate any catalytic activity, it is interesting that even just a short PNA overhang can decrease cleavage efficiency.

The RNA overhang, on the other hand, was shown not to affect the rate of cleavage of RNA 25 by PNAzyme I. This is an encouraging indication for cleavage of longer RNA sequences that would be the targets in potential therapeutic applications. This is also consistent with previous studies that demonstrated that  $\text{Cu}^{2+}$ -neocuproine PNAzymes readily cleave elongated RNAs that are more than double the length of the PNAzyme.<sup>16</sup> An additional point of concern could be interference from other biologically significant metals, however, Na, Mg and K were shown not to interfere with  $\text{Zn}^{2+}$ -PNAzyme-mediated RNA cleavage (see ESI-S5†).

While the rate of cleavage is unaffected by elongation of the short recognition arm, the length of the complex can be more critical for catalytic turnover, *i.e.* for the ability of each PNAzyme molecule to cleave multiple RNA targets. In order to achieve efficient turnover, the cleaved fragments must readily dissociate from the PNAzyme, so that the PNAzyme can bind to the next intact RNA target. The short RNA 16/PNAzyme I complex has been shown to give turnover.<sup>15</sup> The elongated RNA 25/PNAzyme XII system studied herein gave  $34 \pm 2\%$  RNA cleavage in 6 h under turnover conditions (33-fold excess of RNA compared to the PNAzyme, Fig. 7b). While multiple turnover was clearly demonstrated, the extent of cleavage was somewhat lower than the previously reported  $86 \pm 4\%$  for the short complex under the same conditions (Fig. 7a).<sup>15</sup> An explanation could be that elongation of the short recognition arm beyond 7 Watson–Crick base pairs decreases the rate of dissociation of the cleaved

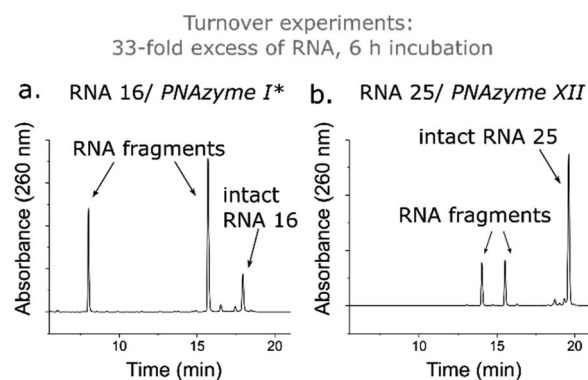


Fig. 7 IEX-HPLC chromatograms showing the extent of cleavage of a 33-fold excess of (a) RNA 16 (14mer) and (b) RNA 25 (18mer) (100  $\mu\text{M}$ ) by their corresponding PNAzymes (3  $\mu\text{M}$ ) after 6 h of incubation in the presence of  $\text{Zn}^{2+}$  (100  $\mu\text{M}$ ) at 37  $^{\circ}\text{C}$ , pH 7. \*The data shown in panel (a) has been previously reported by Luige *et al.*<sup>15</sup>

fragments, thus limiting turnover. Furthermore, these results suggest that further extension of the 7 Watson–Crick base pairs (which include 3 CG or GC pairs) in the long recognition arm would likely also adversely affect turnover.

### Bulge sequence-dependence in the cleavage of 4-nucleotide RNA bulges

Prior reports have indicated that the cleavage of the AAAA bulge can be several times slower than the AAA bulge.<sup>15</sup> While the latter is evidently a better target, the possibility of off-target cleavage of the former should be considered. Moreover, the rate of cleavage of 3-nucleotide bulges is greatly increased when the AAA bulge is replaced with UUA, AUA or GUA bulges (from *ca.* 1 h half-life to *ca.* 16 min at pH 7).<sup>15</sup> It is likely that the sequence of the 4-nucleotide bulges would also have a significant influence on the cleavage rate, but this sequence-dependence requires further investigation.

In the present report, RNAs 5–12 which form 4-nucleotide bulges of various sequences composed of adenosine and uridine in different positions were studied and compared to the cleavage of the AAAA bulge-forming RNA 4 (see Fig. 8). A significant increase in the cleavage rate was observed for RNA bulges where uridine was introduced in the first or second position (RNAs 5–6) or simultaneously in both positions (UUAA, RNA 10). The cleavage rate was also slightly increased when uridine was present in the third position (AAUA, RNA 7). In fact, the fastest cleavage rate was achieved when adenosine was replaced with uridine simultaneously in the first and third

RNA	RNA bulge sequence	% RNA cleaved	
		1 h	3 h
4	-AAAA-	18	51*
5	-UAAA-	39	78
6	-AUAA-	39	77
7	-AAUA-	24	64
8	-AAAU-	14	33
9	-AAUU-	7	20
10	-UUAA-	40	79
11	-UAUA-	55	89
12	-UUUA-		52

Fig. 8 Schematic representations of RNA/PNAzyme complexes where 4-nucleotide RNA bulges are formed, followed by the extent of RNA cleavage observed in each complex after incubation of the RNA/PNAzyme complex (4  $\mu$ M) in the presence of  $Zn^{2+}$  (100  $\mu$ M) at 37  $^{\circ}$ C, pH 7. The structure of "X" (the "molecular scissors") is shown in Fig. 1b. \*The extent of cleavage of RNA 4 by PNAzyme I in 3 h has been previously reported by Luige *et al.*<sup>15</sup>

position (UAUA bulge, RNA 11) in which case the estimated cleavage half-life was below 1 h and the cleavage occurred in a nearly site-specific fashion. The predominant cleavage site in RNA 11 was between the third and fourth bulge nucleotide (5'-

RNA	PNAzyme	% RNA cleaved	
		10 min	20 min
13	I	35*	58*
13	VIII	28	46
20	I	13	22
13	VII	6	11
20	VIII	4	8
16	I	37*	60*
26	VIII	26	47
26	XIII	8	32
27	VIII		0
28	IX		0

Fig. 9 Schematic representations of RNA/PNAzyme complexes containing various mismatches, followed by the extent of RNA cleavage observed in each complex after incubation of the RNA/PNAzyme complex (4  $\mu$ M) in the presence of  $Zn^{2+}$  (100  $\mu$ M) at 37  $^{\circ}$ C, pH 7. The structure of "X" (the "molecular scissors") is shown in Fig. 1b. \*The extent of cleavage of RNAs 13 and 16 by PNAzyme I has been previously reported by Luige *et al.*<sup>15</sup>

UAU/A-3', see ESI-S3†), in contrast to RNA 4 where cleavage fragments resulting from two equally favoured cleavage sites (AA/AA and 5'-AAA/A-3') were detected. While the presence of one or two uridines in the first three positions enhanced the cleavage rate, interestingly, the presence of three uridines in these positions (UUUA bulge in RNA 12) did not appear to affect the cleavage rate compared to the AAAA bulge in RNA 4. Similarly to the previously reported sequence dependence in 3-nucleotide bulges,<sup>15</sup> the preservation of the preferred 5'-U/A-3' or A/A cleavage site was clearly important also in 4-nucleotide bulges, as the presence of uridine in the position next to the GT wobble (AAAU bulge in RNA 8) led to a slight decrease in the cleavage rate and two consecutive uridines next to the GT wobble (AAUU bulge, RNA 9) were even more detrimental to the cleavage rate (Fig. 8).

### Mismatch intolerance in RNA/PNAzyme complexes

Lastly, we investigated the cleavage of mismatched RNA sequences (Fig. 9) to gain a deeper understanding of the specificity of PNAzymes. A terminal mismatch had only a slight negative effect on the cleavage rate, while a single mismatch in the central position in the short recognition arm or two mismatches simultaneously at the central and terminal position reduced the extent of RNA cleavage considerably. Rather interestingly, a mismatch in the long recognition arm in the position adjacent to the bulge-closing base pair had only a modest negative effect, while a GT wobble in the central part of the long recognition arm reduced the cleavage rate somewhat more. The presence of two or more mismatches in the long recognition arm completely abolished the catalytic activity. In summary, the sequence-specific nature of PNAzyme action was clearly reflected in the results. Although some cross reactivity can occur with highly similar sequences, mismatches in the RNA/PNAzyme complexes are clearly discriminative. Taken together with the requirements for the bulge sequence, the results imply high selectivity in target RNA cleavage.

## Conclusions

This study has shown that the RNA cleavage activity is critically dependent on the sequence of the RNA/PNAzyme complex. The cleavage of the AAA bulge occurred with estimated half-lives ranging from approximately 1–1.5 h to 4 h and 11 h, depending on whether the bulge was closed with a CG, GC, UA or AT base pair in the long recognition arm, respectively. Moreover, 2-nucleotide bulges of the same sequence (UA) were cleaved more efficiently when the complex contained the CG and GC bulge-closing base pairs instead of UA and AT pairs. Thus, the bulge-closing base pair was suggested to have a critical role in defining the structural arrangement of the bulge nucleotides, while interactions involving the base pair preceding the bulge-closing base pair are also likely to be important.

The bulge-closing base pair in the short recognition arm was equally influential, as the highest rates were obtained with a GT wobble, and both AT and GC Watson–Crick base pairs had

a negative effect on the cleavage rate. The activity was also affected by the succeeding base pair.

These sequence–activity relationships shed light on the structural demands for efficient cleavage of RNA targets by the Zn<sup>2+</sup>-dimethyl-dppz-PNAzymes. For rapid site-specific cleavage, the RNA targets are suggested to require the sequence 5'-Py-C-bulge-G-C-3' (where Py = pyrimidine, and the bulge sequence can be UUA, AUA or GUA), although the feasibility of the RNA sequence Pu-G (Pu = purine) instead of Py-C preceding the bulge and G-U following the bulge could also be evaluated. Variations in the rest of the sequence are likely to be tolerated. Furthermore, the RNA/PNAzyme complex can be longer than the standard 11 base pairs, albeit the catalytic turnover might be affected if either of the cleaved fragments forms more than 7 base pairs with the PNAzyme. Consequently, these structural demands must be taken into consideration in the selection of RNA target sequences for PNAzymes in future studies. Although RNA targets forming 3-nucleotide bulges are clearly more efficiently cleaved, potential off-target sequence cleavage should be considered, as the PNAzymes were shown to also cleave 4-nucleotide RNA bulges of select sequences with down to 1 h half-lives.

## Experimental

All reagents and solvents used were of analytical commercial quality. RNA cleavage experiments were performed in nuclease-free water for molecular biology purchased from Sigma-Aldrich (St. Louis, Missouri, United States). Oligoribonucleotides RNA 1–23 were purchased from Dharmacon (Lafayette, CO, USA). RNA 1 and 13 were purchased purified. RNAs 2–12 and 14–23 were purified as previously reported.<sup>15</sup> Oligoribonucleotides RNA 26–28 were purchased purified from integrated DNA Technologies, Inc (Coralville, Iowa, USA). RNA 13, 16, 24 and 25 were synthesised as previously reported.<sup>15</sup> Concentrations of RNA sequences were determined by UV absorption at 260 nm on a Varian Cary 300 UV-Vis dual beam spectrophotometer (Varian, Palo Alto, CA, USA) using extinction coefficients obtained by the nearest neighbour approximation.<sup>26</sup>

Rink Amide resin (ChemMatrix, 0.47 mmol g<sup>-1</sup>) was purchased from Biotage (Uppsala, Sweden). Peptide nucleic acid monomers, Fmoc-PNA-A(Bhoc)-OH, Fmoc-PNA-G(Bhoc)-OH, Fmoc-PNA-C(Bhoc)-OH and Fmoc-PNA-T-OH, were purchased from Link Technologies Ltd (Glasgow, UK). Fmoc-L-Dap(Mtt)-OH (2,3-diaminopropionic acid), Fmoc-L-Lys(Boc)-OH, *N,N'*-diisopropylcarbodiimide (DIC) and 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide hexafluorophosphate (HATU) were purchased from Iris Biotech GmbH (Marktredwitz, Germany). Ethyl cyano(hydroxyimino)acetate (Oxyma) was purchased from Merck-Millipore (Burlington, MA, USA). Acetic anhydride and 2,6-lutidine were purchased from Alfa Aesar (Haverhill, Massachusetts, United States). *N*-Methylpyrrolidone (NMP), dichloromethane (DCM), triisopropylsilane (TIS), piperidine and *N*-methylmorpholine (NMM) were purchased from Sigma-Aldrich (St. Louis, Missouri, USA). Trifluoroacetic acid (TFA) was purchased from Fisher Scientific (Pittsburgh, Pennsylvania, USA). Neocuproine

and 3,4-diaminobenzoic acid were purchased from Fluorochem (Hadfield, UK).

### Preparation of PNAzymes I–XII

PNA oligomers were prepared on a Biotage Initiator+ Alstra microwave peptide synthesiser as previously reported.<sup>15</sup> 3,6-Dimethyl-dipyrido[3,2-*a*:2',3'-*c*]phenazine-11-carboxylic acid was synthesised as previously reported.<sup>15</sup> PNAzyme I has been previously reported.<sup>15</sup> PNA conjugates (PNAzymes II–XII) were prepared as previously reported.<sup>15</sup> The RP-HPLC purity analysis of purified PNA conjugates is available in ESI-S1.† The purified PNA conjugates were identified by mass spectrometry on a Bruker Ultraflex MALDI-TOF mass spectrometer in positive ion mode using a sinapic acid matrix (10 mg mL<sup>-1</sup> sinapic acid in acetonitrile–water (1 : 1, v/v) containing 0.1% TFA).

MALDI-MS *m/z*: **PNAzyme I** C<sub>151</sub>H<sub>181</sub>N<sub>73</sub>O<sub>38</sub> [M + H]<sup>+</sup> calculated 3625.34, observed 3625.36. **PNAzyme II** C<sub>162</sub>H<sub>194</sub>N<sub>80</sub>O<sub>41</sub> [M + H]<sup>+</sup> calculated 3916.56, observed 3916.92. **PNAzyme III** C<sub>161</sub>H<sub>194</sub>N<sub>78</sub>O<sub>41</sub> [M + H]<sup>+</sup> calculated 3876.56, observed 3877.13. **PNAzyme IV** C<sub>162</sub>H<sub>195</sub>N<sub>77</sub>O<sub>42</sub> [M + H]<sup>+</sup> calculated 3891.56, observed 3891.73. **PNAzyme V** C<sub>162</sub>H<sub>194</sub>N<sub>80</sub>O<sub>40</sub> [M + H]<sup>+</sup> calculated 3900.57, observed 3901.06. **PNAzyme VI** C<sub>150</sub>H<sub>180</sub>N<sub>74</sub>O<sub>37</sub> [M + H]<sup>+</sup> calculated 3610.46, observed 3611.66. **PNAzyme VII** C<sub>150</sub>H<sub>181</sub>N<sub>71</sub>O<sub>38</sub> [M + H]<sup>+</sup> calculated 3585.45, observed 3585.60. **PNAzyme VIII** C<sub>150</sub>H<sub>181</sub>N<sub>71</sub>O<sub>38</sub> [M + H]<sup>+</sup> calculated 3585.45, observed 3585.29. **PNAzyme IX** C<sub>151</sub>H<sub>181</sub>N<sub>73</sub>O<sub>37</sub> [M + H]<sup>+</sup> calculated 3609.46, observed 3609.30. **PNAzyme X** C<sub>151</sub>H<sub>182</sub>N<sub>70</sub>O<sub>39</sub> [M + H]<sup>+</sup> calculated 3600.45, observed 3600.53. **PNAzyme XI** C<sub>173</sub>H<sub>207</sub>N<sub>87</sub>O<sub>42</sub> [M + H]<sup>+</sup> calculated 4175.68, observed 4176.44. **PNAzyme XII** C<sub>194</sub>H<sub>234</sub>N<sub>96</sub>O<sub>49</sub> [M + H]<sup>+</sup> calculated 4692.88, observed 4693.48.

Concentrations of PNA conjugates were determined by UV absorption at 260 nm on a Varian Cary 300 UV-Vis dual beam spectrophotometer (Varian, Palo Alto, CA, USA), using extinction coefficients obtained by the nearest neighbour approximation,<sup>26</sup> and the reported extinction coefficient for dipyrrophenazine.<sup>27</sup>

### RNA cleavage experiments

RNA cleavage experiments were performed and analysed as previously reported,<sup>15</sup> but for clarity we requote them in ESI-S2.† Representative IEX-HPLC chromatograms showing the extent of RNA cleavage at specified timepoints are shown in ESI-S3.† RNA cleavage half-lives were estimated from the percentage of RNA cleavage at the given time points using the rate law for first order reactions since our experiments are performed under saturation conditions with respect to PNAzyme and Zn<sup>2+</sup> ion binding.

### Determination of RNA cleavage sites

RNA cleavage experiments were performed as detailed above, but with a 22 hour incubation time. The cleaved fragments were then identified by LC-MS analysis of the quenched reaction aliquots as follows. HRMS spectra were collected by elution of the aliquots of RNA cleavage reactions on an Oligonucleotide BEH C18 (2.1 × 50 mm, 130 Å, 3 μm) column using a 0.4

mL min<sup>-1</sup> linear gradient from 0 to 80% acetonitrile in 8.6 mM TEA + 100 mM HFIP buffer over 15 min at 40 °C on Waters Xevo G2-XS QToF. Ionization mode: ESI negative. Source capillary voltage: 2 kV. Source desolvation temperature: 320 °C. Full scan with fragmentation, mass range: 400–2000 Da/0.5 s, sensitivity mode. Deconvoluted using 5 ppm tolerance. The MS spectra of the cleaved fragments are shown in ESI-S3.

### Circular dichroism (CD) spectroscopy

CD spectra of RNA/PNAzyme complexes were measured between 230 and 370 nm on a Jasco J-1500 CD Spectrometer using 10 mm path length cuvettes. The spectra were recorded as an average of five scans at 25 °C and normalised by subtracting the background buffer scans. The samples of RNA/PNAzyme complexes (4 μM) were analysed in the absence of Zn<sup>2+</sup> ions in 10 mM phosphate buffer containing 100 mM NaCl and 0.1 mM EDTA at pH 7.0. The spectra were smoothed over 5 points. The spectra are shown in ESI-S4.†

### Author contributions

Conceptualisation – O. L., M. M. and R. S. Formal analysis – O. L. and R. S. Funding acquisition – R. S. Investigation – O. L. and K. K. with contributions from P. B., M. B. Methodology – O. L., M. M., R. S. Project administration – O. L., R. S. Resources – U. T., R. S. Supervision – M. B., U. T., R. S. Visualisation – O. L. Writing - original draft – O. L. Writing - review and editing – O. L. and R. S. with contributions from all.

### Disclosures

Since performing this work, Olivia Luige has become an employee of AstraZeneca.

### Conflicts of interest

There are no conflicts to declare.

### Acknowledgements

The authors would like to extend their gratitude to Rouven Stulz and the Separation Science Laboratory at AstraZeneca Gothenburg for their help with RNA synthesis and purification. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 721613. This material reflects only the authors' views and the Union is not liable for any use that may be made of the information contained therein.

### References

- 1 R. J. Roberts, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 5905–5908.
- 2 X. Shen and D. R. Corey, *Nucleic Acids Res.*, 2018, **46**, 1584–1600.



- 3 E. C. Kuijper, A. J. Bergsma, W. W. M. P. Pijnappel and A. Aartsma-Rus, *J. Inherited Metab. Dis.*, 2021, **44**, 72–87.
- 4 J. Hall, D. Hüsken and R. Häner, *Nucleic Acids Res.*, 1996, **24**, 3522–3526.
- 5 T. Niittymäki and H. Lönnberg, *Org. Biomol. Chem.*, 2006, **4**, 15–25.
- 6 A. Kuzuya and M. Komiyama, *Curr. Org. Chem.*, 2007, **11**, 1450–1459.
- 7 A. Ghidini, M. Murtola and R. Strömberg, in *DNA in Supramolecular Chemistry and Nanotechnology*, John Wiley and Sons, Ltd, Hoboken, NJ, USA, 1st edn, 2015, pp. 158–171.
- 8 Y. Staroseletz, S. Gaponova, O. Patutina, E. Bichenkova, B. Amirloo, T. Heyman, D. Chiglintseva and M. Zenkova, *Molecules*, 2021, **26**, 1732.
- 9 Y. Aoki and M. J. A. Wood, *Neuromuscular Disord.*, 2021, 1–16.
- 10 F. Zellmann, L. Thomas, U. Scheffer, R. Hartmann and M. Göbel, *Molecules*, 2019, **24**, 807.
- 11 F. Zellmann and M. W. Göbel, *Molecules*, 2020, **25**, 1842.
- 12 O. Patutina, D. Chiglintseva, E. Bichenkova, S. Gaponova, N. Mironova, V. Vlassov and M. Zenkova, *Molecules*, 2020, **25**, 2459.
- 13 Y. Staroseletz, B. Amirloo, A. Williams, A. Lomzov, K. K. Burusco, D. J. Clarke, T. Brown, M. A. Zenkova and E. V. Bichenkova, *Nucleic Acids Res.*, 2020, **48**, 10662–10679.
- 14 H. Åström, N. H. Williams and R. Strömberg, *Org. Biomol. Chem.*, 2003, **1**, 1461–1465.
- 15 O. Luige, P. P. Bose, R. Stulz, P. Steunenbergh, O. Brun, S. Andersson, M. Murtola and R. Strömberg, *Chem. Commun.*, 2021, **57**, 10911–10914.
- 16 O. Luige, M. Murtola, A. Ghidini and R. Strömberg, *Molecules*, 2019, **24**, 672.
- 17 M. Murtola, A. Ghidini, P. Virta and R. Strömberg, *Molecules*, 2017, **22**, 1856.
- 18 A. Ghidini, M. Murtola and R. Strömberg, *Org. Biomol. Chem.*, 2016, **14**, 2768–2773.
- 19 M. Murtola, M. Wenska and R. Strömberg, *J. Am. Chem. Soc.*, 2010, **132**, 8984–8990.
- 20 M. Murtola and R. Strömberg, *Arkivoc*, 2009, 84–94.
- 21 M. Murtola and R. Strömberg, *Org. Biomol. Chem.*, 2008, **6**, 3837–3842.
- 22 H. Åström and R. Strömberg, *Org. Biomol. Chem.*, 2004, **2**, 1901–1907.
- 23 D. Hüsken, G. Goodall, M. J. J. Blommers, W. Jahnke, J. Hall, R. Häner and H. E. Moser, *Biochemistry*, 1996, **35**, 16591–16600.
- 24 U. Kaukinen, L. Bielecki, S. Mikkola, R. W. Adamiak and H. Lönnberg, *J. Chem. Soc., Perkin Trans. 2*, 2001, 1024–1031.
- 25 T. Hermann and D. J. Patel, *Structure*, 2000, **8**, 47–54.
- 26 J. D. Puglisi and I. Tinoco, *Methods Enzymol.*, 1989, **180**, 304–325.
- 27 E. B. Van der Tol, H. J. Van Ramesdonk, J. W. Verhoeven, F. J. Steemers, E. G. Kerver, W. Verboom and D. N. Reinhoudt, *Chem.–Eur. J.*, 1998, **4**, 2315–2323.