

Synthesis of a Peptidoyl RNA Hairpin via a Combination of Solid-Phase and Template-Directed Chain Assembly

Jennifer Bremer,^[a] Christian Richter,^[b] Harald Schwalbe,^[b] and Clemens Richert^{*[a]}

Dedicated to Ulf Diederichsen

Peptidoyl RNAs are the products of ribosome-free, single-nucleotide translation. They contain a peptide in the backbone of the oligoribonucleotide and are interesting from a synthetic and a bioorganic point of view. A synthesis of a stabilized version of peptidoyl RNA, with an amide bond between the C-terminus of a peptide and a 3'-amino-2',3'-dideoxynucleoside in the RNA chain was developed. The preferred synthetic route used an *N*-Teoc-protected aminonucleoside support and in-

volved a solution-phase coupling of the amino-terminal oligonucleotide to a dipeptido dinucleotide. Exploratory UV-melting and NMR analysis of the hairpin 5'-UUGGCGAAAGCdC-LeuLeu-AA-3' indicated that the peptide-linked RNA segments do not fold in a cooperative fashion. The synthetic access to doubly RNA-linked peptides on a scale sufficient for structural biology opens the door to the exploration of their structural and biochemical properties.

Introduction

Covalently linked hybrids of peptides and nucleic acids are fascinating target molecules that are rich in functional groups and stereogenic centers. Their synthesis combines challenges of peptide and oligonucleotide chemistry.^[1,2] Each of the biopolymers has its own set of side reactions during chain assembly. Further, their intrinsic reactivity requires conditions for deprotection and purification that are orthogonal in many respects.^[3,4] The synthetic challenge is particularly acute for hybrids containing RNA as nucleic acid, since RNA is more labile than DNA and gives lower yields in solid-phase synthesis.^[5]

Several groups have reported successfully overcoming the synthetic challenges of preparing hybrid molecules that contain RNA and a peptide. They all used covalent linkages more stable than the ester bond found in the peptidyl tRNA intermediates of translation.^[6–8] Other synthetic constructs were prepared using squarate-diester-mediated coupling,^[9] boronate-forming reactions,^[10] or native ligation.^[11] Our own synthetic work has focused on phosphoramidate-linked peptide-RNA hybrids called 'peptido RNA', for which a solution-phase coupling approach has recently been reported.^[12]

Peptido RNAs contain the phosphoramidate bond between the 5'-terminal phosphate of the oligoribonucleotide chain and the *N*-terminus of the peptide.^[13,14] This leaves the C-terminus of the peptide free for coupling to the 3'-terminus of another nucleic acid strand, such as the bond found in the peptidyl RNAs of translation. Template-directed coupling reactions were indeed observed by us when dipeptido dinucleotides were allowed to react with 3'-aminoterminal strands on an RNA template.^[15] The resulting doubly linked species were named 'peptidoyl RNAs' to indicate that they contain the covalent links found in peptido and in peptidyl RNAs.

The rate of formation of peptidoyl RNAs was found to depend strongly on the relative position of the two strands on the template and on the amino acid residues of the dipeptide.^[15] Further, the reaction was fastest when no gap was left between the double helices formed by the peptide-bearing strand and the strand coupling at its 3'-terminus. Even more surprisingly, the sterically demanding leucinylleucine dipeptide sequence (LeuLeu) gave the fastest coupling, not GlyGly with its unencumbered chain.^[15] A similar finding was made when dipeptidoyl RNAs were allowed to form in ribosome-free single-nucleotide translation assays, in which a 3'-aminoacylated transfer-nucleotide couples to a 3'-aminoacyl primer on an RNA strand acting as primitive mRNA.^[16] Again, sterically demanding amino acid residues like Val and Leu were among the species reacting fast and efficiently. These unexpected findings were particularly intriguing because of the relevance for the origin of the genetic code^[17–20] and biological homochirality.^[21,22]

We therefore initiated a project aimed at gaining insights into the three-dimensional structure of peptidoyl RNAs. Building on our earlier work involving NMR-monitored binding studies with primer-template systems,^[23–25] we designed hairpin 1 (Figure 1) as our target molecule. This peptidoyl hairpin features a LeuLeu dipeptide in the backbone of the nucleic acid strand and a GAAA tetraloop^[26,27] to connect the termini of what would otherwise be a primer and a template. It is small enough to be

[a] J. Bremer, Prof. C. Richert
Institut für Organische Chemie, University of Stuttgart
70569 Stuttgart (Germany)
E-mail: lehrstuhl-2@oc.uni-stuttgart.de

[b] Dr. C. Richter, Prof. Dr. H. Schwalbe
Institute for Organic Chemistry and Chemical Biology,
Johann Wolfgang Goethe-University
60438 Frankfurt (Germany)

Supporting information for this article is available on the WWW under <https://doi.org/10.1002/cbic.202200352>

This article belongs to a Joint Special Collection dedicated to Ulf Diederichsen.

© 2022 The Authors. ChemBioChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

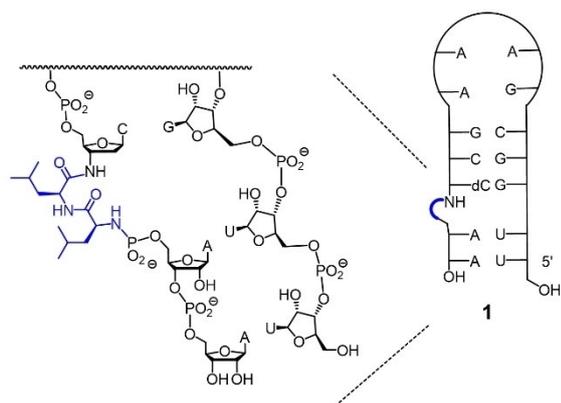


Figure 1. Structure of target molecule 1.

studied by NMR without having to resort to isotopically enriched building blocks. Here we report the synthesis of 1, together with the results of an exploratory study on its folding properties.

Results and Discussion

It was decided that the assembly of the RNA portion of 1 was to occur via solid-phase chain assembly, using the TBDMS method.^[28] The dipeptide H-Leu-Leu-OH is available from a commercial source, obviating the need to prepare it by Boc-based solution-phase synthesis. The main synthetic issue was thus to set up the amide and the phosphoramidate linkages. Figure 2a shows key structures of three synthetic strategies that were explored experimentally.

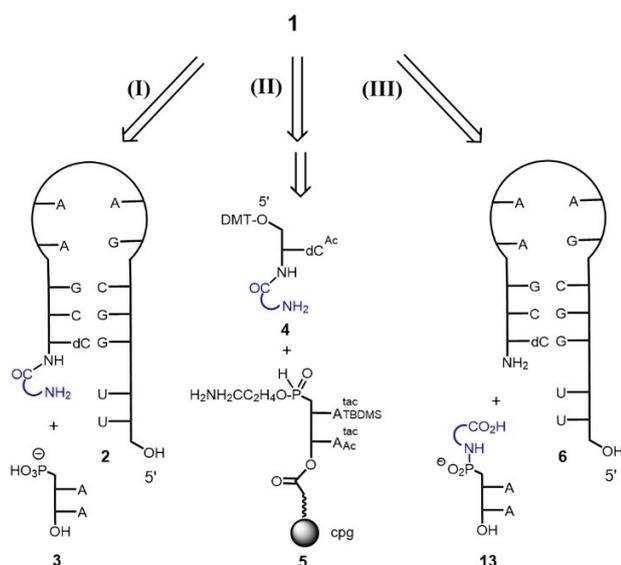


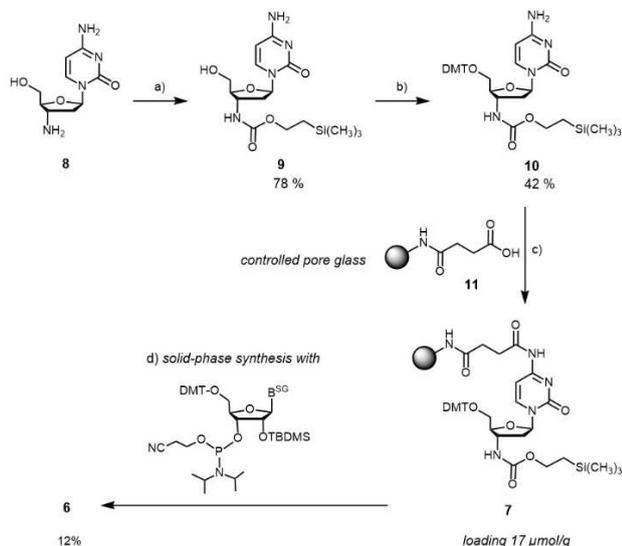
Figure 2. Retrosynthetic considerations and key intermediates of synthetic paths for the preparation of peptidoyl RNA hairpin 1.

Three different approaches for the synthesis of peptidoyl RNA were tested. The first approach involved the assembly of the 5'-terminal segment, including the dipeptide at its 3'-terminus (2), and coupling this species to dinucleotide 3 (path I, Figure 2). To this end, an *N*-trimethylsilyloxyethylcarboxyl (Teoc) protected leucylleucine was coupled to the 3'-position of 3'-amino-2',3'-dideoxycytidine, and the resulting nucleoside was immobilized via a linker to its *N4* position on controlled-pore glass (cpg), followed by RNA chain assembly. The Teoc group was known to be stable during the steps of phosphoramidite-based chain assembly and to be removable with TBAF.^[29] While the overall yield of the peptidyl nucleoside was 17%, and the loading of the support was 10 $\mu\text{mol/g}$, the final coupling of 2 to a preactivated dinucleotide gave no more than traces of 1, as determined by MALDI-TOF analysis of the reaction mixture. Attempts to improve the yield were unsuccessful, and path I was not pursued further.

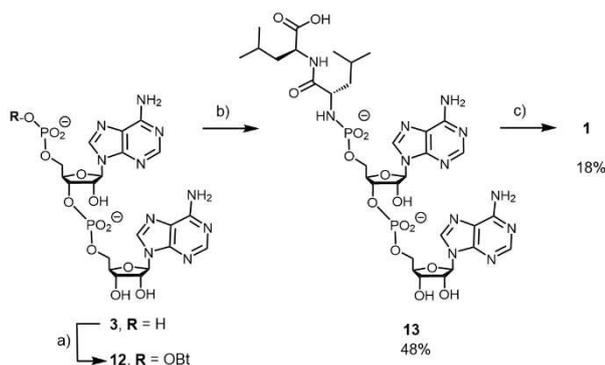
The second approach used path II of Figure 2. Here, the *N*-Teoc protected 3'-dipeptidyl 3'-amino-2',3'-dideoxycytidine residue was *N4* acetylated and 5' DMT-protected, followed by removal of the Teoc group with TBAF, to give nucleosidic building block 4. Separately, an AA dimer was synthesized by automated RNA synthesis and then 5'-phosphitylated, followed by partial hydrolysis with dicyanoimidazole/water to obtain solid-support bound *H*-phosphonate 5. The phosphoramidate linkage was established by redox condensation using CBr_4 and NEt_3 , as previously described for DNA.^[30] Unfortunately, no more than a loading of 3.4 μmol per gram of cpg was achieved, so that subsequent elaboration of the 5'-terminal RNA segment by solid-phase synthesis gave too little product for NMR spectroscopy.

The preferred synthesis followed path III of Figure 2. It uses a template-directed coupling of a dipeptidoyl dinucleotide to 3'-aminoterminally protected hairpin 6 as the final step. The elaboration of 6 via solid support 7 is shown in Scheme 1. As mentioned for the earlier two paths, above, 3'-amino-2',3'-dideoxycytidine (8) was the starting point. The aminonucleoside was Teoc-protected to give 9, which was then tritylated with DMT-Cl in modest yield to afford doubly protected 10. The remaining *N4* amino group of the nucleobase was then used as anchoring point for immobilization on controlled-pore glass, using the methodology of Brown and coworkers.^[31] To accomplish the immobilization of the nucleoside, long-chain alkylamine controlled pore glass (lcaa-cpg) was first treated with succinic anhydride in the presence of DMAP to give 11 and then with 10 in the presence of DIC/HOBt. The nucleoside loading of 7 was found to be 17 $\mu\text{mol/g}$ by trityl assay, and chain assembly by automatic synthesis proceeded uneventfully to give hairpin 6 after basic deprotection followed by Teoc removal with fluoride.

The synthesis of the dipeptido dinucleotide portion was performed using a modification of the method of R uchle et al.,^[12] as shown in Scheme 2. Ribonucleotide dimer 3 was converted to HOBt ester 12, which was then coupled with leucylleucine to yield 13 in 48% overall yield. Coupling between the hairpin and the activated dipeptido dinucleotide to form peptidoyl RNA 1 was then performed in template-



Scheme 1. Synthesis of 3'-aminoterminial hairpin **6**. Conditions: a) 1. Teoc-succinimide, DIPEA, DMF/H₂O; b) DMT-Cl, DMAP, pyridine; c) DIC, HOBt, pentachlorophenol, pyridine; d) 1. RNA chain assembly with commercial phosphoramidite building blocks, B^{PG} = A^{Bz}, C^{Ac} or G^{tBu}, 2. AMA, 55 °C, 3. TBAF, THF.



Scheme 2. Elaboration of dipeptido dinucleotide **13**, and assembly of hairpin **1** by coupling with amino-terminal **2**. a) HOBt, EDC, pH 6, followed by precipitation; b) leucylleucine, H₂O, pH 8.3; c) EDC in 200 mM NaCl, 50 mM MOPS, pH 7.4, 0 °C.

directed fashion in aqueous condensation buffer at 0 °C.^[13] That the coupling of **13** to a 3'-aminoterminial nucleotide is a template-directed reaction had previously been demonstrated experimentally (see Figure 3 in Ref. [15]). While monitoring in the present case indicated well over 50% conversion to product, purification proved challenging. The current protocol for HPLC, using a gradient of acetonitrile in ammonium bicarbonate buffer on a C18 phase gave a yield of 18% of spectroscopically pure **1**. Figure 3 shows a MALDI-TOF mass spectrum and an excerpt of the one-dimensional ¹H-NMR spectrum of our target compound.

Since the NMR spectrum showed well resolved signals and a dominant conformation, a series of two-dimensional NMR spectra was measured for **1**, **6** and unmodified RNA hairpin sequence 5'-UUGGCGAAAGCCAA-3' (**14**) as control, allowing for

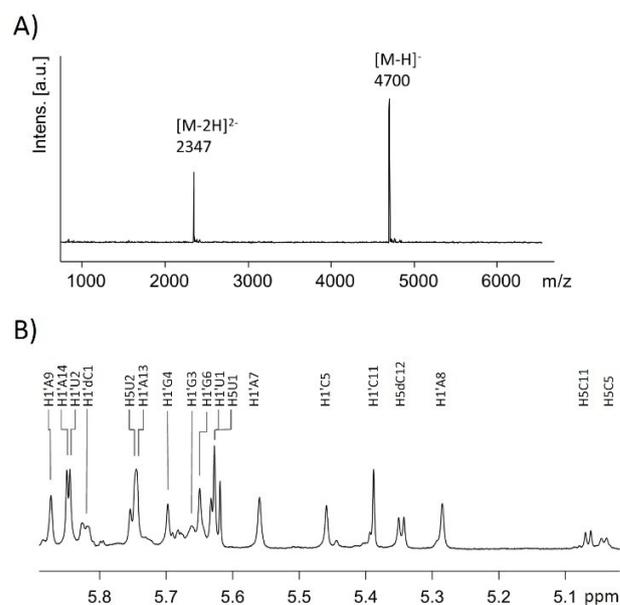


Figure 3. Characterization of **1**. A) MALDI-ToF mass spectrum, and B) region of the ¹H-NMR spectrum in D₂O, 50 mM phosphate buffer, 100 mM NaCl, pH 7.0, 900 MHz, 25 °C showing resonances from the riboses and H5 protons of pyrimidines.

assignment of key resonances for all residues of **1**. Figure 4 shows an excerpt of the NOESY spectrum with assignments indicated on the margins. Additional spectra can be found in the Supporting Information. Well-resolved and sufficiently intense cross peaks were observed for the core RNA portion of the hairpin (G3-dC12) to be identified as folded according to our design, with the expected base pairing for this part of hairpin **1** at room temperature (25 °C). Two intense and one weak imino resonance were observed in the low-field region of spectra measured in H₂O/D₂O (9:1), as expected for a stem with

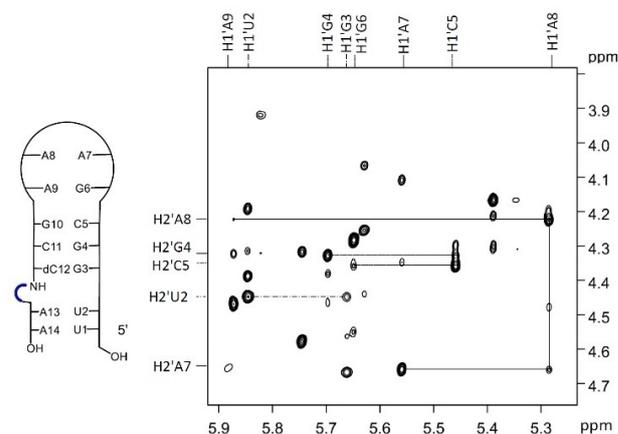


Figure 4. Section of the NOESY-spectrum of **1** in D₂O, 50 mM phosphate 100 mM, NaCl, pH 7.0, 900 MHz, 25 °C, and a mixing time of 300 ms showing cross peaks between anomeric protons and H2' protons used for assignment.

two shielded and one exposed G:C base pair (see Figure S18, Supporting Information).

For the peptide segment, the situation was less clear. For the 3'-terminal region, with its weakly pairing UU/AA dinucleotides and the LeuLeu portion, a single, stably folded conformation would probably not lead to the combination of cross peaks observed (Figure 5). Both internucleotide NOESY cross peaks between dC12 and A13, and between the amino acid residues of LeuLeu and A13 were observed. The former are expected for an uninterrupted, continuously stacked double helix, whereas the latter point to a partially unwound helix. Further, attempts to generate a three-dimensional structure by restrained torsion angle molecular dynamics, using the available NOE constraints and protocols previously applied to similar short helices,^[32–34]

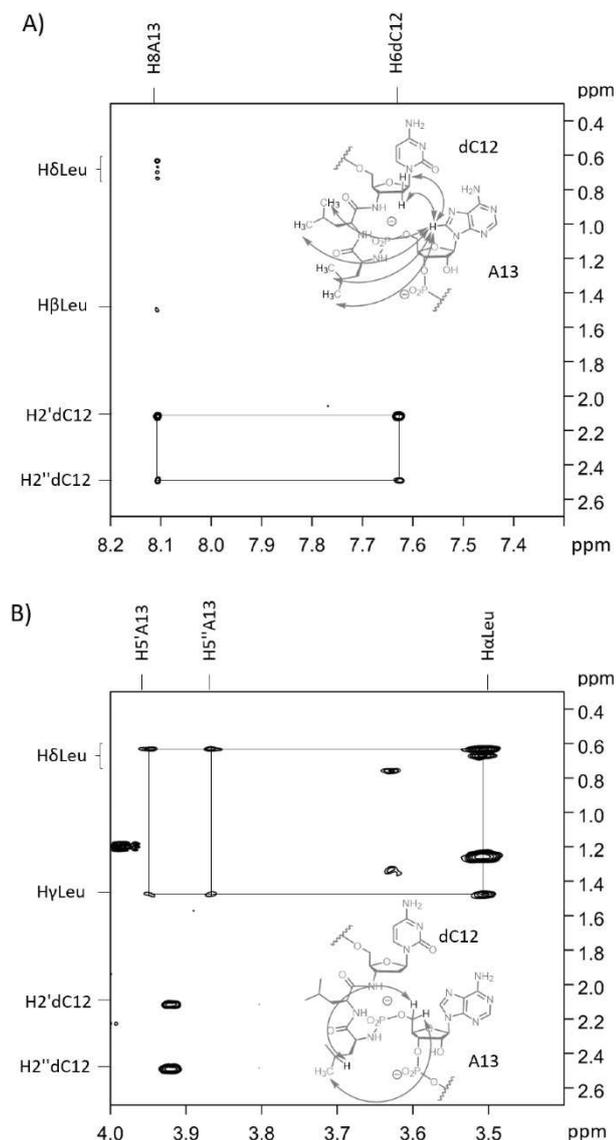


Figure 5. Excerpts of ^1H -NOESY NMR spectrum of **1** showing cross peaks from amino acid residues of LeuLeu or the sugar of dC12 to a nucleotide of the 3'-terminal dinucleotide, with NOEs indicated as grey double-headed arrows in the structural drawings. A) NOEs to H8 of A13, and B) NOEs to H5'/5'' of A13.

did not give any violation-free structures despite numerous computational runs. Therefore, we suspected that **1** was disordered or incompletely folded in its terminal region at 25 °C.

This hypothesis was corroborated by the results of a UV-melting study. When hairpin **1** was subjected to thermal denaturation, a biphasic curve was obtained (Figure 6A). The lower transition had its point of inflection at 21 °C, as seen in the maximum of the first derivative, and the higher temperature transition gave a melting point of 70 °C. By comparison, unmodified RNA hairpin 5'-UUGGCGAAAGCCAA-3' (**14**) showed a single transition, with a melting point of 83 °C, as expected for a cooperative folding/disassembly process involving the entire stem, including the terminal UU/AA segment. (Figure 6B). Melting curves at other salt concentrations, as well as melting curves with pruned versions of **1**, including **6**, confirmed these findings and are presented in chapter 3 of the Supporting Information. Biphasic transitions were observed for **1**, independent of the NaCl content, and a single transition was found for **14**. Together they indicate that the core portion of **1**, including its GAAA tetraloop, folds into a very stable motif, but that the LeuLeu dipeptide segments prevents the AA:UU duplex from folding cooperatively with the remainder of the hairpin when the dipeptide is inserted in the RNA backbone. Instead, the terminal segment now melts independently, with the transition occurring at or below room temperature. Line broadening observed in the ^1H NMR spectra for the H8 signals of A13 and A14 upon cooling from 25 °C to 3 °C (Figure S19, Supporting Information) again confirm this hypothesis.

These findings are interesting because they indicate that reactions anchoring the C-terminus of peptidoyl RNA readily produce products in template-directed fashion that are much less stable thermodynamically than an uninterrupted double helix.^[15] Likewise, the dipeptidoyl RNA products of single-nucleotide translation with sterically demanding aliphatic amino acid residues^[16] may not be formed due to product-development control. Rather, their formation with faster rates than those of less sterically demanding residues^[16] may be the consequence of kinetic control.

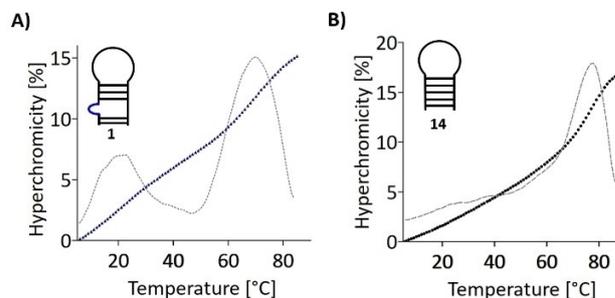


Figure 6. UV-Melting curves of hairpins with first derivatives shown as grey broken lines. A) Peptidoyl hairpin **1**, and B) control hairpin 5'-UUGGCGAAAGCCAA-3' (**14**). Conditions: 5 μM hairpin, in 50 mM phosphate buffer, pH 7, measured at 260 nm and a heating rate of 1 °C/min; first derivatives generated in QtiPlot.

Conclusions

The results presented here show how RNA-peptide hybrid molecules with the peptide embedded in the backbone of the RNA can be prepared on a scale suitable for structural biology studies by a combination of solid-phase chain assembly and a template-directed peptide coupling in aqueous solution. The method should allow for the preparation of a range of different sequences and chain lengths, allowing for an exploration of the chemical and structural properties of peptidoyl RNAs, and their possible role in the prebiotic phase of evolution. This may include structural studies employing X-ray crystallography. One of the questions that will be interesting to address is whether other peptide sequences have a smaller duplex-destabilizing effect than the LeuLeu dipeptide studied here. Studies on DNA with amino acids incorporated in the backbone suggest that this may be the case.^[35] Independent of prebiotically motivated studies, though, our results demonstrate the power of template-directed reactions in synthetic chemistry, complementing better-known usages of such reactions in enzyme-free genetic copying.^[36]

Experimental Section

Protocols, synthetic schemes, and additional spectra can be found in the Supporting Information.

Acknowledgements

The authors thank Franziska Welsch for experimental contributions during an early phase of the project. This work was supported by Volkswagen Foundation (grant Az 92 768) and Deutsche Forschungsgemeinschaft (DFG) project ID 364653263-TRR 235. Work at BMRZ is supported by the State of Hesse. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: NMR spectroscopy · oligonucleotides · peptides · RNA · translation

- [1] C. Richert, P. Grünefeld, *Synlett* **2007**, 0001–0018.
- [2] L. K. McKenzie, R. El-Khoury, J. D. Thorpe, M. J. Damha, M. Hollenstein, *Chem. Soc. Rev.* **2021**, *50*, 5126–5164.
- [3] H. Sun, A. Brik, *Acc. Chem. Res.* **2019**, *52*, 3361–3371.
- [4] L. Baronti, H. Karlsson, M. Marušič, K. Petzold, *Anal. Bioanal. Chem.* **2018**, *410*, 3239–3252.
- [5] R. I. Hogrefe, B. Midthune, A. Lebedev, *Isr. J. Chem.* **2013**, *53*, 326–349.
- [6] S. Terenzi, E. Biala, N. Q. Nguyen-Trung, P. Strazewski, *Angew. Chem.* **2003**, *115*, 3015–3018; *Angew. Chem. Int. Ed.* **2003**, *42*, 2909–2912.
- [7] D. Graber, H. Moroder, J. Steger, K. Trappl, N. Polacek, R. Micura, *Nucleic Acids Res.* **2010**, *38*, 6796–6802.
- [8] A.-S. Geiermann, N. Polacek, R. Micura, *J. Am. Chem. Soc.* **2011**, *133*, 19068–19071.
- [9] M. Fonvielle, N. Sakkas, A. El-Sagheer, E. Braud, S. Cardon, T. Brown, M. Arthur, M. Etheve-Quellejeu, *Angew. Chem.* **2016**, *128*, 13751–13755; *Angew. Chem. Int. Ed.* **2016**, *55*, 13553–13557.
- [10] A. Gimenez Molina, I. Barvik, S. Müller, J.-J. Vasseur, M. Smietana, *Org. Biomol. Chem.* **2018**, *16*, 8824–8830.
- [11] M. McPherson, M. Wright, C. Lohse, *Synlett* **1999**, *S1*, 978–980.
- [12] M. Räu Chile, G. Leveau, C. Richert, *Eur. J. Org. Chem.* **2020**, *45*, 6966–6975.
- [13] M. Jauker, H. Griesser, C. Richert, *Angew. Chem. Int. Ed.* **2015**, *54*, 14564–14569; *Angew. Chem.* **2015**, *127*, 14772–14777.
- [14] H. Griesser, P. Tremmel, E. Kervio, C. Pfeffer, U. E. Steiner, C. Richert, *Angew. Chem. Int. Ed.* **2017**, *56*, 1219–1223; *Angew. Chem.* **2017**, *129*, 1239–1243.
- [15] B. Jash, C. Richert, *Chem. Sci.* **2020**, *11*, 3487–3494.
- [16] B. Jash, P. Tremmel, D. Jovanovic, C. Richert, *Nat. Chem.* **2021**, *13*, 751–757.
- [17] M. Yarus, J. G. Caporaso, R. Knight, *Annu. Rev. Biochem.* **2005**, *74*, 179–198.
- [18] Y. I. Wolf, E. V. Koonin, *Biology Direct* **2007**, *2*, 14.
- [19] M. Yarus, J. J. Widmann, R. Knight, *J. Mol. Evol.* **2009**, *69*, 406–429.
- [20] M. Yarus, *Life* **2017**, *7*, 13.
- [21] D. Beaufile, G. Danger, L. Boiteau, J. C. Rossi, R. Pascal, *Chem. Commun.* **2014**, *50*, 3100–3102.
- [22] O. Doppleb, J. Bremer, M. Bechthold, C. Sanchez, C. Richert, *Chem. Eur. J.* **2021**, *27*, 13544–13551.
- [23] E. Kervio, B. Claasen, U. E. Steiner, C. Richert, *Nucleic Acids Res.* **2014**, *42*, 7409–7420.
- [24] E. Kervio, M. Sosson, C. Richert, *Nucleic Acids Res.* **2016**, *44*, 5504–5514.
- [25] S. Motsch, D. Pfeffer, C. Richert, *ChemBioChem* **2020**, *21*, 2013–2018.
- [26] F. M. Jucker, H. A. Heus, P. F. Yip, E. H. Moors, A. Pardi, *J. Mol. Biol.* **1996**, *264*, 968–980.
- [27] J. M. Blose, K. P. Lloyd, P. C. Bevilacqua, *Biochemistry* **2009**, *37*, 8787–8794.
- [28] S. Agrawal, *Protocols for Oligonucleotides and Analogs*, Humana, Totowa, **1993**, p. 20.
- [29] T. Kottysch, C. Ahlborn, F. Brotzel, C. Richert, *Chem. Eur. J.* **2004**, *16*, 4017–4028.
- [30] J. A. Rojas Stütz, C. Richert, *J. Am. Chem. Soc.* **2001**, *50*, 12718–12719.
- [31] A. H. El-Sagheer, T. Brown, *Proc. Natl. Acad. Sci. USA* **2010**, *35*, 15329–15334.
- [32] K. Siegmund, S. Maheshwary, S. Narayanan, W. Connors, M. Riedrich, M. Printz, C. Richert, *Nucleic Acids Res.* **2005**, *33*, 4838–4848.
- [33] A. Patra, C. Richert, *J. Am. Chem. Soc.* **2009**, *131*, 12671–12681.
- [34] C. Gerlach, B. Claasen, C. Richert, *ChemBioChem* **2014**, *15*, 2584–2589.
- [35] M. Meng, C. Ducho, *Beilstein J. Org. Chem.* **2018**, *14*, 1293–1308.
- [36] C. Deck, M. Jauker, C. Richert, *Nat. Chem.* **2011**, *3*, 603–608.

Manuscript received: June 19, 2022
Revised manuscript received: July 21, 2022
Accepted manuscript online: July 22, 2022
Version of record online: August 5, 2022