

Methods in  
Molecular Biology 2439

Springer Protocols

Gerhard Steger  
Hannah Rosenbach · Ingrid Span  
*Editors*



# DNAzymes

Methods and Protocols

 Humana Press

Volume 2439

# Methods in Molecular Biology

## Series Editor

John M. Walker

*School of Life and Medical Sciences, University of Hertfordshire, Hatfield, Hertfordshire, UK*

For further volumes: <http://www.springer.com/series/7651>

For over 35 years, biological scientists have come to rely on the research protocols and methodologies in the critically acclaimed *Methods in Molecular Biology* series. The series was the first to introduce the step-by-step protocols approach that has become the standard in all biomedical protocol publishing. Each protocol is provided in readily-reproducible step-by-step fashion, opening with an introductory overview, a list of the materials and reagents needed to complete the experiment, and followed by a detailed procedure that is supported with a helpful notes section offering tips and tricks of the trade as well as troubleshooting advice. These hallmark features were introduced by series editor Dr. John Walker and constitute the key ingredient in each and every volume of the *Methods in Molecular Biology* series. Tested and trusted, comprehensive and reliable, all protocols from the series are indexed in PubMed.

---

*Editors*

Gerhard Steger, Hannah Rosenbach and Ingrid Span

## **DNAzymes**

### **Methods and Protocols**



---

*Editors*

Gerhard Steger

Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf,  
Düsseldorf, Germany

Hannah Rosenbach

Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf,  
Düsseldorf, Germany

Ingrid Span

Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf,  
Düsseldorf, Germany

Department of Chemistry and Pharmacy, Friedrich-Alexander Universität,  
Nürnberg-Erlangen, Germany

ISSN 1064-3745

e-ISSN 1940-6029

Methods in Molecular Biology

ISBN 978-1-0716-2046-5

e-ISBN 978-1-0716-2047-2

<https://doi.org/10.1007/978-1-0716-2047-2>

© The Editor(s) (if applicable) and The Author(s), under exclusive license  
to Springer Science+Business Media, LLC, part of Springer Nature 2022

This work is subject to copyright. All rights are solely and exclusively  
licensed by the Publisher, whether the whole or part of the material is  
concerned, specifically the rights of translation, reprinting, reuse of  
illustrations, recitation, broadcasting, reproduction on microfilms or in any  
other physical way, and transmission or information storage and retrieval,  
electronic adaptation, computer software, or by similar or dissimilar  
methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service  
marks, etc. in this publication does not imply, even in the absence of a

specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Humana imprint is published by the registered company Springer Science+Business Media, LLC part of Springer Nature.

The registered company address is: 1 New York Plaza, New York, NY 10004, U.S.A.

---

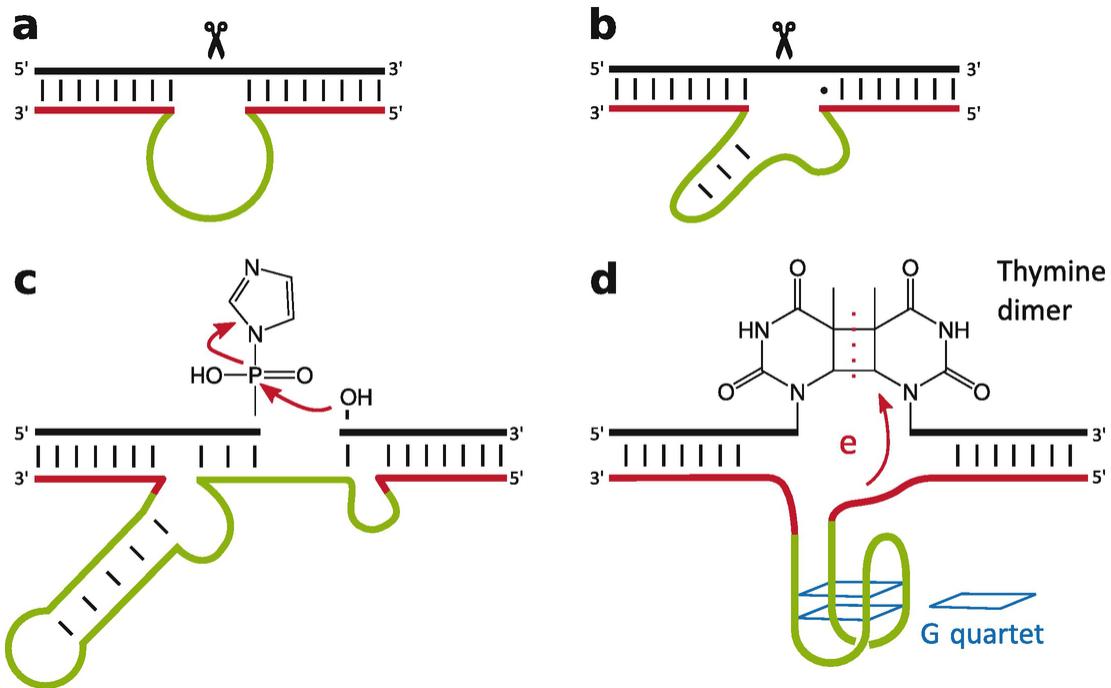
## Preface

Ribozymes—enzymes consisting of only RNA—were well established in 1990: Thomas Czech and coworkers had established the self-splicing of *Tetrahymena* group I intron [1, 2], Sidney Altman and coworkers had described the ribozyme nature of RNase P cleaving off a precursor of tRNA [3, 4], George Bruening and coworkers had detected the self-cleavage activity of satellite RNAs [5], and Robert Symons and coworkers had described the self-cleavage activity of viroids from family *Avsunviroidae* [6]. At this time, most scientists expected that such ribozyme activity would be restricted to RNA because DNA misses several features that were thought to be essential for catalytic activity: DNA occurs mostly not single-stranded and thus rarely forms complex tertiary structures, and DNA lacks the 2'-OH group of RNA. Thus, the first publication on a DNAzyme—DNA with enzymatic activity—by Ronald Breaker and Gerald Joyce [7] in 1994 came as a big surprise. Today, a broad variety of DNAzymes with different catalytic functions are known [8], including RNA cleavage [7], peptide modification [9, 10], phosphorylation [11], thymine dimer photoreversion [12], peroxidation [13], and DNA ligation [14].

Schematic representations of four selected DNAzymes are shown in Fig. 1.

The DNAzyme described by Breaker and Joyce in 1994 and most of the many DNAzymes described afterwards were found by *in vitro* selection starting from a random pool of DNA molecules, called systematic evolution of ligands by exponential enrichment (SELEX) [16]. The principle of SELEX to obtain molecule(s) with a desired function is based on the assumption that a three-dimensional structure, which is required for function, can be formed by many different single-stranded sequences. In theory, a random pool of sequences with 15 up to 50 nucleotides in length and the four different nucleotides comprises  $4^{15} \approx 1.2 \times 10^{18}$  up to  $4^{50} \approx 1.3 \times 10^{30}$  different sequences, covering all possible structures within the limits of the sequence length; in practice this number of random sequences is limited to about  $10^{14}$  to  $10^{15}$  sequences due to additionally required constant sequence parts, for example for amplification, hybridization, and fixation. One key step in the selection process is the separation of nucleic acids with the desired functionality from those which lack this property and therefore should be excluded from the pool in the

next selection round. The other key step is amplification of the relatively low amount of recovered nucleic acids with the desired functionality after each selection round by polymerase chain reaction (PCR) [17, 18], which requires fixed primer binding sites at the ends of the random sequence part.

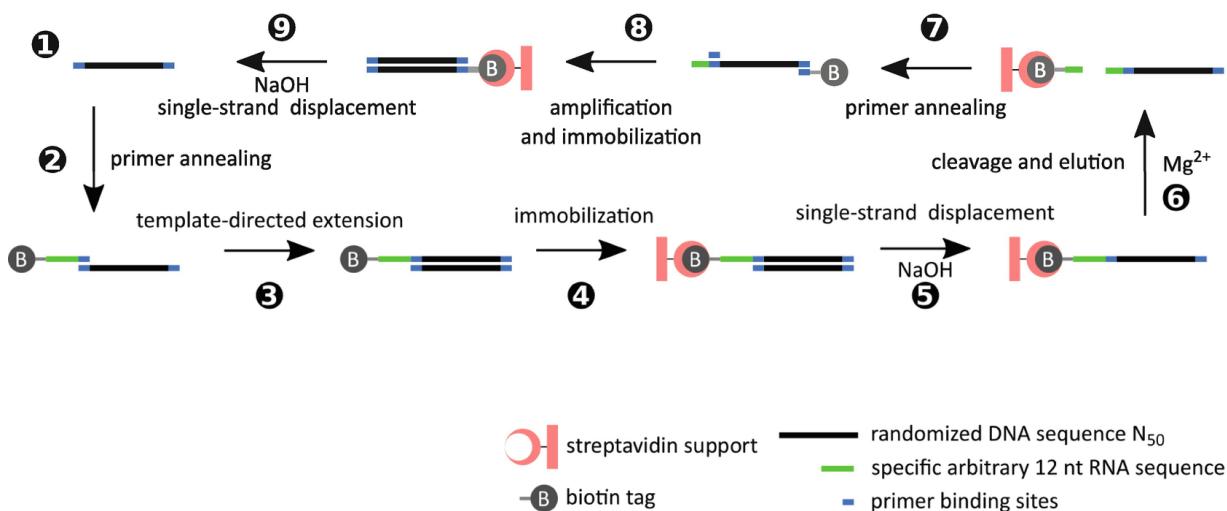


**Fig. 1** Schematic representation of different DNAzymes: The RNA-cleaving DNAzymes **(a)** 10-23 and **(b)** 8-17, **(c)** the DNA-ligating DNAzyme E47, and **(d)** the UV1C DNAzyme with photolyase activity. Catalytically important sequences are represented in green. The binding arms, which can vary in their sequence, are shown in red. The target sequences are shown in black. (Figure based on [15])

For example, the SELEX procedure (Fig. 2), which led in the 8th selection round to clone 17 (8-17 DNAzyme) and in the 10th selection round to clone 23 (10-23 DNAzyme), respectively, started with  $10^{14}$  different DNA molecules consisting of a 5' biotin moiety, followed by a short DNA spacer, 12 target ribonucleotides, few fixed deoxynucleotides, 50 random deoxynucleotides ( $N_{50}$ ), and finally few fixed deoxynucleotides [21]. These chimeric molecules were fixed to a streptavidin-coated support and eluted with a solution containing 10mM MgCl<sub>2</sub> at pH 7.5 and 37 °C. The eluted molecules were 3' fragments, which mostly originated from phosphoester cleavage in the target ribonucleotide stretch were recovered by PCR using the fixed regions surrounding the  $N_{50}$  region as sites for primer

hybridization. Then the now missing 5' parts were reintroduced and these molecules then used in the next round of selection. For further and important details, we refer to the original publication [21]. Figure 2 shows a schematic representation of the *in vitro* selection process for RNA-cleaving DNAzymes.

Similar procedures were used to produce the nowadays great variety of DNAzymes cleaving RNA in the presence of various metal ions and at various conditions [22, 23]. If one aims at other catalytic properties of DNAzymes than RNA cleavage, different immobilization and PCR methods are required; examples of such methods are given in Chapters 1 and 2 by Li et al. and Yang et al. SELEX requires the recovery or modification of single-stranded oligonucleotide sequence pools from mixed template libraries. Chapter 3 by Szokoli et al. provides two protocols for the PCR-based production of ssDNA molecules from low amounts of starting material.



**Fig. 2** *In vitro* selection process for the RNA-cleaving 10-23 DNAzyme. (1) The selection starts from a library of 10<sup>14</sup> randomized sequences with a length of 50 nucleotides (black). The DNA sequences are flanked by defined primer sites (blue) on each site. (2) The DNA sequences are annealed to a primer containing an RNA region of 12 nucleotides (red) as well as a biotin-tag (gray) on the 5'-end. (3) Template-directed extension is performed enzymatically resulting in double-stranded products. (4) The products are immobilized on a streptavidin column. (5) Treatment with NaOH to remove the strand that lacks the biotin-tag. (6) The column is then washed with the reaction buffer containing Mg<sup>2+</sup>. DNA sequences that are capable of adopting the correct folding for the catalysis of an RNA-cleavage reaction detach themselves from the column. (7) A new set of primers is added to the eluted sequences for amplification. (8) The amplification products are immobilized on

a streptavidin column. (9) Treatment with NaOH results in the release of the non-biotinylated strand. The released strand is then used for the next selection round. (Adapted from [19] and [20], based on [21])

For RNA-cleaving DNAzymes, once a DNA motif responsible for catalytic activity has been identified during the SELEX process, this fixed catalytic sequence part can be elongated by the addition of any target-specific “arms,” which bind to the RNA of interest via Watson-Crick base pairing. RNA-cleaving DNAzymes carry a high therapeutic potential by targeting RNA viruses or silencing mRNAs in diseases linked to elevated expression levels of certain proteins or noncoding RNAs. Nevertheless, in order to obtain optimal cleavage rates of the DNAzyme, the binding arms need to be carefully designed. A protocol for the design of an efficient RNA-cleaving DNAzyme with regard to optimal substrate:DNAzyme complex properties is provided in Chapter 4 by Steger and Victor.

Already during the SELEX procedure but even more urgent afterwards, questions about the activity of DNAzymes have to be answered [24]: what are the reaction rate, turnover frequency, catalytic rate constant  $k_{\text{cat}}$ , and Michaelis constant  $K_M$  of “my” DNAzyme? How do the reaction conditions like salt concentration, pH, etc. affect the rate and turnover of the DNAzyme? For this, early protocols used radioactively labeled RNA to determine percentage of cleaved RNA after gel separation of uncleaved and cleaved RNA. More recent protocols using fluorescence-based methods are described in Chapters 5, 6, and 18 by Rosenbach and Steger, Kirchgässler et al., and Kosmann, respectively. Chapter 7 by Wuebben and Schiemann demonstrate the use of electron paramagnetic resonance (EPR) spectroscopy to determine binding constants of  $\text{Mn}^{2+}$  to nucleic acids, which is particularly useful as  $\text{Mn}^{2+}$  can act as a probe for  $\text{Mg}^{2+}$  binding [24].

Although DNAzymes are under investigation for more than two decades, the biophysical and structural characterization of these versatile and diverse catalysts is still at the beginning. Until now, the best characterized DNAzymes are the RNA-cleaving DNAzymes 8-17 and 10-23 [21] as well as the RNA-ligating DNAzyme 9DB1 [25]. Only two high-resolution structures of DNA catalysts exist so far<sup>1</sup>: one structure of the RNA-cleaving 8-17 DNAzyme in complex with a DNA substrate [26] as well as one structure of the RNA-ligating DNAzyme 9DB1 in a post-

catalytic state [25]. Although X-ray crystallography is an excellent method to obtain insights into the structure of macromolecules on an atomic level, this technique is often impeded when applied to nucleic acids due to their surface properties, which includes a regularly ordered and negatively charged phosphate backbone. This may be overcome by co-crystallization of the nucleic acid with a nucleic acid-binding protein. Examples are the use of the RNA-binding protein U1A in Chapter 8 by Rosenbach and Span and of the African Swine Fever Virus Polymerase X in Chapter 9 by Liu et al.

The low chemical diversity of the DNA building blocks, the four deoxynucleotides, as well as flexible regions within the single-stranded DNA catalysts make the application of nuclear magnetic resonance (NMR) spectroscopy a challenging procedure for their structural investigation. These challenges, however, can be overcome as shown in Chapter 10 by Borggraefe and Etzkorn. Molecular modeling and dynamics simulations can give further insights into structural details, activity, and ion binding of DNAzymes; protocols for this method are described in Chapter 11 by Gertzen and Gohlke.

In addition to the abovementioned methods, a broad variety of different biophysical methods can be applied to DNAzymes in order to compensate for the lack of high-resolutinal structural insights and to study the folding and secondary structure of the DNAzymes and interactions with their substrates. Some of those methods are described in Chapters 12–16. While the protocols by Ahunbay et al., Vicino et al., and Bornevasser and Kath-Schorr describe different methods for the labeling of nucleic acids with fluorophores and/or spin labels, the Chapters by Hadzic et al. and Hett and Schiemann provide comprehensive protocols for the analysis and characterization of correspondingly modified nucleic acids using EPR and Förster resonance energy transfer (fluorescence resonance energy transfer; FRET) measurements.

Further biophysical methods that have previously been utilized for the investigation of DNAzymes, but have not been added to this methods book, include—among others—small-angle X-ray scattering (SAXS) [27], isothermal titration calorimetry (ITC) [28], and fluorescence correlation spectroscopy (FCS) [29].

Clearly, RNA-cleaving DNAzymes have therapeutic potential for example as gene silencers. In addition, they are promising tools for the

sensing of metal ions in environmental or biological samples [23]. The protocol provided in Chapter 17 by Huang and Liu describes the purification and application of a phosphorothioate- and fluorescence-modified DNAzyme for metal sensing. Furthermore, RNA-cleaving DNAzymes can be used as biological tools in the process of nucleic acid modification. Chapter 19 by LeVay and Mutschler describes a method to produce a 2', 3'-cyclic phosphate-functionalized RNA using a DNAzyme that catalyzes cleavage of a short 3'-RNA overhang in frozen solution.

This book aims at providing a collection of different protocols for the analysis and characterization of DNAzymes as well as the work with DNAzymes. Methods described in this book are as manifold and diverse as the DNAzymes themselves and range from bioinformatics and molecular dynamics simulations over classical wet-lab protocols for the analysis or modification of nucleic acids to the descriptions of spectroscopic, fluorescence-based, or crystallographic methods to obtain a fundamental understanding of the structure and function of these exquisite catalysts.

## References

1. Cech T, Zaug A, Grabowski P (1981) *In vitro* splicing of the ribosomal RNA precursor of *Tetrahymena*: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell* 27:487–496. [https://doi.org/10.1016/0092-8674\(81\)90390-1](https://doi.org/10.1016/0092-8674(81)90390-1)
2. Cech TR (1989) Nobel lecture: self-splicing and enzymatic activity of an intervening sequence RNA from *Tetrahymena*. <https://www.nobelprize.org/prizes/chemistry/1989/cech/lecture/>. Accessed 4 May 2021
3. Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35:849–857. [https://doi.org/10.1016/0092-8674\(83\)90117-4](https://doi.org/10.1016/0092-8674(83)90117-4)
4. Altman S (1989) Nobel lecture: enzymatic cleavage of RNA by RNA. <https://www.nobelprize.org/prizes/chemistry/1989/altman/lecture/>. Accessed 4 May 2021
5. Prody G, Bakos J, Buzayan J, Schneider I, Bruening G (1986) Autolytic processing of dimeric plant virus satellite RNA. *Science*

231:1577–1580. <https://doi.org/10.1126/science.231.4745.1577>

6. Hutchins C, Rathjen P, Forster A, Symons R (1986) Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid. Nucleic Acids Res 14:3627–3640. <https://doi.org/10.1093/nar/14.9.3627>
7. Breaker R, Joyce G (1994) A DNA enzyme that cleaves RNA. Chem Biol 1:223–229. [https://doi.org/10.1016/1074-5521\(94\)90014-0](https://doi.org/10.1016/1074-5521(94)90014-0)
8. Ponce-Salvatierra A, Boccaletto P, Bujnicki J (2021) DNAmoreDB, a database of DNAzymes. Nucleic Acids Res 49:D76–D81. <https://doi.org/10.1093/nar/gkaa867>
9. Chandrasekar J, Silverman S (2013) Catalytic DNA with phosphatase activity. Proc Natl Acad Sci U S A 110:5315–5320. <https://doi.org/10.1073/pnas.1221946110>
10. Chandrasekar J, Wylder A, Silverman S (2015) Phosphoserine lyase deoxyribozymes: DNA-catalyzed formation of dehydroalanine residues in peptides. J Am Chem Soc 137:9575–9578. <https://doi.org/10.1021/jacs.5b06308>
11. Li Y, Breaker RR (1999) Phosphorylating DNA with DNA. Proc Natl Acad Sci U S A 96:2746–2751. <https://doi.org/10.1073/pnas.96.6.2746>
12. Chinnappen DF, Sen D (2004) A deoxyribozyme that harnesses light to repair thymine dimers in DNA. Proc Natl Acad Sci U S A 101:65–69, <https://doi.org/10.1073/pnas.0305943101>
13. Travascio P, Li Y, Sen D (1998) DNA-enhanced peroxidase activity of a DNA aptamer-hemin complex. Chem Biol 5:505–517. [https://doi.org/10.1016/s1074-5521\(98\)90006-0](https://doi.org/10.1016/s1074-5521(98)90006-0)
14. Cuenoud B, Szostak J (1995) A DNA metalloenzyme with DNA ligase activity. Nature 375:611–614. <https://doi.org/10.1038/375611a0>
15. Silverman S (2016) Catalytic DNA: scope, applications, and biochemistry of deoxyribozymes. Trends Biol Sci 41:595–609. <https://doi.org/10.1016/j.tibs.2016.04.010>

16. Tuerk C, Gold L (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249:505–510. <https://doi.org/10.1126/science.2200121>
17. Saiki R, Scharf S, Faloona F, Mullis K, Horn G, Erlich H, Arnheim N (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230:1350–1354. <https://doi.org/10.1126/science.2999980>
18. Mullis KB (1990) Nobel lecture: the polymerase chain reaction. <https://www.nobelprize.org/prizes/chemistry/1993/mullis/lecture/>. Accessed 4 May 2021
19. Silverman S (2005) *In vitro* selection, characterization, and application of deoxyribozymes that cleave RNA. *Nucleic Acids Res* 33:6151–6163. <https://doi.org/10.1093/nar/gki930>
20. Kumar S, Jain S, Dilbaghi N, Ahluwalia AS, Hassan AA, Kim KH (2019) Advanced selection methodologies for DNAzymes in sensing and healthcare applications. *Trends Biochem Sci* 44(3):190–213. <https://doi.org/10.1016/j.tibs.2018.11.001>
21. Santoro SW, Joyce GF (1997) A general purpose RNA-cleaving DNA enzyme. *Proc Natl Acad Sci U S A* 94(9):4262–4266. <https://doi.org/10.1073/pnas.94.9.4262>
22. Zhang J (2018) RNA-cleaving DNAzymes: old catalysts with new tricks for intracellular and *in vivo* applications. *Catalysts* 8(11):550. <https://doi.org/10.3390/catal8110550>
23. Morrison D, Rothenbroker M, Li Y (2018) DNAzymes: selected for applications. *Small Methods* 2(3):1700319. <https://doi.org/10.1002/smtd.201700319>
24. Rosenbach H, Borggräfe J, Victor J, Wuebben C, Schiemann O, Hoyer W, Steger G, Etzkorn M, Span I (2020) Influence of monovalent metal ions on metal binding and catalytic activity of the

10-23 DNAzyme. Biol Chem 402:99–111. <https://doi.org/10.1515/hsz-2020-0207>

25. Ponce-Salvatierra A, Wawrzyniak-Turek K, Steuerwald U, Höbartner C, Pena V (2016) Crystal structure of a DNA catalyst. Nature 529(7585):231–234. <https://doi.org/10.1038/nature16471>
26. Liu H, Yu X, Chen Y, Zhang J, Wu B, Zheng L, Haruehanroengra P, Wang R, Li S, Lin J, et al. (2017) Crystal structure of an RNA-cleaving DNAzyme. Nat Commun 8(1):1–10. <https://doi.org/10.1038/s41467-017-02203-x>
27. Rosenbach H, Victor J, Borggräfe J, Biehl R, Steger G, Etzkorn M, Span I (2020) Expanding crystallization tools for nucleic acid complexes using U1A protein variants. J Struct Biol 210:107480. <https://doi.org/10.1016/j.jsb.2020.107480>
28. Rosenbach H, Borggräfe J, Victor J, Wuebben C, Schiemann O, Hoyer W, Steger G, Etzkorn M, Span I (2020) Influence of monovalent metal ions on metal binding and catalytic activity of the 10-23 DNAzyme. Biol Chem 402:99–111. <https://doi.org/10.1515/hsz-2020-0207>
29. Su X, Xiao X, Zhang C, Zhao M (2012) Nucleic acid fluorescent probes for biological sensing. Appl Spectrosc 66(11):1249–1262. <https://doi.org/10.1366/12-06803>
30. Borggräfe J, Victor J, Rosenbach H, Viegas A, Gertzen CGW, Wuebben C, Kovacs H, Gopalswamy M, Riesner D, Steger G, Schiemann O, Gohlke H, Span I, Etzkorn M (2022) Time-resolved structural analysis of an RNA-cleaving DNA catalyst. Nature 601 (7891):144–149. <https://doi.org/10.1038/s41586-021-04225-4>

**Gerhard Steger  
Hannah Rosenbach  
Ingrid Span  
Düsseldorf, Germany**

---

## Acknowledgments

We dedicate this book to Prof. em. Dr. Dr. h. c. Detlev Riesner on the occasion of his 80th birthday. He initialized the work with DNAzymes in the Institut für Physikalische Biologie, Heinrich-Heine-University Düsseldorf. His expertise and advice have always been a valuable and essential support for our work. Happy Birthday!

We gratefully acknowledge Prof. em. Dr. John M. Walker for his supportive and careful guidance during the editing process of this book.

We cordially thank all authors for their contribution to this methods book and their fruitful collaboration. Finally, we would like to thank Prof. Dr. R. K. O. Sigel for giving H. R. the opportunity to work on this book during her stay at the University of Zurich, Switzerland.

---

# Contents

## Preface

## Acknowledgments

## Contributors

## Part I Systematic Evolution of Ligands by Exponential Enrichment (SELEX)

### 1 A Protocol for Gold Nanoparticle-Assisted Aptamer Selection for a Small Molecule Porphyrin to Develop DNAzyme

Wenjing Li, Yanwei Cao and Renjun Pei

### 2 Selection of Aptamer for N-Methyl Mesoporphyrin IX to Develop Porphyrin Metalation DNAzyme

Luyan Yang, Yanwei Cao and Renjun Pei

### 3 PCR Methods for the Generation of Catalytic ssDNA

Deni Szokoli, Kristian K. Le Vay and Hannes Mutschler

## Part II Design: Kinetics, Performance, and Turnover

### 4 Design of a DNAzyme

Gerhard Steger and Julian Victor

### 5 Fluorescence-Based Kinetic Measurements for RNA-Cleaving DNAzymes

Hannah Rosenbach and Gerhard Steger

### 6 Stability and Activity of the 10–23 DNAzyme Under Molecular Crowding Conditions

Nina Kirchgässler, Hannah Rosenbach and Ingrid Span

### 7 Quantifying the Number and Affinity of Mn<sup>2+</sup>-Binding Sites with EPR Spectroscopy

Christine Wuebben and Olav Schiemann

## Part III Biophysical Characterization of DNAzymes

### 8 Obtaining Crystals of Nucleic Acids in Complex with the Protein U1A Using the Soaking Method

Hannah Rosenbach and Ingrid Span

### 9 Crystallization and Structural Determination of 8–17 DNAzyme

Hehua Liu, Song Mao, Jia Sheng and Jianhua Gan

**10 Solution NMR Spectroscopy as a Tool to Study DNAzyme Structure and Function**

Jan Borggräfe and Manuel Etzkorn

**11 Molecular Modeling and Simulations of DNA and RNA: DNAzyme as a Model System**

Christoph G. W. Gertzen and Holger Gohlke

**Part IV Further Biophysical Methods**

**12 Single-Molecule Kinetic Studies of Nucleic Acids by Förster Resonance Energy Transfer**

Mélodie C. A. S. Hadzic, Roland K. O. Sigel and Richard Börner

**13 Chemical Dual End-Labeling of Large Ribozymes**

Esra Ahunbay, Fabio D. Steffen, Susann Zelger-Paulus and Roland K. O. Sigel

**14 Spin Labeling of Long RNAs Via Click Reaction and Enzymatic Ligation**

Maria Francesca Vicino, Christine Wuebben, Mark Kerzhner, Michael Famulok and Olav Schiemann

**15 Preparation of Site-Specifically Spin-Labeled RNA by in Vitro Transcription Using an Expanded Genetic Alphabet**

Lisa Bornewasser and Stephanie Kath-Schorr

**16 PELDOR Measurements on Nitroxide-Labeled Oligonucleotides**

Tobias Hett and Olav Schiemann

**Part V Use of DNAzymes**

**17 Sensing Metal Ions with Phosphorothioate-Modified DNAzymes**

Po-Jung Jimmy Huang and Juewen Liu

**18 Enhancement of Activity for Peroxidase-Mimicking DNAzyme by Covalent Attachment of Hemin to G-Quadruplex-Forming Oligonucleotide Using Click Chemistry**

Joanna Kosman

**19 Generation of RNA with 2', 3'-Cyclic Phosphates by Deoxyribozyme Cleavage in Frozen Solutions**

Kristian K. Le Vay and Hannes Mutschler

## **Index**

---

## **Contributors**

### **Esra Ahunbay**

Department of Chemistry, University of Zurich, Zurich, Switzerland

### **Jan Borggräfe**

Institute of Physical Biology, Heinrich Heine University Düsseldorf,  
Düsseldorf, Germany

Institute of Biological Information Processing, IBI-7: Structural  
Biochemistry, Forschungszentrum Jülich, Jülich, Germany

### **Richard Börner**

Laserinstitut Hochschule Mittweida, University of Applied Sciences  
Mittweida, Mittweida, Germany

### **Lisa Bornewasser**

Department of Chemistry, University of Cologne, Köln, Germany

### **Yanwei Cao**

CAS Key Laboratory of Nano-Bio Interface, Suzhou Institute of Nano-Tech  
and Nano-Bionics, Chinese Academy of Sciences, Suzhou, China

### **Manuel Etzkorn**

Institute of Physical Biology, Heinrich Heine University Düsseldorf,  
Düsseldorf, Germany

Institute of Biological Information Processing, IBI-7: Structural  
Biochemistry, Forschungszentrum Jülich, Jülich, Germany

Jülich Center for Structural Biology (JuStruct), Forschungszentrum Jülich,  
Jülich, Germany

### **Michael Famulok**

Life & Medical Sciences Institute (LIMES), Chemische Biologie, c/o  
Kekulé-Institut für organische Chemie, Bonn, Germany

### **Jianhua Gan**

Shanghai Public Health Clinical Center, State Key Laboratory of Genetic  
Engineering, Collaborative Innovation Center of Genetics and  
Development, School of Life Sciences, Fudan University, Shanghai, China

**Christoph G. W. Gertzen**

Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany  
Center for Structural Studies (CSS), Heinrich Heine University Düsseldorf, Düsseldorf, Germany

**Holger Gohlke**

Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany  
John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), Institute of Biological Information Processing (IBI-7: Structural Biochemistry) and Institute of Bio- and Geosciences (IBG-4: Bioinformatics), Forschungszentrum Jülich GmbH, Jülich, Germany

**Mélodie C. A. S. Hadzic**

Department of Chemistry, University of Zurich, Zurich, Switzerland

**Tobias Hett**

Institute of Physical and Theoretical Chemistry, Rheinische Friedrich-Wilhelms-University, Bonn, Germany

**Po-Jung Jimmy Huang**

Department of Chemistry, Waterloo Institute for Nanotechnology, University of Waterloo, Waterloo, ON, Canada

**Stephanie Kath-Schorr**

Department of Chemistry, University of Cologne, Köln, Germany

**Mark Kerzhner**

Life & Medical Sciences Institute (LIMES), Chemische Biologie, c/o Kekulé-Institut für organische Chemie, Bonn, Germany

**Nina Kirchgässler**

Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany

**Joanna Kosman**

Department of Bioanalytical Chemistry, Faculty of Chemistry, Adam Mickiewicz University, Poznań, Poland

**Kristian K. Le Vay**

Biomimetic Systems, Max Planck Institute of Biochemistry, Martinsried, Germany

Faculty of Chemistry and Chemical Biology, TU Dortmund University, Dortmund, Germany

**Wenjing Li**

School of Nano-Tech and Nano-Bionics, University of Science and Technology of China, Hefei, China

CAS Key Laboratory of Nano-Bio Interface, Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences, Suzhou, China

**Hehua Liu**

Shanghai Public Health Clinical Center, State Key Laboratory of Genetic Engineering, Collaborative Innovation Center of Genetics and Development, School of Life Sciences, Fudan University, Shanghai, China

**Juewen Liu**

Department of Chemistry, Waterloo Institute for Nanotechnology, University of Waterloo, Waterloo, ON, Canada

**Song Mao**

Department of Chemistry and The RNA Institute, University at Albany, State University of New York, Albany, NY, USA

**Hannes Mutschler**

Biomimetic Systems, Max Planck Institute of Biochemistry, Martinsried, Germany

Faculty of Chemistry and Chemical Biology, TU Dortmund University, Dortmund, Germany

**Renjun Pei**

School of Nano-Tech and Nano-Bionics, University of Science and Technology of China, Hefei, China

CAS Key Laboratory of Nano-Bio Interface, Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences, Suzhou, China

**Hannah Rosenbach**

Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany

**Olav Schiemann**

Institute of Physical and Theoretical Chemistry, Rheinische Friedrich-Wilhelms-University, Bonn, Germany

**Jia Sheng**

Department of Chemistry and The RNA Institute, University at Albany, State University of New York, Albany, NY, USA

**Roland K. O. Sigel**

Department of Chemistry, University of Zurich, Zurich, Switzerland

**Ingrid Span**

Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany

**Fabio D. Steffen**

Department of Chemistry, University of Zurich, Zurich, Switzerland

**Gerhard Steger**

Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany

**Deni Szokoli**

Biomimetic Systems, Max Planck Institute of Biochemistry, Martinsried, Germany

Faculty of Chemistry and Chemical Biology, TU Dortmund University, Dortmund, Germany

**Maria Francesca Vicino**

Institute of Physical and Theoretical Chemistry, Rheinische Friedrich-Wilhelms University, Bonn, Germany

**Julian Victor**

Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf,  
Düsseldorf, Germany

**Christine Wuebben**

Institute of Physical and Theoretical Chemistry, Rheinische Friedrich-Wilhelms-University, Bonn, Germany

**Luyan Yang**

CAS Key Laboratory of Nano-Bio Interface, Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences, Suzhou, China

**Susann Zelger-Paulus**

Department of Chemistry, University of Zurich, Zurich, Switzerland

---

## Footnotes

<sup>1</sup> During typesetting of the book, Borggräfe et al. [30] elucidated structure and detailed mechanism of the 10–23 DNAzyme.

# **Part I**

## **Systematic Evolution of Ligands by Exponential Enrichment (SELEX)**

# 1. A Protocol for Gold Nanoparticle-Assisted Aptamer Selection for a Small Molecule Porphyrin to Develop DNAzyme

Wenjing Li<sup>1, 2</sup>, Yanwei Cao<sup>2</sup> and Renjun Pei<sup>1, 2</sup>✉

(1) School of Nano-Tech and Nano-Bionics, University of Science and Technology of China, Hefei, China  
(2) CAS Key Laboratory of Nano-Bio Interface, Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences, Suzhou, China

✉ Renjun Pei  
Email: [rjpei2011@sinano.ac.cn](mailto:rjpei2011@sinano.ac.cn)

## Abstract

DNAzymes can be obtained by in vitro selection, when specific analytes such as porphyrins are used as targets. In this process, the immobilization of targets serving as a major step is not suitable for the selection of DNAzymes because small molecules like porphyrins lack active groups for immobilization. Here, we develop a gold nanoparticle (AuNPs)-assisted aptamer selection for the small molecule porphyrin Zinc(II)-Protoporphyrin IX (ZnPPIX), which can realize the selection of DNAzyme without an immobilization step, and the detailed in vitro selection process in solution is described.

**Key words** DNAzyme – Aptamer – Gold nanoparticle – In vitro selection – Porphyrin

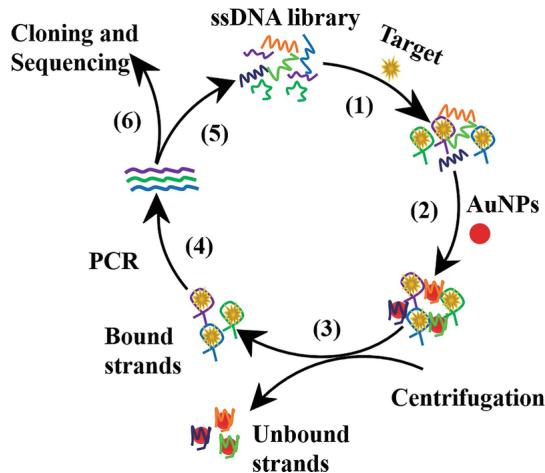
## 1 Introduction

The systematic evolution of ligands by exponential enrichment (SELEX) [1] is a combinatorial biology method used to obtain DNA/RNA aptamer molecules that can specifically interact with various analytes with high sensitivity and selectivity [2]. Specifically, these DNA aptamers that can be combined with targets like porphyrins usually exhibit catalytic properties, termed DNAzymes. However, different from large size targets such as biomacromolecules and cells, the aptamers selected for small molecule targets (such as porphyrins [3, 4], ATP [5], etc.) usually show relative lower affinities. Owing to the small differences of size, weight, and charge properties between target-binding complexes and unbound sequences, the general nitrocellulose filter-based SELEX method is not suitable for small molecule targets. SELEX methods for small molecule targets typically involve an immobilization step to facilitate separation that needs to chemically modify target molecules and then attach them to a solid matrix to separate the target-binding oligonucleotides from a large number of unbound oligonucleotides [6]. However, the immobilization of target molecules may not only alter their original conformations, but may induce the generation of steric hindrance at the interface leading to difficulty or even failure to the selection. Meanwhile, it is difficult to immobilize these small molecules without active groups.

Note that the graphene oxide (GO)-SELEX has proven to be an effective method for the selection of aptamers against small molecule targets without the immobilization step [7–9]. In such a process, the ssDNAs can adsorb on GO by π–π stacking interaction and can detach from nanomaterials by target molecules. Similar to GO, gold nanoparticles (AuNPs) can also adsorb ssDNA [10]. Specifically, the oligonucleotides folded by target-binding could be desorbed from the surface of AuNPs, resulting in the aggregation of AuNPs in salt solution accompanied by a solution color transition from red to purple-blue [11]. Based on this principle, various colorimetric aptasensors with simplicity and high sensitivity have been developed by using AuNPs [12–14].

In this work, we developed a gold nanoparticle-assisted aptamer selection for a small molecule porphyrin to develop DNAzymes. Briefly, the random ssDNA library was incubated with high concentration of targets, then the mixture was incubated with AuNPs to completely adsorb free ssDNA sequences; the target-binding sequences, which were not adsorbed on the surface of AuNPs, were separated from unbound sequences by centrifugation. The recovery rate of the selected ssDNA pool was performed to assess the enrichment efficiency of each round to determine if further selection is needed. This protocol has two obvious advantages: (1) there is no need to modify

the target molecule, so the screening process is further simplified and the cost is reduced; (2) relatedly, the use of unmodified small molecules can maximize the interaction with oligonucleotides, which may increase the binding affinities. Figure 1 displays an overview of our strategy for the selection of aptamers by using AuNPs as the separation matrix and Zinc(II)-Protoporphyrin IX (ZnPPIX) as the target molecule. Using this strategy, we successfully obtained the aptamer called ZnP1 and its truncated aptamer ZnP1.2 for ZnPPIX. Significantly, ZnP1-hemin and ZnP1.2-hemin complexes showed higher peroxidase activities than hemin alone, indicating that the selected aptamer has potential to be a peroxidase mimicking DNAzyme [15].



**Fig. 1** Overview of the in vitro selection process (from [15] reproduced with permission from Springer). (1) Incubation of the random ssDNA library with the target. (2) Adding AuNPs to the preincubated mixture of random ssDNA library and target. (3) Separation of the bound library elements from the unbound sequences. (4) Amplification of the binding sequences. (5) Generation and purification of the selection library. (6) Cloning and sequencing

## 2 Materials

### 2.1 Reagents and Buffers

All solutions are prepared with sterile Milli-Q water ( $18.2 \text{ M}\Omega \text{ cm}$ ) unless otherwise noted.

1. 1 mM ZnPPIX: Dissolve 0.01 g ZnPPIX (Frontier Scientific) in 3.19 mL dimethyl sulfoxide (DMSO), and then dilute 200  $\mu\text{L}$  of this stock solution with Milli-Q water to 1 mL, store at  $4^\circ\text{C}$  in the dark for use.
2. 100  $\mu\text{M}$  Fe(III) Protoporphyrin IX (Hemin): Add 0.01 g hemin in 3.06 mL DMSO, and then dissolve 20  $\mu\text{L}$  of this stock solution with Milli-Q water to 1 mL, store at  $4^\circ\text{C}$  in the dark.
3. 100  $\mu\text{M}$  N-methyl mesoporphyrin IX (NMM): Add 0.01 g NMM (Frontier Scientific) in 3.44 mL DMSO, and then dilute 20  $\mu\text{L}$  of this stock solution with Milli-Q water to 1 mL, store at  $4^\circ\text{C}$  in the dark.
4. 50 mM 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) diammonium salt (ABTS): Add 0.13717 g ABTS in 5 mL Milli-Q water and store at  $-20^\circ\text{C}$  in the dark.
5. 20 mM  $\text{H}_2\text{O}_2$ : Dilute 2.02  $\mu\text{L}$  30%  $\text{H}_2\text{O}_2$  (9.9 M,  $\rho = 1122 \text{ g/L}$ ) with Milli-Q water to 1 mL.
6. 1% Hydrogen Tetrachloroaurate(III) ( $\text{HAuCl}_4 \cdot 3\text{H}_2\text{O}$ ): Add 0.1159 g  $\text{HAuCl}_4 \cdot 3\text{H}_2\text{O}$  in 10 mL Milli-Q water.
7. Polymerase chain reaction (PCR) amplification reagents: 10  $\times$  PCR buffer ( $\text{Mg}^{2+}$  plus), dNTP mixture (2.5 mM each), Taq polymerase (5 U/ $\mu\text{L}$ ) (TaKaRa).
8. pMD18-T Vector Cloning Kit (TaKaRa).
- 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer: 100 mM HEPES, pH 7.4.
- 9.

10. SELEX buffer: 20 mM HEPES, pH 7.4, 60 mM NaCl.
- 11.
12. Binding buffer: 20 mM HEPES, pH 7.4, 140 mM NaCl.
13. Hemin-Binding (HB) buffer: 20 mM HEPES, pH 7.4, 60 mM NaCl, 0.5% Triton X-100 (w/v), and 1% DMSO (v/v).
14. 1× phosphate-buffered saline (PBS): Dissolve 8 g NaCl, 0.2 g KCl, 1.44 g Na<sub>2</sub>HPO<sub>4</sub>, 0.24 g KH<sub>2</sub>PO<sub>4</sub> in 1 L Milli-Q water, pH 7.4.
15. 3 M NaOAc: Add 12.3 g sodium acetate in 40 mL Milli-Q water, adjust the pH to 5.2 with glacial acetic acid, and add Milli-Q water to bring the volume up to 50 mL.
16. 1× Tris-buffered ethylenediaminetetraacetic acid (TBE): Add 53.9 g Tris-base, 27.5 g boric acid, 3.7224 g ethylenediaminetetraacetic acid (EDTA), in 5 L Milli-Q water, pH 8.0.
17. 4% agarose gel: Add 2 g agarose to an Erlenmeyer flask containing 50 mL 1 × TBE buffer, boil the sample in a microwave oven until it is clear and transparent, cool it to 65 °C, and pour into a plastic plate with a hole comb inserted.
18. Lysogeny broth (LB) liquid medium: Add 5 g Tryptone, 2.5 g Yeast Extract, 5 g NaCl into 500 mL Milli-Q water and autoclave at 121 °C and 0.1 MPa for 30 min, cool down to room temperature (RT), and then store at 4 °C.
19. LB solid medium: Add 0.9 g agar powder to every 60 mL of liquid LB medium, and autoclave at 121 °C for 30 min, and then heat and melt the LB solid medium. When it is cooled to below 56 °C, add 100 µL ampicillin (100 mg/mL) into 100 mL LB, mix and pour into the plate.

## 2.2 Oligonucleotides

1. Random ssDNA library with a random region of 40 nucleotides (Lib): 5'-ATACCAGCTTATTCAATT-40N-AGATAGTAAGTGCAATCT-3'.
2. Forward primer (p1): 5'-ATACCAGCTTATTCAATT-3'.
3. Reverse primer (p2): 5'-AGATTGCACTTACTATCT-3'.
4. 5' biotinylated reverse primer (bio-p2): 5'-Biotin-AGATTGCACTTACTATCT-3'.
5. ZnP1:  
5'ATACCAGCTTATTCAATTGGCGGGGGTTGCTCTACTTGATGATCTGCGCTATGCCGTAGATAGTAA  
3'.
6. ZnP1.2: 5'-GGCGGGGGTTGCTCTACTTGATGATCTGCGCTATGCCGT-3'.

## 2.3 Equipments

1. UV-Vis spectrophotometer.
2. Transmission electron microscope.
3. Amicon Ultra 0.5 mL 10 kDa Centrifugal filter.

4. Fluorescence spectrophotometer.
  5. PCR cycler.
  6. Micro Bio-Spin Chromatography Columns, empty.
- 

### 3 Methods

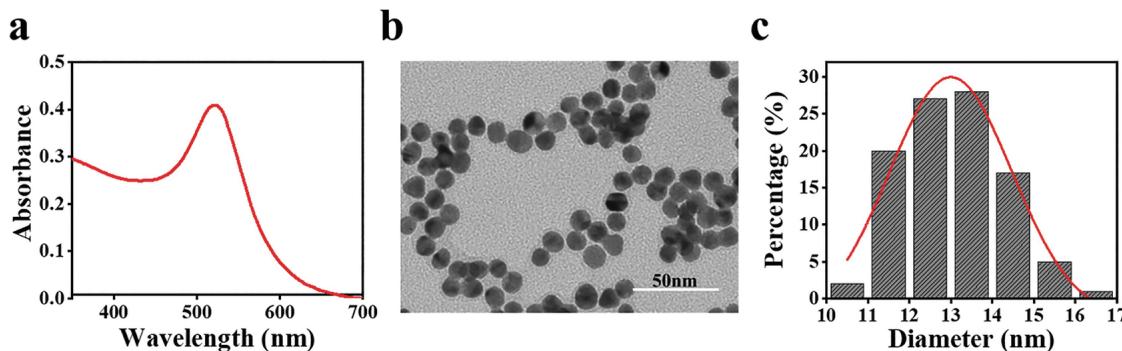
#### 3.1 Synthesis of Gold Nanoparticles

13 nm diameter AuNPs can be synthesized by the citrate reduction of  $\text{HAuCl}_4$ .

1. Bring an aqueous solution of tetrachloroauric acid (2 mL of 1% (w/v)  $\text{HAuCl}_4$  in 100 mL of Milli-Q water) to boil in a round-bottom flask with a magnetic stirrer and electric heating.
2. Add sodium citrate (8 mL of 1% (w/v) solution to the flask while stirring rapidly to yield colloidal gold particles (0.018% of (w/v)  $\text{HAuCl}_4$ ). Perform the synthesis of gold nanoparticles under boiling condition maintained by electric heating and simultaneous mixing by magnetic mixer at 1000 rpm.
3. When the color of the liquid changed to red, heat the solution for another 30 min and then cool to room temperature (RT). The AuNPs should be stored at 4 °C for future research.

#### 3.2 Characterization of Gold Nanoparticles

1. Concentrate 10 mL of the synthesis of citrate-stabilized AuNPs solution at  $9391 \times g$  for 15 min at 4 °C to 1 mL sterile Milli-Q water.
2. Detect the absorption band of AuNPs in the visible region by UV-Vis spectroscopy, and obtain the size and morphology of the particles by a transmission electron microscope (TEM). Figure 2 presents an example of data that was obtained.
3. Determine the concentration of these AuNPs used for experiments to be 16.7 nM by UV-Vis spectroscopy, based on an extinction coefficient of  $2.7 \times 10^8 \text{ M}^{-1} \text{ cm}^{-1}$  at  $\lambda = 520 \text{ nm}$  for 13 nm AuNPs.



**Fig. 2** Characterization of AuNPs. (a) The UV-Vis spectrum of particles. The freshly prepared AuNPs show a maximum surface plasmon resonance (SPR) absorption peak at 520 nm. (b) The TEM image and (c) the particle size distribution of AuNPs. The particles have a uniform size and morphology with an average size of about 13 nm. (Figures reproduced from [15] with permission from Springer)

#### 3.3 Selection Procedure Assisted by Gold Nanoparticles

1. For the first round of selection, dissolve the random library (200 pmol) in binding buffer, and then heat at 95 °C for 5 min and quickly cool on ice for 10 min, which is conducive to the formation of secondary structures of ssDNAs. Each 100 µL of binding solution contains 20 µL of 100 mM HEPES buffer (pH 7.4), 7 µL of 2 M NaCl, 2 µL of 100 µM Lib, and 71 µL Milli-Q water (see Note 1).

2. Dilute 20  $\mu$ L 1 mM ZnPPIX with binding buffer to a final concentration of 200  $\mu$ M. Then mix the 100  $\mu$ L solution containing library from **step 1** with 100  $\mu$ L 200  $\mu$ M ZnPPIX in a molar ratio of 1:100 and incubate at RT for 1 h in the dark.
3. Add 1 mL of 16.7 nM AuNPs to the above solution to a final volume of 2 mL in SELEX buffer and further incubate for 75 min at RT. The preparation of the selection solution is as follows: 16.7 nM AuNPs 1 mL, 100 mM HEPES buffer (pH 7.4) 360  $\mu$ L, 2 M NaCl 46  $\mu$ L, Milli-Q water 394  $\mu$ L, and the solution from **step 2** (see **Note 2**).
4. Divide the solution into 500  $\mu$ L per tube, and then centrifuge at  $6010 \times g$  for 1 h. Concentrate and purify the supernatant by centrifugal filter. Every 500  $\mu$ L can be concentrated to 50  $\mu$ L (see **Note 3**).
5. Calculate the accurate concentration of ssDNA on the basis of Lambert–Beer law by using the molar absorption coefficient and measuring the absorbance at 260 nm with a spectrophotometer.
6. Calculate the recovery rate by the ratio of the amount of ssDNA in the enriched pool to the added amount of ssDNA (see **Note 4**).

### 3.4 Amplification of the Binding Sequences

1. Pre-amplification: Perform 6 rounds of PCR pre-amplification of the library obtained from the above selection to prevent ssDNA loss caused by optimization of the number of PCR rounds and improper operation. The pre-amplification PCR reaction system is as follows: 10 $\times$  PCR buffer 10  $\mu$ L, 2.5 mM dNTPs 4  $\mu$ L, 50  $\mu$ M p1 2  $\mu$ L, 50  $\mu$ M Biotin-p2 2  $\mu$ L, template 2  $\mu$ L, TaKaRa Taq 0.5  $\mu$ L, Milli-Q water 79.5  $\mu$ L. The condition of PCR is 3 min at 95 °C, following by 6 cycles of 30 s at 94 °C, 30 s at 48 °C, and 15 s at 72 °C, and a final extension step for 5 min at 72 °C.
2. Cycle number optimization: Proceed 6, 8, 10, 12 rounds of amplification with the pre-amplified product as a template for re-amplification. The preparation method of the PCR round number optimization solution is as follows: 10 $\times$  PCR buffer 5  $\mu$ L, 2.5 mM dNTPs 2  $\mu$ L, 50  $\mu$ M p1 1  $\mu$ L, 50  $\mu$ M Biotin-p2 1  $\mu$ L, template 1  $\mu$ L, TaKaRa Taq 0.25  $\mu$ L, Milli-Q water 39.75  $\mu$ L. The amplified products are subjected to 4% agarose gel electrophoresis to choose the best PCR amplification cycles (see **Note 5**).
3. Large-scale PCR: In order to obtain at least 100 pmol of PCR product for the next round of selection, perform 20 individual 50  $\mu$ L PCR reactions simultaneously. Subject the pre-amplification products to PCR amplification again by the biotinylated reverse primer as follows: 10 $\times$  PCR buffer 100  $\mu$ L, 2.5 mM dNTPs 40  $\mu$ L, 50  $\mu$ M p1 20  $\mu$ L, 50  $\mu$ M Biotin-p2 20  $\mu$ L, template 20  $\mu$ L, TaKaRa Taq 5  $\mu$ L, Milli-Q water 795  $\mu$ L. Amplify the template under optimal cycles.

### 3.5 Generation and Purification of Selection Library

1. Incubate the PCR product with streptavidin-agarose beads for 30 min and add the mixture to the micro-spin column for sedimentation.
2. Wash the column with PBS three times to remove unbound DNA.
3. Denature the DNA in the column by addition of 0.1 M NaOH solution for 2 min, the desired ssDNA strands can be isolated from streptavidin-agarose beads and repeat this step once.
4. Neutralize the pH to 7.0–7.5 with 1.5 M HCl solution.
5. Purification by ethanol precipitation: Add 1/10 volume of 3 M NaOAc (pH 5.2) to the solution and mix the DNA solution thoroughly; then add two times volume of pre-cooled ethanol, mix well again, and place it at –20 °C for more than 30 min; next centrifuge at  $18,407 \times g$  for 25 min, and carefully pour out the supernatant; finally add 3/4 tube with 70% ethanol,  $18,407 \times g$  for 25 min, and carefully pour out the supernatant.

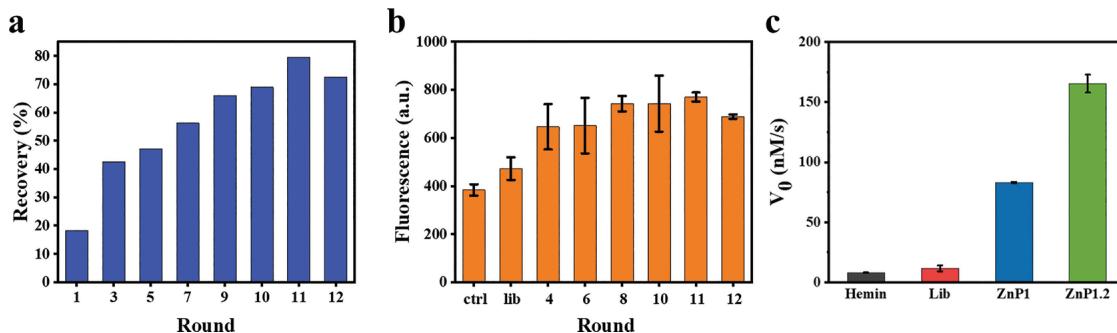
6. Place the uncapped tube in a hot blast cycle oven at 65 °C to evaporate the remaining liquid.
7. Dissolve the collected ssDNA library in 100 µL Milli-Q water, measure the absorbance with a spectrophotometer, and the concentration of ssDNA can be calculated by the molar absorption coefficient of the library.

### 3.6 Repeating Selection Cycle

After round 1, mix 100–200 pmol purified ssDNA pool with ZnPPIX with a molar ratio of 1:100 according to the identical procedure described above. Use the collected ssDNA pool as a secondary library for the next round of selection. Repeat the process until the affinity of the selected pool with the target no longer increases (see Notes 6–8).

### 3.7 Enrichment Monitoring

1. When determining if further selection is needed two things are required to assess the enrichment efficiency of each round: Evaluate the recovery rate during the selection process (see Subheading 3.3, step 6). Figure 3a presents an obtained example. Investigate the fluorescence enhancement of the enriched pool after each round for NMM to further determine the enrichment efficiency.
2. Dissolve the ssDNA pool of a given round in SELEX buffer, then heat to 95 °C for 5 min and cool down to RT.
3. Mix 1 µM of ssDNA pool with 1 µM of NMM in 100 µL SELEX buffer and incubate for 1 h in the dark at RT before measurement.
4. Set the excitation wavelength to 399 nm and record the fluorescence emission spectrum from 550 nm to 750 nm on a spectrofluorometer. Figure 3b presents an obtained example.



**Fig. 3** (a) Recovery rate of the enriched ssDNA pool. The ratio gradually increases with the increased number of the selected rounds. Until the 11th round, the recovery rate reaches a maximum, indicating the successful enrichment of the 11th pool. (b) Fluorescence intensity of NMM upon binding with the enriched ssDNA pool. Fluorescence intensity of NMM gradually increases before 11th rounds, and the fluorescence of the 12th round is slightly lower than that of the 11th round, which further confirms the successful enrichment of the pool. (c) Analysis of peroxidase activity. Hemin alone and hemin-library complexes have low native peroxidase activity, while both ZnP1 (full-length aptamer) and ZnP1.2 (truncated aptamer) exhibit enhanced peroxidase activity, of which ZnP1.2 is 22 times higher than that of hemin alone. Error bars represent the standard deviation from three separate trials. (Figures reproduced from [15] with permission from Springer)

### 3.8 Cloning and Sequencing

After confirming the enrichment, amplify the chosen ssDNA pool with unmodified primers, then clone the PCR product into *Escherichia coli* by PMD18-T TA cloning kit as follows:

1. Add 3 µL PCR product, 1 µL PMD18-T vector, 5 µL Solution I, and 1 µL Milli-Q water to one tube and mix them, and then ligate at 16 °C for 3 h in the PCR cycler.
2. Centrifuge the ligated product briefly and place it in ice-bath while thawing a tube of BL21 cells on ice. Add

- the ligated product into BL21 cells carefully and mix cells gently. Then place the cells on ice for 30 min.
3. After heat shock at exactly 42 °C in water bath for 90 s, immediately place it on ice to cool for 3–5 min without shaking.
4. Add 900 µL of fresh LB liquid medium into the mixture, and shake vigorously (200 rpm) for 45 min at 37 °C to resuscitate the cells. At this time, heat and melt the LB solid medium, when it is cooled to below 56 °C, add 100 µL ampicillin (100 mg/mL) into 100 mL LB, mix and pour into the plate.
5. After the resuscitation, centrifuge the sample at 6010 × g for 5 min. Discard the supernatant (about 800 µL), and resuspend the precipitate in 200 µL remained LB. Spread the precipitate onto a plate and incubate overnight at 37 °C.
6. Select 35 colonies randomly and put them separately into 10 mL centrifuge tubes with 2.5 mL ampicillin-containing LB liquid medium. Cultivate in a constant-temperature shaker at 37 °C for 5 h with shaking at 200 rpm.
7. Send the clones for sequencing.

### 3.9 Measurement of Peroxidase Activity

After sequencing, evaluate the peroxidase activity of the obtained sequences by using the ABTS-H<sub>2</sub>O<sub>2</sub> system for characterization.

1. Dissolve DNA in SELEX buffer, denature at 95 °C for 5 min, and then cool down to RT.
2. Mix the hemin (1 µM) with 2 µM of ssDNA in 80 µL HB buffer and incubate them for 1 h at RT.
3. Add 10 µL 20 mM H<sub>2</sub>O<sub>2</sub> and 10 µL 5 mM ABTS to above solution to initiate the reaction.
4. Measure the absorbance versus time profiles with an UV-Vis spectrophotometer at 414 nm for 60 s using 1 cm path-length quartz cuvettes (see Note 9).
5. Calculate the initial rate ( $V_0$ ) from the first 10 s of the reaction time curve ( $\Delta\varepsilon = 36,000 \text{ M}^{-1} \text{ cm}^{-1}$ ). Figure 3c presents an obtained example.

---

## 4 Notes

1. The size of the 40 nucleotides position generated by randomly mixing A, C, G, and T bases is chosen according to our experience, because it can provide adequate results for the majority of the intended applications.
2. The salt concentration in buffer and the incubation time of ssDNA library and target should be optimized to obtain the best adsorption conditions of DNA on AuNPs. In our experiment, target is incubated with AuNPs in SELEX buffer in the presence of 60 mM NaCl for 75 min before centrifugation to minimize the non-specific, non-adsorbed ssDNA.
3. The reason for concentration and purification of the supernatant after separating the binding sequences is that the components of the collected solution are complex. For example, the solution containing an excessive amount of target has a certain degree of influence on subsequent PCR. Concentrating and purifying the selected library before doing PCR, and adding a small volume of concentrated library as a PCR template can reduce its impact on PCR amplification efficacy.
4. The recovery rate of the selected ssDNA pool after each round can be used to assess the enrichment efficiency of each round to determine if further selection is needed. The selected pool can be concentrated before PCR amplification, and its amount should be determined, then recovery rate can be calculated by the ratio of the amount of ssDNA in the enriched pool to the added amount of ssDNA. The stronger the binding of the selected pool to target, the relatively more amount of ssDNA would be collected, correspondingly making a

higher recovery rate.

5.

The PCR reaction can be modified to allow the production of single-stranded oligonucleotides. If the PCR reaction produces aberrant products such as higher molecular weight DNA products, smear, etc. several tests reactions with various cycle numbers can be realized.

6.

In vitro selection is a reiterative protocol, with multiple rounds of selection, even though only the first round is described in detail here. The number of rounds is affected by trade-offs between sequence counts, the range of affinities represented, and selection bias. In general, data from previous rounds of selection cover a wider range of relative affinities, while selected DNAs from later rounds are biased toward higher affinity motifs. However, the counts of each individual sequence are low in the first few rounds of selection, so more rounds might be necessary in some cases.

7.

Cleanliness and good laboratory practices are important for the success of in vitro selection. It is important to avoid cross-contamination during the selection procedure.

8.

It is imperative to save a half of the ssDNA pool at -20 °C for archival purposes. If subsequent amplification or selection reactions fail for any reason, the procedure can be started from the previous round.

9.

The catalytic activity reflects the oxidation rate of ABTS by H<sub>2</sub>O<sub>2</sub> through the generation of oxidized product ABTS<sup>•+</sup>. After adding ABTS and H<sub>2</sub>O<sub>2</sub>, it is necessary to measure immediately.

## Acknowledgments

This work was financially supported by the Natural Science Foundation of China (21575154, 21775160, 81801837, 31800685) and the Science Foundation of Jiangsu Province (BE2016680, BE2018665, BK20180250, BK20180258, BK20180261).

---

## References

1. Tuerk C, Gold L (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249:505–510  
[Crossref]
2. Stoltenburg R, Reinemann C, Strehlitz B (2007) SELEX—A (r)evolutionary method to generate high-affinity nucleic acid ligands. *Biomol Eng* 24:381–403  
[Crossref]
3. Li Y, Geyer R, Sen D (1996) Recognition of anionic porphyrins by DNA aptamers. *Biochemistry* 35:6911–6922  
[Crossref]
4. Travascio P, Li Y, Sen D (1998) DNA-enhanced peroxidase activity of a DNA aptamer-hemin complex. *Chem Biol* 5:505–517  
[Crossref]
5. Huizenga DE, Szostak JW (1995) A DNA aptamer that binds adenosine and ATP. *Biochemistry* 34:656–665  
[Crossref]
6. Ruscito A, DeRosa MC (2016) Small-molecule binding aptamers: selection strategies, characterization, and applications. *Front Chem* 4:14  
[Crossref]
7. Park J-W, Tatavarty R, Kim DW, Jung H-T, Gu MB (2012) Immobilization-free screening of aptamers assisted by graphene oxide. *Chem Commun* 48:2071–2073  
[Crossref]
8. Nguyen VT, Kwon YS, Kim JH, Gu MB (2014) Multiple GO-SELEX for efficient screening of flexible aptamers. *Chem Commun* 50:10513–10516  
[Crossref]
9. Gu H, Duan N, Wu S, Hao L, Xia Y, Ma X, Wang Z (2016) Graphene oxide-assisted non-immobilized SELEX of okdaic acid aptamer and the analytical application of aptasensor. *Sci Rep* 6:21665

[\[Crossref\]](#)

10. Liu J (2012) Adsorption of DNA onto gold nanoparticles and graphene oxide: surface science and applications. *Phys Chem Chem Phys* 14:10485–10496  
[\[Crossref\]](#)
11. Li H, Rothberg L (2004) Colorimetric detection of DNA sequences based on electrostatic interactions with unmodified gold nanoparticles. *Proc Natl Acad Sci U S A* 101:14036–14039  
[\[Crossref\]](#)
12. Yang C, Wang Y, Marty JL, Yang X (2011) Aptamer-based colorimetric biosensing of Ochratoxin A using unmodified gold nanoparticles indicator. *Biosens Bioelectron* 26:2724–2727  
[\[Crossref\]](#)
13. Peng Y, Li L, Mu X, Guo L (2013) Aptamer-gold nanoparticle-based colorimetric assay for the sensitive detection of thrombin. *Sens Actuators B* 177:818–825  
[\[Crossref\]](#)
14. Liu J, Bai W, Niu S, Zhu C, Yang S, Chen A (2014) Highly sensitive colorimetric detection of 17 $\beta$ -estradiol using split DNA aptamers immobilized on unmodified gold nanoparticles. *Sci Rep* 4:7571  
[\[Crossref\]](#)
15. Li W, Luo Y, Gao T, Yang L, Wang J, Pei R (2019) In vitro selection of DNA aptamers for a small-molecule porphyrin by Gold nanoparticle-based SELEX. *J Mol Evol* 87:231–239  
[\[Crossref\]](#)

## 2. Selection of Aptamer for N-Methyl Mesoporphyrin IX to Develop Porphyrin Metalation DNAzyme

Luyan Yang<sup>1</sup>, Yanwei Cao<sup>1</sup> and Renjun Pei<sup>1, 2</sup> 

- (1) CAS Key Laboratory of Nano-Bio Interface, Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences, Suzhou, China  
(2) School of Nano-Tech and Nano-Bionics, University of Science and Technology of China, Hefei, China

 **Renjun Pei**

Email: [rjpei2011@sinano.ac.cn](mailto:rjpei2011@sinano.ac.cn)

### Abstract

Mesoporphyrin IX (MPIX) contains a planar macrocycle center that can interact with various divalent metal ions through the exposed binding sites, leading to the metalation of MPIX. The DNA aptamers for porphyrin molecules usually display different catalytic functions (termed deoxyribozymes or DNAzymes), which can accelerate such chemical reactions. Inspired by this, an affinity chromatography selection approach was designed for identifying a porphyrin metalation DNAzyme. In our experiment, *N*-methyl mesoporphyrin IX (NMM), an analog of MPIX, is used as the target molecule, owing to its stable and high fluorescence enhancement after combining with specific oligonucleotides. Our results showed that the selected aptamer Nm1 is capable of binding to NMM with a low micromolar dissociation constant ( $0.75 \pm 0.08 \mu\text{M}$ ) and displays a catalytic activity for MPIX metalation with 3.3-fold rate enhancement. The

protocol for isolation of such a porphyrin metalation DNAzyme is described in detail here.

**Key words** Aptamer – Capture-SELEX – DNAzyme – Metalation – Porphyrin

---

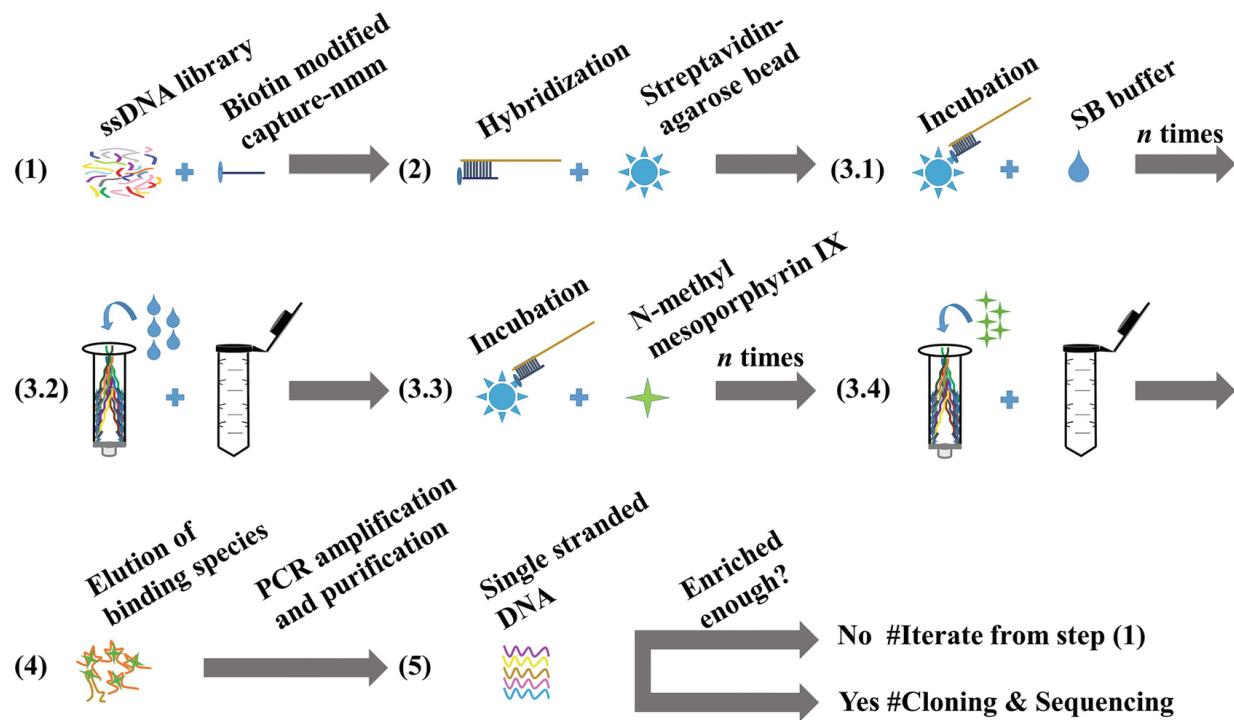
## 1 Introduction

Nucleic acid aptamers are short, single-stranded molecules that can specifically recognize an enormously wide range of targets, thanks to their characteristic spatial structures obtained by the systematic evolution of ligands by exponential enrichment (SELEX) method [1, 2]. A large number of aptamers with specific recognition functions have been developed over the past three decades [3–5]. Among them, some RNA or DNA aptamers against porphyrin molecules usually display catalytic activities, such as peroxidase activity or porphyrin metalation [6–9], termed ribozymes or deoxyribozymes (abbreviated as DNAzymes, respectively). However, the catalytic molecules involving catalytic activities are RNAzymes [10, 11] or G-quadruplexes [8, 12], which are unstable or greatly affected by environmental factors, such as pH or the concentration of metal ions. Thus, it is imperative to obtain a kind of DNAzymes with stable conformations.

In this work, we utilized a modified affinity chromatography method to get a porphyrin metalation DNAzyme. This capture-SELEX method is established by hybridizing a random ssDNA library with a short biotin-modified capture sequence on the streptavidin-coated beads, and can realize the separation of binding and unbinding oligonucleotides by incubation and elution with the target. The unbound sequences still stay in the selection column, but the binding sequences will be released, PCR-amplified, and then survive into the next selection round. Thus, the interaction between collected binding sequences and target should be much stronger than the interaction between collected binding sequences with the complement domain in this capture sequence. Compared to the conventional method of immobilizing the target on a solid substrate matrix, this capture-SELEX method avoids the possible change of properties after modification, the inaccessibility of partial target surface, and is very suitable for screening aptamers against these small molecules that have complex structures or are difficult to immobilize. Until now, aptamers for various small molecules,

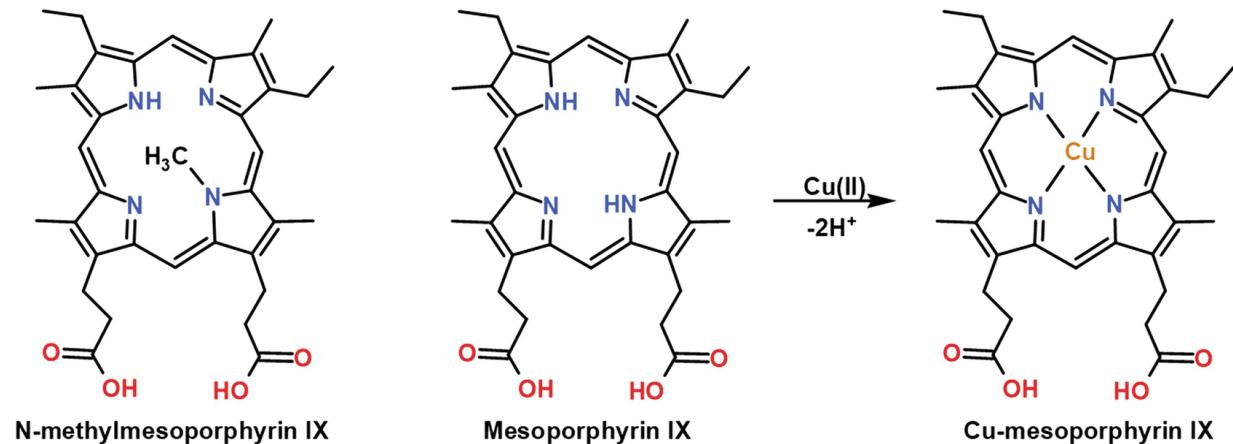
such as penicillin [13], ractopamine [14], and furaneol [15], have been successfully selected by using this method, since it can reduce or even eliminate non-specific binding during SELEX.

In vitro selection for the DNAzyme described here was performed by using NMM, an analog of Mesoporphyrin IX (MPIX) [16, 17], as the target molecule. This is because NMM serving as a commercially available and unsymmetrical anionic porphyrin usually exhibits the stable and high fluorescence enhancement when being combined with a specific ligand [16, 17], which can be utilized to easily monitor the screening process. The schematic diagram of the overall capture-SELEX is represented in Fig. 1 and the structure of NMM is shown in Fig. 2. Finally, we successfully isolated aptamer Nm1 that exhibits catalytic activity for the metal insertion reaction of mesoporphyrin IX and can be developed as a new light-up fluorescent probe.



**Fig. 1** Schematic illustration of the selection procedure. (1) Hybridization of capture-nmm and the random ssDNA library. (2) Incubation of the DNA mixture with the streptavidin-coated agarose beads in the selection column. (3.1) Washing the column with the SB buffer several times ( $n$ ) to remove the uncaptured DNA sequences and to prepare the column for SELEX; (3.2) collecting the eluates as  $S_n$  samples; (3.3) adding *N*-methyl mesoporphyrin IX solutions several times ( $n$ ) and incubating with the column to bind with the pre-incubated ssDNA library; (3.4) collecting the eluates as  $N_n$  samples. (4) Collecting the

elutions by target solutions to obtain the complexes of target and binding sequences and monitoring the selection progress by using collected elutions and additional wash. (5) Cloning and sequencing are applied for isolation of aptamers when the pool is enriched enough



**Fig. 2** The structure of *N*-methyl mesoporphyrin IX and the insertion reaction of Cu(II) into mesoporphyrin IX. The central macrocycle is often called the porphinato core; the presence of the methyl group in NMM distorts the planar macrocycle. The idea of the DNAzyme is likely to accelerate the metalation reaction by distorting the planar porphyrin toward the transition state of metalation

## 2 Materials

### 2.1 Oligonucleotides

Handling of HPLC-purified oligonucleotides is described below (see **Note 1**).

1. Lib-30N-nmm (the 76-nt ssDNA library containing 30 random bases):  
GGAGGCTCTCGGGACGAC-N<sub>30</sub>-  
GTCGTCCCGATGCTGCAATCGTAAGAAT.
2. Capture-nmm: GTCGTCCCGAGAGGCCATA-biotin.
3. P1-nmm (forward primer): GGAGGCTCTCGGGACGAC.
4. P2-nmm (reverse primer): ATTCTTACGATTGCAGCATCGGGAC.
5. Biotin-p2-nmm: biotin-ATTCTTACGATTGCAGCATCGGGAC.

6. Nm1:  
GGACGACCGACCGAAGGGAGGGATGGGTACATCTGTCGTCC.

## 2.2 Solutions and Buffers

Some solutions need to be autoclaved at 121 °C, 0.1 MPa for 30 min, including Tris–HCl buffer, SB buffer (SELEX buffer), washing buffer, MB buffer (Metalation buffer), and LB medium (Lysogeny broth medium, liquid and solid). All solutions and buffers are prepared with sterile pure water and then stored at 4 °C unless noted otherwise.

1. 200 mM Tris–HCl buffer: Add 4.85 g tris(hydroxymethyl)aminomethane (Tris) in 150 mL pure water, adjust the pH to 7.4 with concentrated HCl (12 N), and adjust the volume to 200 mL.
2. SB buffer: Firstly, prepare 5 M NaCl (2.922 g NaCl in 10 mL pure water), 1 M KCl (0.7455 g KCl in 10 mL pure water), and 1 M MgCl<sub>2</sub> (0.9521 g MgCl<sub>2</sub> in 10 mL pure water). Then take 10 mL of 200 mM Tris–HCl buffer, 3 mL of 5 M NaCl, 500 µL of 1 M KCl, and 200 µL of 1 M MgCl<sub>2</sub> to prepare 100 mL SB buffer with a final concentration of 20 mM Tris–HCl, 150 mM NaCl, 5 mM KCl, and 2 mM MgCl<sub>2</sub>, pH 7.4.
3. Washing buffer: 20 mM Tris–HCl, pH 7.4.
4. Stock solutions of NMM and MPIX: Firstly, dissolve 5 mg NMM and 10 mg MPIX in 1 mL of dimethyl sulfoxide (DMSO) for 8.6 mM NMM and 15.6 mM MPIX, respectively. Then prepare 500 µL of 2 mM NMM and MPIX, and store them in the dark at –20 °C.
5. Solutions of NMM: Dissolve the 2 mM NMM solutions in SB buffer to 100 µM, 20 µM, 15 µM, 10 µM, etc., that are stored at –20 °C in the dark and ready for use (*see Note 2*).
6. Solution of MPIX: Dissolve 2 mM MPIX into MB buffer. Need to use 5 µL of 2 mM MPIX to a final concentration of 100 µM in metalation reaction.

**reacuon.**

7.  $1 \times$  TBE buffer: Dissolve 107.8 g Tris, 55 g boric acid, and 7.4448 g EDTA Na<sub>2</sub> in 10 L pure water, which is the running buffer of gel electrophoresis.
8. 4% agarose gel: Add 2 g agarose powder into 50 mL  $1 \times$  TBE buffer in an Erlenmeyer flask, and heat in a microwave until completely dissolved to transparency. Pour the agarose solution into a gel tray with a 3 mm thick comb in place immediately after cooling to about 60 °C. When the gel is completely solidified, store it at 4 °C (*see Note 3*).
9. 20 bp DNA Marker and 6× loading buffer are used for gel electrophoresis (TaKaRa).
10. 10× PCR buffer with Mg<sup>2+</sup>, 2.5 mM dNTPs mixture, and 5 U/ $\mu$ L rTaq polymerase are used for PCR amplification (TaKaRa).
11. LB medium (liquid): Add 5 g tryptone, 2.5 g yeast extract, and 5 g sodium chloride into 500 mL of pure water, and shake the container to dissolve.
12. LB medium (solid): Add 0.9 g agar powder to 60 mL of LB liquid medium and place in an Erlenmeyer flask.
13. 100 mg/mL ampicillin solution: Add 0.1 g ampicillin sodium to 1 mL pure water, mix and store at -20 °C.
14. 200 mM Tris-HOAc buffer: Add 4.85 g Tris in 150 mL pure water, adjust the pH to 7.4 with concentrated HOAc (17.5 N), and make the volume to 200 mL.
15. MB buffer: Take 1 mL of 200 mM Tris-HAc buffer, 300  $\mu$ L of 5 M NaCl, 50  $\mu$ L of 1 M KCl, 20  $\mu$ L of 1 M MgCl<sub>2</sub>, 50 mg of Triton X-100, and 100  $\mu$ L of DMSO to prepare 10 mL MB buffer with a final concentration of 25 mM Tris-HAc, 150 mM NaCl, 5 mM KCl, 2 mM MgCl<sub>2</sub>, 0.5% Triton X-100 (w/v), and 1% DMSO (v/v), pH 7.4.

16. UV-Vis spectrophotometer.
  17. PCR cycler.
  18. Horizontal electrophoresis tank.
  19. Luminescent image analyzer.
  20. Fluorescence spectrometer.
- 

### 3 Methods

#### 3.1 Selection Procedure

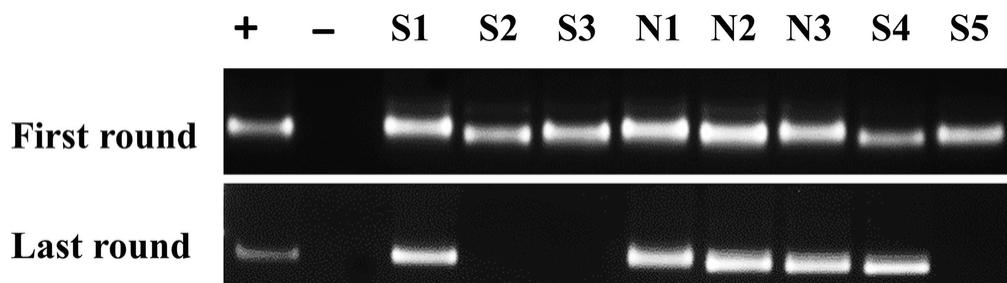
1. Hybridization of lib-30N-nmm and capture-nmm: A ssDNA library with  $10^{14-15}$  different oligonucleotides that are composed of a 30-nt random region and a biotinylated capture sequence at 3' end is employed. For a given round, hybridize capture-nmm with the ssDNA library (1000 pmol lib-30N-nmm at first round) at a molar ratio of 2:1 in 200  $\mu$ L SB buffer, seal the tube, and denature in a water bath at 95 °C for 5 min, and then naturally cool to room temperature.
2. Pretreatment of a column: Add 200  $\mu$ L streptavidin agarose beads in an empty column (micro bio-spin empty chromatography column) and pre-equilibrate it with 200  $\mu$ L SB buffer thrice.
3. Incubation of the hybridized DNA mixture and the pre-equilibrium column: For the first round, incubate the hybridized DNA sample with the column for 30 min. This combination allows the ssDNA library to be indirectly immobilized on the affinity column.
4. Competitive elution: In the following,  $n$  represents the number of screening rounds; in each repeat eight eluates are sampled in 2 mL centrifuge tubes. Firstly, wash the column with SB buffer three times to remove the unbound DNA and label eluates as  $nS1$ ,  $nS2$ ,  $nS3$  samples; secondly, perform positive selection with NMM solution (100  $\mu$ M at first round) three times. Specific binding occurs between the flexible oligonucleotide and the target at potential binding sites. The DNAs

~~single nucleotide and the target at potential binding sites. The DNA~~  
binding to NMM can be competitively eluted from the affinity column and corresponding eluates are labeled as *nN1*, *nN2*, *nN3* samples; thirdly, wash twice with SB buffer and label the eluates as *nS4*, *nS5* sample.

5. Small-scale PCR reaction: Prepare 50  $\mu$ L PCR sample containing 10  $\mu$ L 10 $\times$  PCR buffer, 4  $\mu$ L of 2.5 mM dNTPs, 2  $\mu$ L of 50  $\mu$ M p1-nmm, 2  $\mu$ L of 50  $\mu$ M p2-nmm, 1  $\mu$ L template, 0.25  $\mu$ L rTaq, and 39.75  $\mu$ L pure water. Specifically, the template is either the positive control (5 nM lib-30N-nmm), negative control (pure water), or samples (*nS1*, *nS2*, *nS3*, *nN1*, *nN2*, *nN3*, *nS4*, and *nS5* elutions; *see step 4*). Set the PCR program in a PCR cycler as follows: pre-denaturation at 95 °C for 5 min, cycles of denaturation at 92 °C for 15 s, annealing at 58 °C for 30 s, and extension at 72 °C for 20 s (cycle number range: 6–14). Optimize the cycle numbers of PCR, and perform all subsequent experiments under the optimized PCR conditions (*see Note 4*).
6. Enrichment monitoring: Mix 10  $\mu$ L of each sample with 2  $\mu$ L of 6 $\times$  loading buffer containing SYBR Green I before adding to a 4% agarose gel. Carry out the electrophoresis at 150 V for 40 min in a horizontal electrophoresis tank. Stain the small-scale PCR products with SYBR Green I by gel electrophoresis, and monitor the enrichment of the ssDNA library according to the brightness difference of each band by a luminescent image analyzer. The screening will stop iteration when the sub-library is enriched. Two models of elutions of the first and the last round are shown in Fig. 3.
7. Large-scale PCR reaction: Prepare 1 mL PCR sample containing 100  $\mu$ L 10 $\times$  PCR buffer, 40  $\mu$ L of 2.5 mM dNTPs, 20  $\mu$ L of 50  $\mu$ M p1-nmm, 20  $\mu$ L of 50  $\mu$ M biotin-p2-nmm, 20  $\mu$ L template, 5  $\mu$ L rTaq, and 795  $\mu$ L pure water. Amplify the template, which is the mixture of *nN1*, *nN2*, *nN3*, and *nS4* (*see step 4*), by the optimized PCR conditions (*see step 5*) (*see Note 5*).
8. Generation and purification of sub-library: Add 200  $\mu$ L streptavidin agarose beads into a column and pre-equilibrate it with washing buffer three times (*see Note 6*). Incubate the column with PCR-amplified dsDNA with a non-use chain biotinylated at 3' end for 20 min three times. Further wash it five times by 500  $\mu$ L washing buffer to remove

unbound elements. The strong interaction between streptavidin and biotin helps us to separate the dsDNA under denaturing condition. Incubate 150  $\mu$ L of 0.1 M NaOH with the column for 90 s twice to make a biotin-labeled chain still binding to streptavidin beads, while a biotin-free chain is detached and then these ssDNAs are neutralized with 1.5 M HCl to pH 7.4 for the next round (see Note 7).

9. Repeat steps 1–8 for the next round: Pay attention to control the amount of ssDNA pool from 500 to 200 pmol and NMM from 100 to 10  $\mu$ M, and gradually reduce the amount of them, until we get an enriched library specifically binding to NMM (see Note 8).



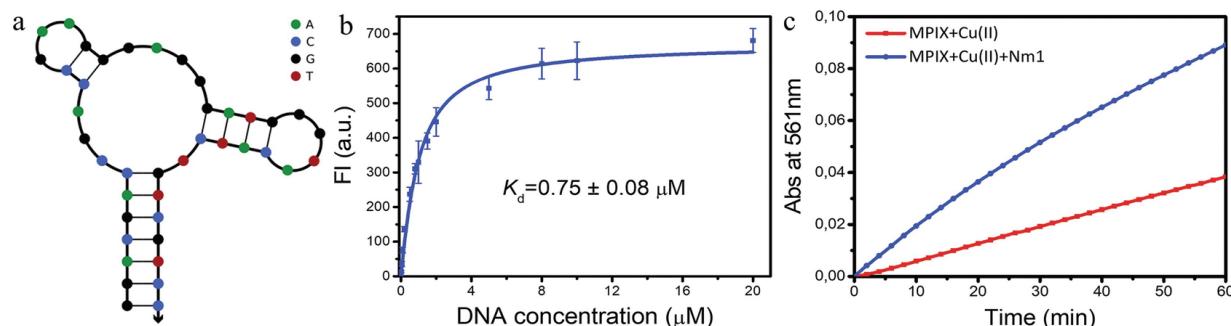
**Fig. 3** Two examples of elutions from the first and last rounds. In the initial round, the brightness of all the elution bands is at a similar intensity. After enough rounds of selection, the brightness of S2 and S3 bands seems to remarkably decrease but the N bands show strong fluorescent intensity. The significant differences between S3 and N1 bands in 4% agarose gel electrophoresis prompt us to stop the SELEX process and begin cloning

## 3.2 Cloning and Sequencing

1. After 11 rounds of screening, PCR-amplify the DNA enriched library with unmodified primers for cloning. According to the small-scale PCR conditions, check the dsDNA product for gaining appropriate brightness by gel electrophoresis.
2. Add 1  $\mu$ L pMD18-T vector and 5  $\mu$ L Solution I from pMD18-T TA Cloning Kit (TaKaRa), 2  $\mu$ L PCR product and 1  $\mu$ L pure water to a 300  $\mu$ L tube. After gently mixing, perform ligation reaction at 16 °C for 2–4 h in a PCR cycler.
3. Centrifuge the ligated product briefly and place it on ice. Gently add the ligation solution to 100  $\mu$ L BL-21 competent cells on ice for

15 min.

4. After heating the above liquid for 90 s at 42 °C, immediately put it on ice to cool for 3–5 min. Be careful not to shake during this period because of the transformation.
5. Add 900 µL of fresh liquid LB medium to the above liquid and resuscitate at 37 °C, 200 rpm for 60 min. At this time, melt the solid LB medium and when the temperature drops below 55 °C, add 60 µL, 100 mg/mL ampicillin solution into 60 mL solid LB medium, mix and pour the plate.
6. Spin the samples at  $6010 \times g$  for 5 min, discard 800 µL of supernatant and resuspend the cells in the remaining solution. Gently blow the recovered cells repeatedly, spread them over the plate, and incubate the plate at 37 °C overnight (see **Note 9**).
7. Pick about 30 monoclonal colonies on the plate and place them separately in 1.5 mL centrifuge tubes. Cultivate at 200 rpm until the solution is turbid, take 1 mL of each solution and send them for sequencing.
8. Analyze DNA secondary structures of various DNA sequences [18] to find truncated and optimized candidate sequences, then select and synthesize aptamer candidate chains. The secondary structure of Nm1 is shown in Fig. 4a.



**Fig. 4** (a) Predicted secondary structure of NMM aptamer by using the NUPACK software. (b) Binding curve of Nm1. (c) The formation of metalloporphyrin is monitored by the spectral value change at 561 nm (No aptamer, red square; Nm1, blue circle). Data

lines are obtained from three separate trials. (Adapted with permission from Ref. 20. Copyright 2019 American Chemical Society)

### 3.3 Fluorescent Measurement

1. Pretreatment of DNA sequence: Dissolve the selected oligonucleotides in SB buffer, heat at 95 °C for 5 min, and slowly cool to room temperature for 30 min (*see Note 10*).
2. Mix 5 μM candidate aptamer and 1 μM NMM in 100 μL SB buffer, and incubate at 25 °C for 30 min before measurement. Then measure the fluorescence spectra of the mixture of 5 μM DNA strand and 1 μM NMM at RT.
3. Set the excitation wavelength to 399 nm, the emission spectrum in the wavelength range of 550–750 nm in 1 nm increment, and the slit widths to 10 nm on a fluorescence spectrophotometer.
4. The spectrum shows a considerable fluorescence enhancement after the addition of aptamer Nm1. Therefore, collect the fluorescence intensity of NMM at the maximum emission wavelength of 610 nm and use in subsequent calculations.

### 3.4 Job–Plot Curve

1. Perform the fluorescence titration method to obtain the binding ratio of Nm1 and NMM by controlling the total concentration of NMM and Nm1 at 5 μM.
2. Change the DNA percentage in the mixture from 0% to 100%; that means 0, 0.25, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 4.75, 5 μM of Nm1 and correspondingly 5, 4.75, 4.5, 4, 3.5, 3, 2.5, 2, 1.5, 1, 0.5, 0.25, 0 μM of NMM in 100 μL SB buffer, respectively.
3. After setting the concentration of Nm1 and NMM, measure the fluorescence intensity at 610 nm according to fluorescent measurement method (*see Subheading 3.3*).
4. Make a job–plot scatter plot [19], and determine the stoichiometric

binding ratio through executing two linear fitting lines by Origin software. The binding ratio of Nm1 and NMM is 1.

### 3.5 Determination of $K_d$

1. Calculate the  $K_d$  of aptamer-target complex by measuring the fluorescence of 0.5  $\mu\text{M}$  NMM with increasing concentrations of Nm1 (0, 0.01, 0.05, 0.1, 0.2, 0.5, 0.8, 1, 1.5, 2, 5, 8, 10, 20  $\mu\text{M}$ ) in 100  $\mu\text{L}$  SB buffer (*see Note 11*).
2.  $K_d$  is obtained by fitting the nonlinear binding curve using the following equation:

$$F - F_0 = \frac{F_{\max} - F_0}{2} \left[ \frac{K_d}{C_{\text{dye}}} + 1 + \frac{nC_{\text{DNA}}}{C_{\text{dye}}} - \sqrt{\left( \frac{K_d}{C_{\text{dye}}} + 1 + \frac{nC_{\text{DNA}}}{C_{\text{dye}}} \right)^2 - 4 \frac{nC_{\text{DNA}}}{C_{\text{dye}}}} \right]$$

where  $F$  and  $F_0$  are the fluorescent intensity of NMM with and without aptamers;  $n$  is the binding ratio of aptamer and dye (*see Subheading 3.4*);  $C_{\text{DNA}}$  and  $C_{\text{dye}}$  represent the concentration of Nm1 and 0.5  $\mu\text{M}$  NMM, respectively [20]. Use the nonlinear fitting in Origin software to obtain  $F_{\max}$  and  $K_d$ . Figure 4b shows the  $K_d$  of isolated aptamer Nm1.

### 3.6 Metal Insertion Reaction

1. The diagram of the porphyrin metalation reaction is shown in Fig. 2. Measure the UV-Vis spectra of porphyrin and aptamer-porphyrin after the addition of  $\text{Cu(OAc)}_2$  by UV-Vis spectroscopy.
2. Heat 5  $\mu\text{M}$  DNA aptamer in MB buffer at 95 °C for 5 min, and then slowly cool to RT. Add 5  $\mu\text{L}$  of 2 mM MPIX and incubate with the pretreated DNA aptamer for 30 min to prepare the MPIX-Nm1 sample.
3. Mix 200  $\mu\text{M}$   $\text{Cu(OAc)}_2$  solution with 100  $\mu\text{M}$  MPIX solution and the MPIX-Nm1 sample (*see step 2*), respectively.
4. Initial rate of metalation is obtained by calculating the absorbance change at 561 nm in the first 10 min. The product Cu(II)-MPIX is

monitored by the small change in Q band and exhibits absorption

difference at 561 nm with a molar absorptivity of  $18.6 \text{ mM}^{-1} \text{ cm}^{-1}$  [21].

5. Time-dependent absorbance changes are shown in Fig. 4c. Investigate the initial rate of insertion by the absorbance increase at 561 nm to assess the activity of DNAzyme. Evidently, Nm1 can accelerate the insertion reaction of Cu(II) into MPIX (see Note 12).
- 

## 4 Notes

1. The lyophilized DNA powder provided by the company is centrifuged and dissolved in sterile Milli-Q water. We obtain the molar absorption coefficient ( $\epsilon$ ) of each DNA strand [22], measure the absorbance at 260 nm by UV-Vis spectrophotometer, and then calculate the concentration by the Beer–Lambert law. The final concentration of ssDNA is 100  $\mu\text{M}$  and the stock solution should be stored at  $-20^\circ\text{C}$ .
2. Measure the concentration of NMM by UV-Vis spectrophotometer (the extinction coefficient of NMM is  $1.45 \times 10^5 \text{ M}^{-1} \text{ cm}^{-1}$  at 379 nm). Be careful to avoid light when using NMM. If the operation is improper, repeat the measurement using original stock solution.
3. The gel can be stored for ~1 month and be used repeatedly if the internal sample runs out or to the top of each lane, which can quite save time and materials.
4. Two fixed sequences flanking the randomized region in the library sequence are designed for primer annealing. The purpose of small-scale PCR is to find a suitable cycle number for large-scale PCR amplification and to obtain amplified DNA products for agarose gel electrophoresis. If the number of cycles cannot be determined, one can try 6, 8, 10, and 12 PCR cycles, determine the concentration of DNA products with an agarose gel, and estimate the maximum optimal number of cycles.  
Importantly, biotin-p2-nmm is used for the next step to produce

5. ssDNA. One needs to perform gel electrophoresis to ensure that a sufficient amount of dsDNA is obtained.
6.  

Washing buffer is used to remove the unbound DNA and promote buffer exchange completely to avoid magnesium hydroxide precipitation.
7. From the next screening, the collected ssDNA pool can use the same protocol, but its concentration should be measured by UV-Vis spectrophotometer, and the amount of sub-library and capture-nmm should be recalculated for the next cycle.
8. A successful screening depends on many factors. In the selection process, it is important to prevent the samples from being contaminated by anyone and anywhere. Careful inspection, analysis, and adjustment of the screening parameters are conducive to the smooth running of each step and contribute to the success.
9.  
All operations should be performed carefully and gently; try to work in a clean bench. A successful spread plate should have isolated bacterial colonies evenly distributed on the plate.
10.  
The amount of sequence is pre-calculated (*see Subheading 3.3, step 2*), which means that the volume of NMM should be deducted from the total volume of 100  $\mu$ L in advance.
11.  
In order to obtain a more accurate  $K_d$  value, more data points should be measured near the end of the reaction.
12. We isolate aptamers for NMM, a distorted porphyrin, to develop porphyrin metalation DNAzyme since the porphyrin distortion is an important step in the metallization of porphyrin. Besides, theoretically, not only aptamers with catalytic effect can be obtained, but it is also possible to obtain non-catalytic sequences. Therefore, we need to select sequences with the analysis of affinity, selectivity, and specificity when detecting the catalysis of aptamers. Note that the accuracy and specificity of selected aptamers can be further enhanced by adding negative selection.

## Acknowledgments

This work was financially supported by the Natural Science Foundation of China (21575154, 21775160, 81801837, 31800685), the CAS/SAFEA International Innovation Teams program, and the Science Foundation of Jiangsu Province (BE2016680, BE2018665, BK20180250, BK20180258, BK20180261).

---

## References

1. Ellington AD, Szostak JW (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature* 346(6287):818–822  
[\[Crossref\]](#)
2. Tuerk C, Gold L (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249(4968):505–510  
[\[Crossref\]](#)
3. Hermann T, Patel DJ (2000) Adaptive recognition by nucleic acid aptamers. *Science* 287(5454):820–825  
[\[Crossref\]](#)
4. Tan W, Wang H, Chen Y, Zhang X, Zhu H, Yang C, Yang R, Liu C (2011) Molecular aptamers for drug delivery. *Trends Biotechnol* 29(12):634–640  
[\[Crossref\]](#)
5. Yan S-R, Foroughi MM, Safaei M, Jahani S, Ebrahimpoor N, Borhani F, Baravati NRZ, Aramesh-Boroujeni Z, Foong LK (2020) A review: recent advances in ultrasensitive and highly specific recognition aptasensors with various detection strategies. *Int J Biol Macromol* 155:184–207  
[\[Crossref\]](#)
6. Willner I, Shlyahovsky B, Zayats M, Willner B (2008) DNAzymes for sensing, nanobiotechnology and logic gate applications. *Chem Soc Rev* 37(6):1153–1165  
[\[Crossref\]](#)
7. Teller C, Shimron S, Willner I (2009) Aptamer–DNAzyme hairpins for amplified biosensing. *Anal Chem* 81(21):9114–9119  
[\[Crossref\]](#)
8. Li W, Li Y, Liu Z, Lin B, Yi H, Xu F, Nie Z, Yao S (2016) Insight into G-quadruplex-hemin DNAzyme/RNAzyme: adjacent adenine as the intramolecular species for

remarkable enhancement of enzymatic activity. *Nucleic Acids Res* 44(15):7373–7384  
[Crossref]

9. Norouzi A, Ravan H, Mohammadi A, Hosseinzadeh E, Norouzi M, Fozooni T (2018) Aptamer–integrated DNA nanoassembly: a simple and sensitive DNA framework to detect cancer cells. *Anal Chim Acta* 1017:26–33  
[Crossref]
10. Conn MM, Prudent JR, Schultz PG (1996) Porphyrin metalation catalyzed by a small RNA molecule. *J Am Chem Soc* 118(29):7012–7013  
[Crossref]
11. Ren R, Wang L-L, Ding T-R, Li X-M (2014) Enzyme-free amplified detection of nucleic acids based on self-sustained replication of RNase and its application in tumor cell detection. *Biosens Bioelectron* 54:122–127  
[Crossref]
12. Li Y, Sen D (1996) A catalytic DNA for porphyrin metallation. *Nat Struct Biol* 3(9):743–747  
[Crossref]
13. Paniel N, Istamboulié G, Triki A, Lozano C, Barthelmebs L, Noguer T (2017) Selection of DNA aptamers against penicillin G using capture-SELEX for the development of an impedimetric sensor. *Talanta* 162:232–240  
[Crossref]
14. Duan N, Gong W, Wu S, Wang Z (2017) An ssDNA library immobilized SELEX technique for selection of an aptamer against ractopamine. *Anal Chim Acta* 961:100–105  
[Crossref]
15. Komarova N, Andrianova M, Glukhov S, Kuznetsov A (2018) Selection, characterization, and application of ssDNA aptamer against furaneol. *Molecules* 23(12):3159  
[Crossref]
16. Guo L, Nie D, Qiu C, Zheng Q, Wu H, Ye P, Hao Y, Fu F, Chen G (2012) A G-quadruplex based label-free fluorescent biosensor for lead ion. *Biosens Bioelectron* 35(1):123–127  
[Crossref]
17. Yett A, Lin LY, Beseiso D, Miao J, Yatsunyk LA (2019) N-methyl mesoporphyrin IX as a highly selective light-up probe for G-quadruplex DNA. *J Porphyr Phthalocyanines* 23:1195–1215  
[Crossref]

18. Zadeh JN, Steenberg CD, Bois JS, Wolfe BR, Pierce MB, Khan AR, Dirks RM, Pierce NA (2011) NUPACK: analysis and design of nucleic acid systems. *J Comput Chem* 32(1):170–173  
[\[Crossref\]](#)
19. Huang CY (1982) Determination of binding stoichiometry by the continuous variation method: the job plot. In: Purich DL (ed) *Methods enzymol*, vol 87. Academic Press, Cambridge, pp 509–525
20. Yang L, Ding P, Luo Y, Wang J, Lv H, Li W, Cao Y, Pei R (2019) Exploration of catalytic nucleic acids on porphyrin metalation and peroxidase activity by in vitro selection of aptamers for N-methyl mesoporphyrin IX. *ACS Comb Sci* 21(2):83–89  
[\[Crossref\]](#)
21. Yang J, Bowser MT (2013) Capillary electrophoresis–SELEX selection of catalytic DNA aptamers for a small-molecule porphyrin target. *Anal Chem* 85(3):1525–1530  
[\[Crossref\]](#)
22. Owczarzy R, Tataurov AV, Wu Y, Manthey JA, McQuisten KA, Almabazi HG, Pedersen KF, Lin Y, Garretson J, McEntaggart NO (2008) IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Res* 36(2):163–169  
[\[Crossref\]](#)

### 3. PCR Methods for the Generation of Catalytic ssDNA

Deni Szokoli<sup>1, 2</sup>, Kristian K. Le Vay<sup>1, 2</sup> and Hannes Mutschler<sup>1, 2</sup>✉

(1) Biomimetic Systems, Max Planck Institute of Biochemistry, Martinsried, Germany

(2) Faculty of Chemistry and Chemical Biology, TU Dortmund University, Dortmund, Germany

✉Hannes Mutschler

Email: [hannes.mutschler@tu-dortmund.de](mailto:hannes.mutschler@tu-dortmund.de)

#### Abstract

The ability to produce single-stranded DNA on a preparative scale from low amounts of starting templates is necessary for most research involving deoxyribozymes, but is particularly important for performing *in vitro* selections. While the production of single-stranded RNA is straightforward by means of *in vitro* transcription, the enzymatic production of single-stranded DNA (ssDNA) on a preparative scale is often difficult. Nevertheless, several methods for the production of ssDNA have been published over the years. Here, we present two PCR methods that we find to be particularly effective, fast, and affordable, which we have adapted for our own needs.

**Key words** Polymerase chain reaction (PCR) – ssDNA production – 10–23 DNAzyme – ssDNA library recovery – DNAzyme preparation

---

### 1 Introduction

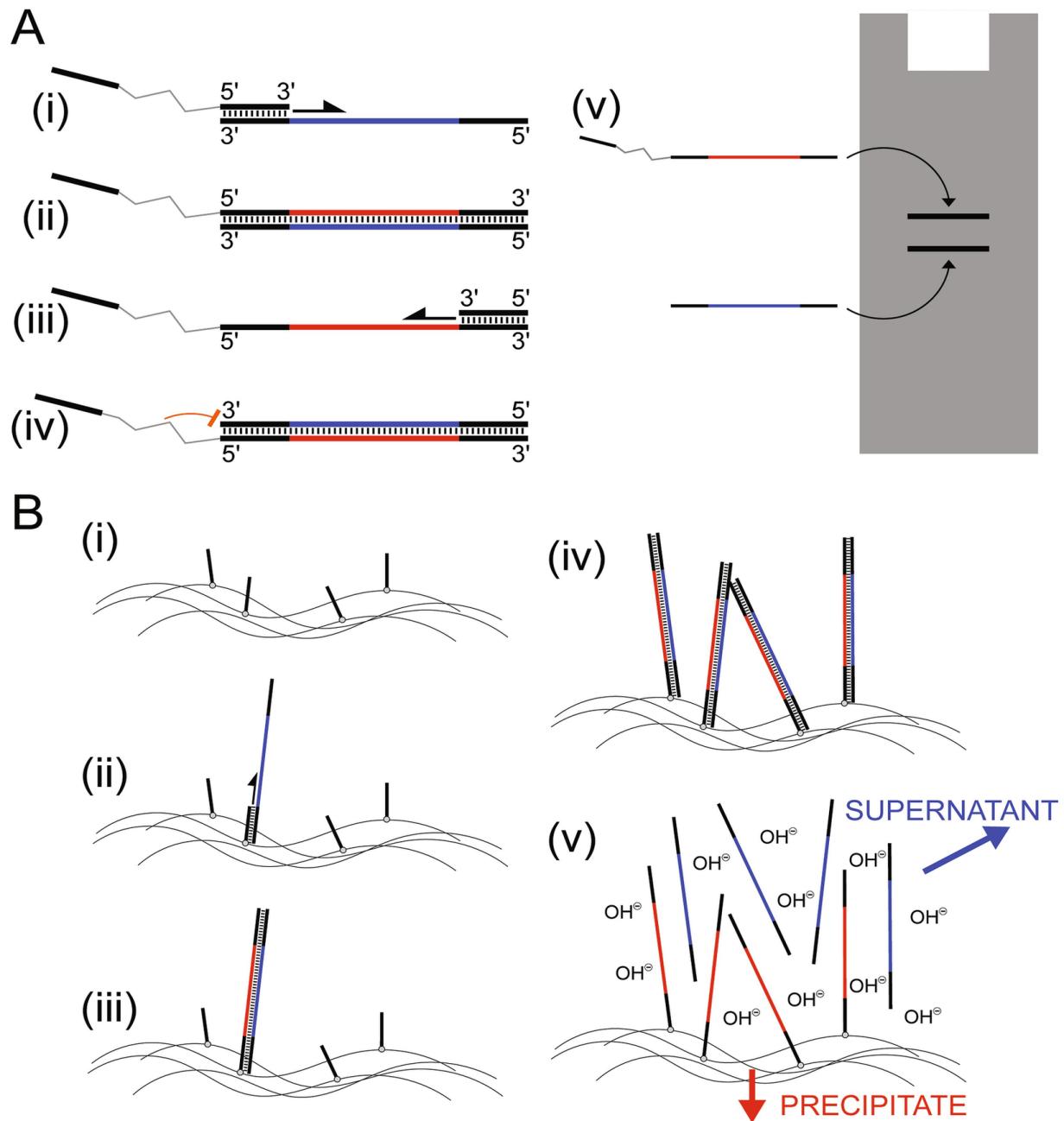
In almost all life forms, DNA predominantly exists in the double-stranded form. The relative absence of naked ssDNA in living systems makes acquiring it by enzymatic means non-trivial, and so producing DNA molecules such as deoxyribozymes that can only fold into their active form when single-stranded is a challenging task.

There are presently several methods by which one can produce ssDNA, the most straightforward of which is solid-phase synthesis using phosphoramidite building blocks [1]. While this method yields large amounts of pure, single-stranded oligonucleotides, there are limitations, such as difficulty in producing strands longer than 200 nucleotides (nt) [2]. Moreover, solid-phase synthesis is not suitable when single-stranded oligonucleotides need to be amplified from a pool of mixed templates, which is of particular importance for the recovery or modification of functional sequence pools from libraries during *in vitro* selection experiments.

Several publications describe the use of modified polymerase chain reactions (PCR), combined with additional enzymatic steps, in order to generate ssDNA in high yields. Typical examples include the use of exonucleases, such as lambda exonuclease, to selectively degrade one of the two strands of the PCR product [3]. Other methods rely on techniques such as asymmetric PCR, which involves using an excess of one primer to generate different amounts of each strand, effectively producing dsDNA as well as excess ssDNA [4].

More advanced methods have also been developed based on chemically modified primers. William and Bartel used primers with internal triethylene glycol linkers that prevent 3'-polymerase elongation beyond a defined sequence [5]. Using these “primer-terminators” during PCR (here named

PriTer-PCR) yields dsDNA products with strands of unequal length, which can be separated by denaturing polyacrylamide gel electrophoresis (PAGE) (see Fig. 1a).

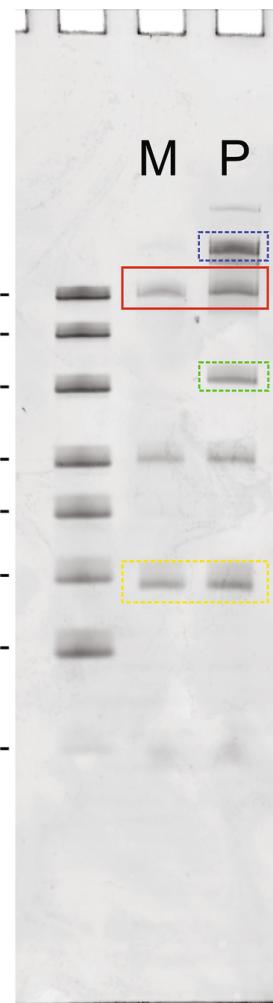


**Fig. 1** Illustrations of PCR techniques. **(a)** PriTer-PCR: (i) A 5'-A<sub>20</sub>-hexaethyleneglycol modified primer anneals to the ssDNA and is extended by a DNA polymerase. (ii) A complementary strand which is, e.g., 20 nt longer is produced. (iii) An unmodified primer anneals to the modified template generated in the previous step and is extended by a DNA polymerase. (iv) The primer is extended up to the spacer, but cannot be extended past it, thus terminating the extension. This produces a product 20 nt shorter than the template. (v) The PCR product can be separated using denaturing PAGE, producing a faster migrating band for the strand of interest, and a slower migrating band for the 5'-A<sub>20</sub>-hexaethyleneglycol modified complement. The faster migrating band of interest is excised and the DNA from the gel slice is extracted. **(b)** MeRPy-PCR: (i) One primer is covalently bound to a linear polyacrylamide polymer. (ii) The template oligonucleotide anneals to the polymer-bound primer, and the primer is extended by a DNA polymerase. (iii) A dsDNA product is generated. The newly generated, complementary strand is covalently linked to the polymer. (iv) After several

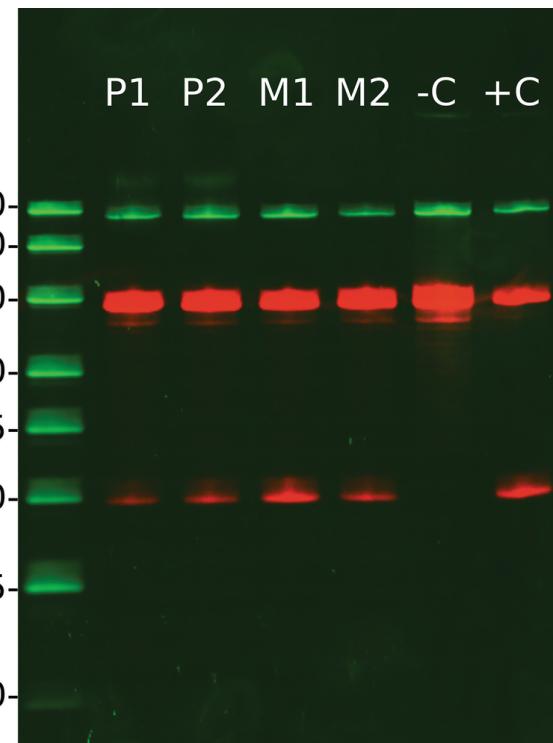
rounds, the original template has been amplified and there is an abundance of dsDNA product, linked by one strand to the polymer. (v) The DNA duplex is denatured by addition of a base, releasing the strand of interest into the solution. The polymer is selectively precipitated, leaving the strand of interest in the supernatant. The ssDNA of interest can now easily be recovered

A recent method developed by the Shih laboratory allows the production of single-stranded DNAs >200 nt (see Fig. 1b) [6]. Methanol-responsive Polymer PCR (MeRPy-PCR) involves the incorporation of one modified PCR primer into a polymer of linear polyacrylamide. A PCR reaction is performed using the polymer-bound primer and a second unmodified primer, the product of which is dsDNA in which the undesired strand is chemically linked to the polyacrylamide polymer. Denaturation of the dsDNA, followed by the precipitation of the polyacrylamide polymer, allows the target strand to be extracted from the supernatant.

We have adapted both MeRPy-PCR and PriTer-PCR methods for the production of active, single-stranded deoxyribozymes from DNA libraries obtained during *in vitro* selection, and present our protocols in this chapter. As an example, we included the preparation of a 10–23 deoxyribozyme using both methods (see Fig. 2). RNA cleavage activity of the deoxyribozyme obtained by MeRPy-PCR and PriTer-PCR is reproducible and comparable to the activity of chemically synthesized ssDNA (see Fig. 3). Generally, the PriTer-PCR method is faster to set up, but has some drawbacks (see Note 1). The MeRPy-PCR method requires more initial preparation, and therefore has a higher barrier to entry, but is otherwise straightforward to use and has fewer drawbacks.



**Fig. 2** PCR products using an example 56 nt long 10–23 template. The first lane contains a molecular weight standard, annotated with numbers representing the length of each band in nucleotides. Lane 2 marked with “M” contains the product of a MeRPy-PCR (without prior clean-up). The third lane marked with “P” contains the PriTer-PCR products. The bands in the red box highlight the products of the expected length of the 10–23 ssDNA construct that should be excised as described in Subheading 3.9. The band in the blue box just above the product in lane “P” is the 5'-A<sub>20</sub>-hexaethyleneglycol modified product. Take care not to excise any of this band when extracting (see Note 21). The band in the green box is the 40 nt long 5'-dA<sub>20</sub>-hexaethyleneglycol modified (-)-strand PriTer primer. The bands in the yellow box are the unmodified 20 nt long (+)-strand primers. All other bands are of unknown origin, and are either a side-product of DNA synthesis, or the result of nonspecific PCR amplification



**Fig. 3** 10–23 DNAzyme RNA cleavage assay. Bands in green are imaged in the SYBR emission channel. Bands in red are imaged in the Cy5 channel. The leftmost lane contains an annotated molecular weight standard, annotated with oligonucleotide length in nt. The green bands at 56 nt represent the 10–23 DNAzyme. The top red bands are uncleaved RNA substrate. The lower red bands are cleaved RNA substrate. The lanes labeled “P1” and “P2” are replicates for a 10–23 DNAzyme generated by two different PriTer-PCRs. The lanes labeled “M1” and “M2” are replicates for a 10–23 DNAzyme generated by two different MeRPy-PCRs. The lanes labeled “-C” and “+C” are negative and positive controls, respectively. Positive control contains RNA substrate, chemically synthesized 10–23 DNAzyme ordered from a commercial supplier, and reaction buffer. Negative control contains RNA substrate, chemically synthesized 10–23 DNAzyme, but no reaction buffer. PCRs were performed using 10–23 reverse complement as input templated to ensure all activity comes from newly generated strands (see Note 21)

## 2 Materials

### 2.1 Oligonucleotides

1. All oligonucleotides were purchased. The supplier (e.g., Integrated DNA Technologies) must be able to provide 5'-acrydite modification for the MeRPy (-)-strand primer and an internal hexaethylene glycol modification for the PriTer-PCR (-)-strand primer. All oligonucleotides were ordered with standard desalting purification, except those with fluorescent tags, which were ordered with HPLC purification. For the example of the 10–23 DNAzyme given in this chapter, the following oligonucleotide sequences were used (see Note 2):

- (a) 10–23 DNAzyme sequence:  
TGTGGTGTGATATTGCCGGCAGGCTAGCTACAAACGACGGATGGTAGTATGGTGGC
- (b) 10–23 DNAzyme template sequence:  
GCCACCATACTACCATCCCGTCGTTAGCTAGCCTGCCGAATATCACACCAACA
- (c) 5'-Cy5 fluorophore-tagged 10–23 RNA substrate sequence:  
rCrArUrArCrUrArCrCrArUrCrCrGrArUrGrCrCrGrArArUArUrCrArCrArC.
- (d) Primer sequences: 5'-A<sub>20</sub>-hexaethyleneglycol-GCCACCATACTACCATCCCG, 5'-acrydite-GCCACCATACTACCATCCCG, TGTGGTGTGATATTGCCGGC.
2. Integrated DNA Technologies 10/60 Ladder was used as size standard on all gels.

## 2.2 Polyacrylamide Gel Electrophoresis Stocks and Buffers

1. Tris-Borate-Ethylenediaminetetraacetic acid (EDTA) (TBE) Running buffer (10× stock): 0.89 M Tris base, 0.89 M boric acid, 20 mM EDTA, pH 8. Prepare 108 g Tris base, 55 g boric acid and 40 mL 500 mM EDTA pH 8. Dissolve in a 1 L beaker with a magnetic stirrer in 500 mL ultra-pure water, such as that obtained from a Merck Millipore Milli-Q water filtration system. Adjust pH to 8 using HCl. Transfer solution to 1 L graduated cylinder and fill to 1 L with pure water.
2. Tris-Borate-Urea (TBU) acrylamide gel stock: 8 M urea, 20% (w/v) acrylamide/bis-acrylamide 19:1, 1× TBE running buffer. Weigh 240.24 g of urea into 1 L glass beaker. Add 250 mL 40% (w/v) acrylamide/bis-acrylamide 19:1 to beaker, along with 50 mL of 10× TBE. Dissolve by mixing (*see Note 3*).
3. 10% (w/v) ammonium persulfate (APS), make fresh or store at –20 °C.
4. *N,N,N,N'*-Tetramethyl-ethylenediamine (TEMED), store at 4 °C.

## 2.3 Nucleic Acid PAGE Loading Buffer

1. Dissolve 5 mg of bromophenol blue in 9.99 mL formamide. Add 10 µL 500 mM EDTA pH 8.

## 2.4 Additional Reagents and Materials

1. 70% ethanol.
2. Methanol, store at –20 °C.
3. Isopropanol.
4. 5 mg/mL glycogen.
5. Phusion® Hot Start Flex 2× Master Mix.

6. 40% (w/v) acrylamide (*see Note 4*).
7. 5% (w/v) APS: 25 mg APS in 0.5 mL pure water.
8. 5% (w/v) TEMED: 6.5  $\mu$ L TEMED in 93.5  $\mu$ L pure water.
9. 3 M Sodium acetate pH 5.
10. SYBR-Gold or equivalent DNA stain.
11. 0.22  $\mu$ m Cellulose Acetate SpinX columns.
12. MeRPy polymerization (TBE) buffer (5 $\times$  stock): 500 mM Tris base, 500 mM boric acid, 10 mM EDTA pH 8.2. Prepare by mixing 3.028 g of Tris base, 1.546 g boric acid, and 1 mL of 500 mM EDTA pH 8. Add enough pure water to fully dissolve, adjust pH to 8.2, and fill to final volume of 50 mL.
13. MeRPy purification (TE) buffer (10 $\times$  stock): 50 mM Tris, 1 mM EDTA pH 8. Prepare by mixing 0.303 g of Tris base and 100  $\mu$ L of 500 mM EDTA pH 8. Add enough pure water to fully dissolve, adjust pH to 8, and fill to final volume of 50 mL.
14. Basic denaturing buffer: 200 mM NaOH, 2 mM EDTA. Prepare by mixing 10 mL of 1 M NaOH and 200  $\mu$ L of 500 mM EDTA pH 8, then fill with pure water to final volume of 50 mL.

## 2.5 Equipment

1. Tube rotator for 1.5 mL and 2 mL reaction tubes.
2. Refrigerated centrifuge capable of achieving 20,000  $\times$  *g* and temperatures at least as low as 3 °C.
3. Gel imaging system.
4. Slab gel casting kits: 20  $\times$  20 cm Maxi-gel, 2 mm spacer.
5. PCR thermocycler.
6. Nanodrop or similar UV-Vis spectrometer for measuring nucleic acid concentrations.
7. Tabletop mini-centrifuge.
8. Vortexer.
9. Blue-light gel transilluminator.

---

## 3 Methods

### 3.1 Primer Design

Design primers to the following specifications:

1. 18–20 nt in length.
2. Minimal secondary structure.
3. Primers should neither be self-complementary, nor mutually complementary.
4. Minimal melting temperature difference between (+)-strand and (−)-strand primer (up to 2 °C is acceptable).
5. Two consecutive strong bases (G or C) at the 3' end of the oligonucleotide.
6. For MeRPy-PCR, the (−)-strand primer must contain a 5'-acrydite modification for incorporation into linear polyacrylamide polymer.
7. For PriTer-PCR, the (−)-strand primer must begin with a 5'-A<sub>20</sub> (other lengths are permissible), followed by an internal hexaethyleneglycol modification, and end with the desired primer sequence.

### **3.2 Polymerization of Primer-Free Linear Polyacrylamide for MeRPy-PCR**

For comparison to polymerization of (−)-strand MeRPy primer (*see Note 5*).

1. Prepare a 2 mL glass vial with septum cap.
2. Add the following reagents in order: 100 µL 5× MeRPy polymerization buffer, 335.25 µL pure water, 62.5 µL 40% (w/v) acrylamide, 1.25 µL 20% (w/v) sodium acrylate.
3. Close cap and mix by vortexing for 10 s (*see Note 6*).
4. Pierce septum cap by inserting a needle which is attached to a nitrogen gas source.
5. Loosen the cap half-way, so that air can pass without obstruction, but without the cap falling off.
6. Make sure that the needle is submerged in the solution and the tip is near the bottom of the vial.
7. Gently open the gas valve until bubbles can be seen coming out of the needle at a rate of around 5 bubbles per second.
8. Purge solution for 20 min.
9. After purging, unscrew septum cap from vial completely while continuing to purge the solution (take care not to remove needle from solution), add 0.5 µL of 5% (w/v) TEMED and swirl the vial to mix the solution. Add 0.5 µL of 5% (w/v) APS, swirl once more, then screw the cap back on half-way.
10. Continue purging for 15 min.
11. Tighten cap completely but leave the nitrogen pressure on. This will result in the bubbling

slowly stopping.

12.

Incubate tightly closed vial with nitrogen pressure overnight at room temperature.

13.

The solution should be highly viscous in the morning, which indicates that the reaction was successful. The needle can now be safely removed from the vial. Proceed to purification immediately (*see* Subheading 3.5), or store at -20 °C if you wish to purify it at a later time.

### 3.3 Polymerization of (-)-Strand MeRPy Primer (*See Note 5*)

1. Prepare a 2 mL glass vial with septum cap.
2. Add the following reagents in order: 100 µL 5× MeRPy polymerization buffer, 285.25 µL pure water, 50 µL 1 mM acrydite-tagged (-)-strand primer, 62.5 µL 40% (w/v) acrylamide, 1.25 µL 20% (w/v) sodium acrylate.
3. Close cap and mix by vortexing for 10 s (*see Note 6*).
4. Pierce septum cap by inserting a needle which is attached to a nitrogen gas source.
5. Loosen the cap half-way, so that air can pass without obstruction, but without the cap falling off.
6. Make sure needle is submerged in the solution and the tip is near the bottom of the vial.
7. Gently open the gas valve until bubbles can be seen coming out of the needle at a rate of around five bubbles per second.
8. Purge solution for 20 min.
9. After purging, unscrew septum cap from vial completely while continuing to purge the solution (take care not to remove needle from solution), add 0.5 µL of 5% (w/v) TEMED, and swirl the vial to mix the solution. Add 0.5 µL of 5% (w/v) APS, swirl once more, then screw the cap back on half-way.
10. Continue purging for 15 min.
11. Tighten cap completely but leave the nitrogen pressure on. This will result in the bubbling slowly stopping.
12. Incubate tightly closed vial with nitrogen pressure overnight at room temperature.
13. The solution should be highly viscous in the morning. This indicates the reaction was successful. The vial can now be safely removed from the needle. Proceed to purification protocol (*see* Subheading 3.4) or store at -20 °C if you wish to purify it at a later time.

### 3.4 Purification Protocol for MeRPy Primer Polymer

1. Add 350 µL of 1× MeRPy purification buffer to the vial containing the polymerized MeRPy primer.

2. Vortex for 2 min.
3. Transfer contents of the vial to a 50 mL centrifuge tube (*see Note 7*).
4. Add 0.5 mL of 1× MeRPy purification buffer to the empty vial and vortex for 30 s to rinse out the remaining solution.
5. Transfer the contents of the vial into the 50 mL tube (*see Note 7*).
6. Repeat **steps 4** and **5** three times, to ensure that all of the polymer is washed out and transferred into the tube.
7. Add an additional 2.15 mL of 1× MeRPy purification buffer. The total volume of added 1× MeRPy purification buffer should now be 4.5 mL.
8. Vortex 50 mL tube containing the polymer solution at high speed for 10 min (*see Note 8*).
9. Add 25 µL of 5 M NaCl and vortex for 10 s.
10. Fill a disposable syringe with 5 mL of cold methanol (stored at -20 °C, *see Note 9*). Vortex the 50 mL Falcon tube containing the polymer solution with cap removed at low speed to avoid spillage. While the tube is vortexing, slowly drip in 5 mL of cold methanol. As the last few mL of methanol are added, the solution should turn turbid due to the precipitated polymer.
11. Transfer the tube to ice and incubate for 2 min.
12. Centrifuge for 5 min at 150 × g and 4 °C.
13. A visible pellet should form after centrifugation. Decant supernatant.
14. Add 4.5 mL of pure water, and vortex for 10 min. Make sure the pellet is completely dissolved.
15. Add 0.5 mL of 10× MeRPy purification buffer.
16. Add 30 µL of 5 M NaCl.
17. Repeat **steps 10–13** to reprecipitate the polymer and obtain a pellet.
18. Add 4.5 mL pure water and vortex for 10 min. Make sure the pellet is completely dissolved.
19. Add 0.5 mL 10× MeRPy purification buffer.
20. Vortex for 10 s.
21. Store at -20 °C.

### 3.5 Purification Protocol for Linear Polyacrylamide Polymer

1. Add 350  $\mu$ L of 1 $\times$  MeRPy purification buffer to the vial containing the linear polyacrylamide polymer.
2. Vortex for 2 min.
3. Transfer contents of vial to a 50 mL centrifuge tube (*see Note 7*).
4. Add 0.5 mL of 1 $\times$  MeRPy purification buffer to the empty vial and vortex for 30 s.
5. Transfer the content of the vial into the centrifuge tube (*see Note 7*).
6. Repeat **steps 4** and **5** three times, to ensure that all of the polymer is washed out and transferred into the tube.
7. Add an additional 2.15 mL of 1 $\times$  MeRPy purification buffer. The total volume of added 1 $\times$  MeRPy purification buffer should now be 4.5 mL.
8. Vortex 50 mL tube containing the polymer solution at high speed for 10 min.
9. Add 25  $\mu$ L of 5 M NaCl and vortex for 10 s.
10. Fill a disposable syringe with 5 mL of methanol (stored at -20 °C) (*see Note 9*). Vortex the 50 mL tube containing polymer solution with cap removed at low speed to avoid spillage. While the tube is vortexing slowly drip in the 5 mL of cold methanol. As the last few mL of methanol are added, the solution should turn turbid from the precipitated polymer.
11. Transfer the tube to ice and incubate for 2 min.
12. Centrifuge for 5 min at 150  $\times$  g and 4 °C.
13. A visible pellet should form after centrifugation. Decant supernatant.
14. Add 4.5 mL of pure water, and vortex for 10 min. Make sure the pellet is completely dissolved.
15. Add 0.5 mL of 10 $\times$  MeRPy purification buffer.
16. Add 30  $\mu$ L of 5 M NaCl.
17. Repeat **steps 10–13** to precipitate the polymer and obtain a pellet.
18. Add 2.5 mL pure water and vortex for 10 min. Make sure the pellet is completely dissolved. Final concentration is 1% (w/v) of linear polyacrylamide polymer.
19. Store at -20 °C.

### 3.6 Preparation of Preparative Denaturing PAGE Nucleic Acid Gels

1. Prepare the gel plates for casting by placing 2 mm thick spacers between the front and back plate, and clamping them tightly on the bottom and sides to prevent leakage.

2. Pipette 45 mL TBU acrylamide gel stock into a 50 mL centrifuge tube.
3. Add 45  $\mu$ L of TEMED. Screw on cap and swirl around to mix. Take care not to create bubbles. Reopen cap.
4. Add 450  $\mu$ L of 10% APS. Quickly screw on cap and swirl to mix. Take care not to create bubbles.
5. Quickly pour the contents of the tube into the sealed gel plates. You can do this with a 50 mL serological pipette, or directly pour from the tube. Take care to avoid air bubbles between the plates as the solution polymerizes (*see Note 10*).
6. Quickly insert a 2 mm thick, wide-welled comb in order to create large wells suitable for large volumes used in extractions. The well width used in our example is 15 mm, producing a well with the approximate dimensions: 10 mm  $\times$  15 mm  $\times$  2 mm.
7. Allow to polymerize for at least 2 h, but ideally overnight (*see Note 11*).

### 3.7 MeRPy-PCR and PriTer-PCR Protocol

1. In a 200  $\mu$ L PCR tube, mix: 5  $\mu$ L 10  $\mu$ M (+)-strand primer, 4  $\mu$ L 10 nM template oligo, 50  $\mu$ L of Phusion<sup>®</sup> Hot Start Flex 2 $\times$  Master Mix (*see Notes 12 and 13*).
2. For MeRPy-PCR, add: 5  $\mu$ L of 10  $\mu$ M polymerized (-)-strand primer obtained from Subheading 3.4 (*see Note 14*).
3. For PriTer-PCR, add: 5  $\mu$ L 10  $\mu$ M (-)-strand-A<sub>20</sub>-hexaethyleneglycol primer.
4. Fill to final volume of 100  $\mu$ L with nuclease-free water. Mix well by pipetting slowly. Centrifuge briefly on tabletop mini-centrifuge to get rid of any bubbles.
5. Place the tubes into a thermocycler and run the following program:

Phase I (1 cycle):

(a) 98 °C for 30 s (initial denaturation).

Phase II (Variable number of cycles. *See Note 15*):

(a) 98 °C for 10 s (denaturing).

(b) 64 °C for 30 s (annealing) (*see Note 16*).

(c) 72 °C for 15 s (extension) (*see Note 17*).

Phase III (1 cycle):

(a) 72 °C for 1 min (final extension).

(b) Hold at 4 °C.

6. Proceed to Subheading 3.8 for ssDNA extraction from MeRPy-PCR, or Subheading 3.9 for ssDNA extraction from PriTer-PCR. An analytical gel image of the unpurified MeRPy-PCR products of the 10–23 DNAzyme example is shown in Fig. 2.

### 3.8 ssDNA Extraction from MeRPy-PCR

1. Transfer the MeRPy-PCR product into a 1.5 mL reaction tube.
2. Add 1/3 volume, in this case 33.33  $\mu$ L, of 1% (w/v) linear polyacrylamide to the MeRPy-PCR product.
3. Mix well by vortexing for 20 s.
4. Add 1 volume, in this case 133.33  $\mu$ L, of cold methanol to the reaction tube and vortex for 20 s. This should precipitate the polymer along with the bound dsDNA PCR product.
5. Incubate for a minimum of 1 min on ice.
6. Centrifuge at 2000  $\times g$  and 3 °C for 5 min.
7. Carefully remove the supernatant by pipetting.
8. Resuspend pellet in 100  $\mu$ L of nuclease-free water by vortexing for 10 min, or until the pellet is completely dissolved.
9. Add 0.22 volumes, or in this case 22  $\mu$ L, of basic denaturing buffer and vortex for 20 s.
10. Incubate at room temperature for at least 1 min to denature the dsDNA product.
11. Add 1 volumes, or in this case 122  $\mu$ L, of ice-cold methanol and vortex for 20 s to precipitate polymer.
12. Incubate on ice for 1 min.
13. Centrifuge at 2000  $\times g$  and 3 °C for 5 min.
14. Transfer supernatant to fresh 1.5 mL reaction tube.
15. Centrifuge supernatant at 20,000  $\times g$  and 3 °C for 10 min.
16. Carefully transfer supernatant into fresh 1.5 mL reaction tube to ensure no carry-over of polymer.
17. Perform an isopropanol precipitation on the supernatant by adding 0.1 volume, or in this case 24.4  $\mu$ L, 3 M sodium acetate (pH 5), followed by roughly 1.5 volumes, or in this case 400  $\mu$ L ice-cold isopropanol, and 0.2–0.3  $\mu$ L glycogen (5 mg/mL).
18. Centrifuge sample for 1 h at 20,000  $\times g$  and 3 °C to precipitate DNA (see Note 18).
19. Decant supernatant or remove by pipetting, taking care not to disturb the pellet.
20. Carefully add 1 mL 70% ice-cold ethanol.
21. Centrifuge at 20,000  $\times g$  for 10 min at 3 °C.

22. Optional: Repeat **steps 19–21** if strict desalting is required.
23.
  - Remove supernatant by pipetting, ensuring that nearly no ethanol remains in the tube.
24. Place open tube in benchtop heating block at 37 °C until the pellet is completely dry.
25. Resuspend pellet in nuclease-free water or your buffer of choice. Recommended volumes range from 10 to 20 µL.
26. Quantify yield by determining sample absorbance at 260 nm and calculate concentration using a sequence-specific extinction coefficient.
27. Store at –20 °C or –80 °C.

### 3.9 PriTer-PCR ssDNA Extraction

1. In a 1.5 mL or 0.5 mL tube, mix 100 µL of the PriTer-PCR reaction mixture from Subheading 3.7 with 100 µL of denaturing gel loading buffer (*see Note 19*). Mix thoroughly by pipetting or vortexing for 10 s. Incubate PCR product in a heat block at 95 °C for at least 5 min before loading on extraction gel.
2. Prepare an oligonucleotide length standard by mixing it with loading buffer in a comparable ratio to the PCR product. Mix 1–5 µL of the ladder stock with an equal volume of denaturing gel loading buffer, heat to 95 °C for 3 min, then cool on ice. Load approximately 0.3–0.5 µg of ladder per 15 mm wide lane.
3. Pre-run extraction gel from Subheading 3.6 at 25 W for 10 min. Switch off the power supply and flush the wells with running buffer at least twice using a needle and syringe to remove accumulated urea from wells.
4. Carefully load your sample and your size standard. In this example, we load 100–200 µL of sample in gel loading buffer (1:1 ratio) per well (10 mm × 15 mm × 2 mm). Higher volumes of sample can be divided between multiple wells.
5. Run the gel at 5 W for about 5 min, or until you see that all of the dye has migrated into the gel.
6. When the samples have entered the gel completely, run the gel at the appropriate power level. For the 10–23 DNAzyme used here which is 56 nt in length (Fig. 2), and a 20% gel the proper settings are 30 W and a running time of 50 min, or until the blue dye has reached the last 10% of the gel's height.
7. Remove the gel from the plates and soak for 10 min in 200 mL of running buffer and 2 µL of SYBR Gold stain.
8. Drain and wash the gel, then wrap in plastic wrap. Use a blue-light transilluminator and yellow filter to locate the bands of the (+)-strand, which runs with length of target, and of its complementary (–)-strand, which has a lower electrophoretic mobility due to the 5'-A<sub>20</sub>-hexaethyleneglycol modification. An analytical gel image for the PriTer-PCR preparation of the

model 10–23 DNAAzyme is shown in Fig. 2.

9. Either carefully cut the lower (+)-strand band with a clean scalpel, or mark its location with a permanent marker on the plastic wrap for scalpel excision outside of the transilluminator. Take care not to include any of the slower migrating band in the gel slice.
10. Transfer gel slice into 2 mL round bottomed tube and crush it (see Note 20).
11. Add approximately 1–2 gel slice volumes of 0.3 M sodium acetate (pH 5) to the tube and rotate or shake overnight to elute ssDNA from crushed gel slice.
12. Remove crushed gel by transferring contents of tube into a 0.22  $\mu$ m Cellulose Acetate SpinX column and centrifuging at 16,000  $\times g$  for 5 min.
13. Transfer eluate into fresh 1.5 mL centrifuge tube and add 0.2  $\mu$ L glycogen and 1 mL of cold isopropanol. Mix by vortexing for 10 s.
14. Centrifuge tube for 1–1.5 h at 3 °C at 20,000  $\times g$ . If a pellet did not form, centrifuge another 30 min.
15. Carefully decant or pipette out supernatant. Take care not to disturb the pellet.
16. Wash by carefully adding 1 mL of cold 70% ethanol and centrifuge another 10 min at 3 °C at 20,000  $\times g$ . Then carefully decant once more.
17. Optionally repeat step 16.
18. Dry pellet by leaving open tube in 37 °C benchtop thermal block until the pellet is completely dry.
19. Resuspend in desired volume of nuclease-free water or buffer of choice. Recommended volumes range from 10 to 20  $\mu$ L.
20. Quantify yield by determining sample absorbance at 260 nm and calculate concentration using a sequence-specific extinction coefficient.
21. Store at –20 °C or –80 °C.

---

## 4 Notes

1. PriTer-PCR is not a very reliable method for library recovery during an in vitro selection, as the gel extraction protocol is fertile ground for short non-functional parasitic sequences to overtake the population. For example, in a selection with a library of length  $N$ , if parasites emerge with length  $N-20$ , then the (–)-strand will migrate within the same band as the selected oligos, as it behaves as if it were a 20 nt longer oligonucleotide. This contamination can be avoided by taking particular care during PCR not to produce any nonspecific products. Such parasites can also be removed from a pool by using a longer dA sequence at the 5' end of the (–)-strand primer after the parasites emerge, or simply using a dA sequence so long that it does not allow for parasite emergence. MeRPy does not suffer from these issues, and while a more difficult method to get setup and running, is a better choice for sensitive selections.

2. For the purposes of demonstrating the effectiveness of these protocols and showing that the DNAzymes produced by the methods demonstrated in this chapter are functional, we used, in addition to the oligonucleotides mentioned in Subheading 2.1, item 1, the following materials and methods, in brief:
  - (a) 10–23 DNAzyme reaction buffer: 10 mM MgCl<sub>2</sub>, 1 M NaCl, 50 mM Tris–HCl pH 7.5.
  - (b) Cleavage assays were performed by incubating 1 pmol of DNAzyme with 10 pmol of RNA substrate in 1× reaction buffer, at 37 °C for 2 h. Reactions were quenched by addition of loading buffer from Subheading 2.3.
  - (c) An analytical 20% urea PAGE (same as in Subheading 3.6, except 1 mm thick) was used to visualize RNA cleavage (Fig. 3).
3. To dissolve urea, it is necessary to heat as well as stir. Acrylamide is a dangerous chemical, so take the proper precautions when handling it. Take care to place beaker with magnetic stir bar on heated magnetic stirrer in an active fume hood. Make sure to cover the beaker with aluminum foil for safety.
4. The acrylamide used here must be completely free of bis-acrylamide.
5. Procedure to polymerize MeRPy primer and primer-free linear polyacrylamide are identical except for the addition of acrydite-tagged primer and a reduced amount of pure water in comparison to primer-free linear polyacrylamide, and can be done in parallel, and use the same 5% (w/v) APS and TEMED stock solutions.
6. Unless otherwise specified always vortex on highest speed.
7. Due to the high viscosity of the solution this can be difficult, so using a pipette is discouraged. Instead, fill a 20 mL disposable syringe with air and attach a needle that has been bent by ~90°. (This can be achieved by using the needle's plastic cap.) Turn the vial upside-down above the open 50 mL reaction tube and insert the needle of the syringe into the solution. Depress the syringe plunger to inject air and expel the solution from the vial.
8. Vortexing for longer periods is made much easier with use of a vertical tube holder.
9. It is safest to work in a fume hood when handling methanol.
10. Some leftover gel mix will remain in the tube. Screw cap back onto tube. Content of the tube is a good indicator on the polymerization state of the gel.
11. Placing gel into a 60 °C incubator will increase the rate of polymerization. Additionally, the rise in gel temperature makes the pre-running of the gel unnecessary.
12. Other polymerases may be used. In this case the protocol may need to be adapted to the manufacturer's specifications.
13. Taq polymerase may be a poor choice for DNAzyme production, partially because of its comparatively high error rate, but primarily due to its tendency to add terminal thymine overhangs.

~~comparatively high error rate, but primarily due to its tendency to add terminal mismatch overhangs~~ to its products, which may interfere with certain DNAzyme systems that rely on homogenous 3'-ends.

14. Determining the final concentration of the polymerized MeRPy primer is difficult. Here, we assume a 100% efficient reaction, although real yields may be lower and MeRPy primer input may need to be increased slightly.
15. A quantitative PCR reaction might be helpful to determine the number of cycles that are necessary for your particular amplification. This is especially the case if you are using this protocol to recover a selection library. In the case of a selection experiment, it is essential not to reach the plateau phase of PCR to prevent sequence reshuffling, so the number of Phase II cycles must be adjusted appropriately. In the case of the example protocol, with 0.5 µM primer concentrations and 4 fmol of template input, 8 cycles are typically sufficient. If you are not performing a selection this is not a concern, and you can cycle between 30 and 35 times.
16. Make sure to use the appropriate annealing temperature for your set of primers. During an in vitro selection it is generally better to use higher annealing temperatures, sometimes even above the  $T_m$  of the primers, to reduce nonspecific amplification.
17. Extension time can be shorter. Even as short as 5 s may be enough for most DNAzymes, which are shorter than 0.1 kb. During a selection it is preferable not to use a longer extension time than needed to prevent nonspecific amplification.
18. Shorter centrifugation times may be used to save time if producing the highest yield is not a concern.
19. The electrophoretic mobility of the PCR product may be slightly altered compared to size standards due to the presence of salts.
20. The plunger of a 1 mL disposable syringe is a particularly effective and convenient tool for crushing gel slices. Take care not to crush gel too much (into a paste) as it will clog the Cellulose Acetate SpinX column.
21. Analytical denaturing PAGE was prepared as per specifications of Subheading 3.6, except scaled down to 25 mL, and cast into a 1 mm thick gel, with smaller wells.

---

## References

1. Hughes RA, Ellington AD (2017) Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. *Cold Spring Harb Perspect Biol* 9:a023812. <https://doi.org/10.1101/cshperspect.a023812> [Crossref][PubMed][PubMedCentral]
2. Caruthers MH (2011) A brief review of DNA and RNA chemical synthesis. *Biochem Soc Trans* 39(2):575–580 [Crossref]
3. Higuchi RG, Ochman H (1989) Production of single-stranded DNA templates by exonuclease digestion following the polymerase chain reaction. *Nucleic Acids Res* 17:5865. <https://doi.org/10.1093/nar/17.14.5865> [Crossref][PubMed][PubMedCentral]
4. Gyllensten UB, Erlich HA (1988) Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. *Proc Natl Acad Sci U S A* 85:7652–7656. <https://doi.org/10.1073/pnas.85.20.7652>

[Crossref][PubMed][PubMedCentral]

5. Williams KP, Bartel DP (1995) PCR product with strands of unequal length. Nucleic Acids Res 23:4220–4221. <https://doi.org/10.1093/nar/23.20.4220>  
[Crossref][PubMed][PubMedCentral]
6. Minev D, Guerra R, Kishi JY et al (2019) Rapid in vitro production of single-stranded DNA. Nucleic Acids Res 47(22):11956–11962. <https://doi.org/10.1093/nar/gkz998>  
[Crossref][PubMed][PubMedCentral]

## **Part II**

# **Design: Kinetics, Performance, and Turnover**

## 4. Design of a DNAzyme

### Prediction of mRNA Regions Accessible to a DNAzyme

Gerhard Steger<sup>1</sup>✉ and Julian Victor<sup>1</sup>✉

(1) Institut für Physikalische Biologie, Heinrich-Heine-Universität  
Düsseldorf, Düsseldorf, Germany

✉ **Gerhard Steger (Corresponding author)**  
Email: [steger@biophys.uni-duesseldorf.de](mailto:steger@biophys.uni-duesseldorf.de)

✉ **Julian Victor**  
Email: [victor@hhu.de](mailto:victor@hhu.de)

#### Abstract

The efficiency of RNA-cleaving DNAzymes depends on a large extent on complex formation with their RNA targets. We describe available prediction tools that should help in the design of efficient DNAzymes and show some experimental methods to test the predictions. The main example is for a 10–23 DNAzyme, but the procedure works as well for the 8–17 DNAzyme family.

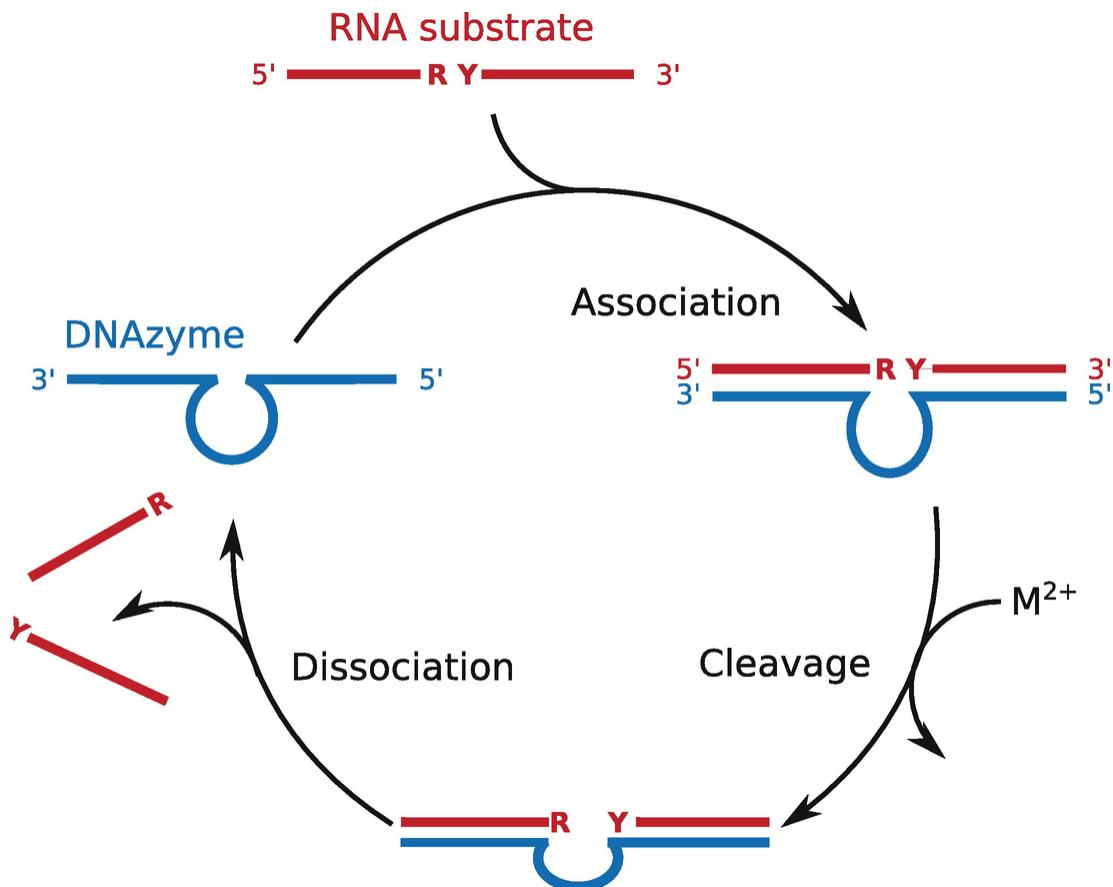
**Key words** RNAfold – RNAPlfold – poland – DINAMElt

---

## 1 Introduction

The catalytic RNA-cleavage activity of a DNAzyme from the 10–23 family is shown in the scheme of Fig. 1. Briefly, the DNAzyme associates with the target RNA to form a complex. Next, the DNAzyme:RNA complex binds divalent cation(s) to form the catalytically active enzyme:substrate complex

that is able to cleave the RNA. Note that the catalytic activity of some DNAzymes does not depend on divalent cations [1, 2]. Finally, the enzyme:product complex of DNAzyme and cleaved RNA dissociates into its components, and the free DNAzyme is able to re-enter the catalytic cycle.



**Fig. 1** Reaction scheme for DNAzyme-mediated RNA cleavage. A DNAzyme, for example, the 10–23 DNAzyme, associates with the RNA target by pairing of its recognition arms, the catalytic core of the DNAzyme cleaves the target in the presence of divalent metal ion  $M^{2+}$  between a purine R and a pyrimidine Y, and finally the ternary complex dissociates

There are further requirements for a DNAzyme to exhibit high substrate sequence specificity and selectivity, high catalytic efficiency, and rapid catalytic turnover [3]. Such requirements, not directly obvious from the scheme of Fig. 1, are at least the following:

1. Most DNAzymes require a certain dinucleotide at the cleavage site of the target RNA; for example, a purine ( $R = \{A, G\}$ ) and a pyrimidine

(Y = {C, U}) are required by the 10–23 DNAzyme (Fig. 1). These nucleotides are on the one hand a part of the catalytic center and on the other hand the point of cleavage, which is either transesterification or hydrolysis on the central phosphate [4, 5].

The sequence constraint by an RY dinucleotide on a possible cleavage position is only minor: if the target RNA would have a random base composition ( $p_{N \in \{A,C,G,U\}} = 0.25$ ), an RY dinucleotide is found in every four nucleotides, and, for example, a GU is found in every 16 nucleotides.

2. The binding arms of the DNAzyme have to bind to the target RNA. For this, at least three consequences have to be taken into account:
  - (a) With increasing length of the binding arms, the thermodynamic binding strength increases and favors RNA:DNAzyme complex formation.
  - (b) A stable complex of uncleaved RNA and DNAzyme results also in a stable complex of cleaved RNA and DNAzyme that disfavors dissociation of cleaved RNA and DNAzyme. This results in lower turnover and lower catalytic efficiency of the DNAzyme.
  - (c) With increasing length of the binding arms, the chance for unwanted intra- and/or intermolecular basepairings by the DNAzyme is increasing; that is, the binding arms and/or the DNAzyme core may basepair to each other or to the DNAzyme's core loop. Such pairings apparently lower the availability of binding arms for interaction with the target. An example for such an unwanted interaction is the first crystal structure of a DNAzyme:RNA complex where two such complexes formed two four-way junctions instead of a catalysis-relevant structure [6].
3. The target region of the (m)RNA should have a low degree of structure. The lower the intramolecular basepairing probability of the target region, the higher is the probability for the binding arms of the DNAzyme to access the target and form the DNAzyme:RNA complex. In case of an mRNA as target, the target's most stable secondary structure, which is a thermodynamic feature, might not be of relevance

but rather a kinetically fast folding structure [7, 8]. For this, there are at least two reasons: firstly, many mRNAs are short-lived [9] and, secondly, a ribosome will denature any mRNA structure during translation; each of these features hinders mRNA to reach their global structures at thermodynamic equilibrium but favors sequentially folded mRNAs with only locally stable structures.

In the next sections, we will show selected computational tools that are able to predict the abovementioned features and thus should help a user to design a DNAzyme with optimal features.

## 2 Materials

The computing effort and memory usage for most programs predicting RNA secondary structure like **RNAfold** (see Table 1) scale with  $O(n^3)$  and  $O(n^3)$ , respectively, for a sequence of lengths  $n$ . **RNAplfold** scales with  $O(n \cdot L^2)$  time and  $O(n + L^2)$  memory for any fixed window size  $L$  (see Methods, step 4). The **poland** program scales with  $O(n^3)$  time and  $O(n)$  memory. That is, a standard laptop is sufficient for calculations of sequences with the typical length of mature mRNAs but be aware that doubling the sequence length will result in a factor of 8 in computing time by **RNAfold**.

**Table 1** Programs and web sites. The mentioned site for a program links either to a site for download or web usage

Program name, Service	Site	Ref.
<b>Sequence conversion</b>		
readseq	<a href="https://sourceforge.net/projects/readseq/">https://sourceforge.net/projects/readseq/</a>	[10]
	<a href="https://packages.ubuntu.com/eoan/readseq">https://packages.ubuntu.com/eoan/readseq<sup>a</sup></a>	
SeqRet	<a href="https://www.ebi.ac.uk/Tools/sfc/emboss_seqret/">https://www.ebi.ac.uk/Tools/sfc/emboss_seqret/</a>	[11]
<b>Nucleic acid secondary structure prediction</b>		
Vienna package	<a href="http://rna.tbi.univie.ac.at/">http://rna.tbi.univie.ac.at/</a>	[12]

Program name, Service	Site	Ref.
RNAfold	<a href="http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi">http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi</a>	[13]
RNAPlfold	<a href="https://www.tbi.univie.ac.at/RNA/">https://www.tbi.univie.ac.at/RNA/</a>	[14]
RNAcofold	<a href="http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAcofold.cgi">http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAcofold.cgi</a>	[15]
DINAMelt Web Server		
Mfold	<a href="http://unafold.rna.albany.edu/?q=mfold/DNA-Folding-Form">http://unafold.rna.albany.edu/?q=mfold/DNA-Folding-Form</a>	[16]
DINAMelt	<a href="http://unafold.rna.albany.edu/?q=DINAMelt/Homodimer-simulations">http://unafold.rna.albany.edu/?q=DINAMelt/Homodimer-simulations</a>	[17, 18]
RNAstructure package	<a href="https://rna.urmc.rochester.edu/RNAstructure.html">https://rna.urmc.rochester.edu/RNAstructure.html</a>	[19]
<b>Double strand with mismatches</b>		
poland	<a href="http://www.biophys.uni-duesseldorf.de/html/local/POLAND/poland.html">http://www.biophys.uni-duesseldorf.de/html/local/POLAND/poland.html</a>	[20, 21]
<b>Graphics</b>		
gle	<a href="http://glx.sourceforge.net/">http://glx.sourceforge.net/</a>	
R2R	<a href="https://sourceforge.net/projects/weinberg-r2r/">https://sourceforge.net/projects/weinberg-r2r/</a>	[22]
prob.pl	<a href="https://github.com/StegerG/DesignDz">https://github.com/StegerG/DesignDz</a>	
perl	<a href="https://www.perl.org/get.html">https://www.perl.org/get.html</a>	
<b>Databases for sequence retrieval</b>		
GenBank	<a href="https://www.ncbi.nlm.nih.gov/nucleotide/">https://www.ncbi.nlm.nih.gov/nucleotide/</a>	[23]
European nucleotide archive (ENA)	<a href="https://www.ebi.ac.uk/ena/browser/home">https://www.ebi.ac.uk/ena/browser/home</a>	[24]
DNA databank of Japan (DDBJ)	<a href="http://getentry.ddbj.nig.ac.jp/">http://getentry.ddbj.nig.ac.jp/</a>	[25]
<b>Description of sequence formats</b>		
Vienna	<a href="https://www.tbi.univie.ac.at/RNA/tutorial/">https://www.tbi.univie.ac.at/RNA/tutorial/</a>	
FASTA	<a href="https://en.wikipedia.org/wiki/FASTA_format">https://en.wikipedia.org/wiki/FASTA_format</a>	

<sup>a</sup>The Debian/Ubuntu version allows for direct conversion into Vienna format

In the following, we focus on the usage of programs from the Vienna RNA package and DINAMelt Web Server; similar programs are available from the RNAstructure package and web server (*see* Table 1). All mentioned programs run on Linux, macOS, and Windows, and are available as web services.

A text editor is necessary for modification of sequence files.

The `perl` program, which is described in Note 1, wraps all programs mentioned in Methods, steps 2, 3, 4, and 5, and produces a tabular output and a `gle` script to graphically show the output (for example, *see* Fig. 4). This graphics output, however, might be produced by many graphics programs that are able to read comma-separated value (CSV) files.

---

### 3 Methods

In the following, we describe the use of certain programs that allow a user to get the desired information as asked for in the Introduction:

1. Get your mRNA sequence; for example, download the sequence from GenBank (*see* Table 1).  
In the following, the GenBank reference sequence NM\_000311.3 for the *Homo sapiens* prion protein (PrP) is used as an example.
2. Convert the sequence into Vienna or FASTA format (Table 1).  
The first line of a Vienna-formatted sequence file starts with “>” followed by a header, for example, the sequence ID, followed by a single line containing the sequence (*see* Fig. 2, lines 3–4).

```

1 $ readseq -f20 -oPrP.vie NM_000311.gb
2 $ more PrP.vie
3 > NM_000311
4 attaaagatgattttacagtcaatgagccacgtcaggagcgatggcacccgcaggcgg...
5 $ RNAfold --noLP -p < NM_000311_3.vie > PrP_rnafold.log
6 $ more PrP_rnafold.log
7 > NM_000311
8 auuaaagaugauuuuuacagucaaugagccacguucaggagcgauggcacc ...
9 .(((((((((((((((((...((((..((....)) ... (-877.40)
10 .((((((((((((((...,((((..(((..((....)) ... [-923.09]
11 .((((((((((.((((...((...((..((....)) ... {-749.46 d=392.26}
12 frequency of mfe structure in ensemble 6.36852e-33;
13 ensemble diversity 607.34
14 $ RNAPlfold --noLP -W 80 -L 40 -u 10 < NM_000311_3.vie

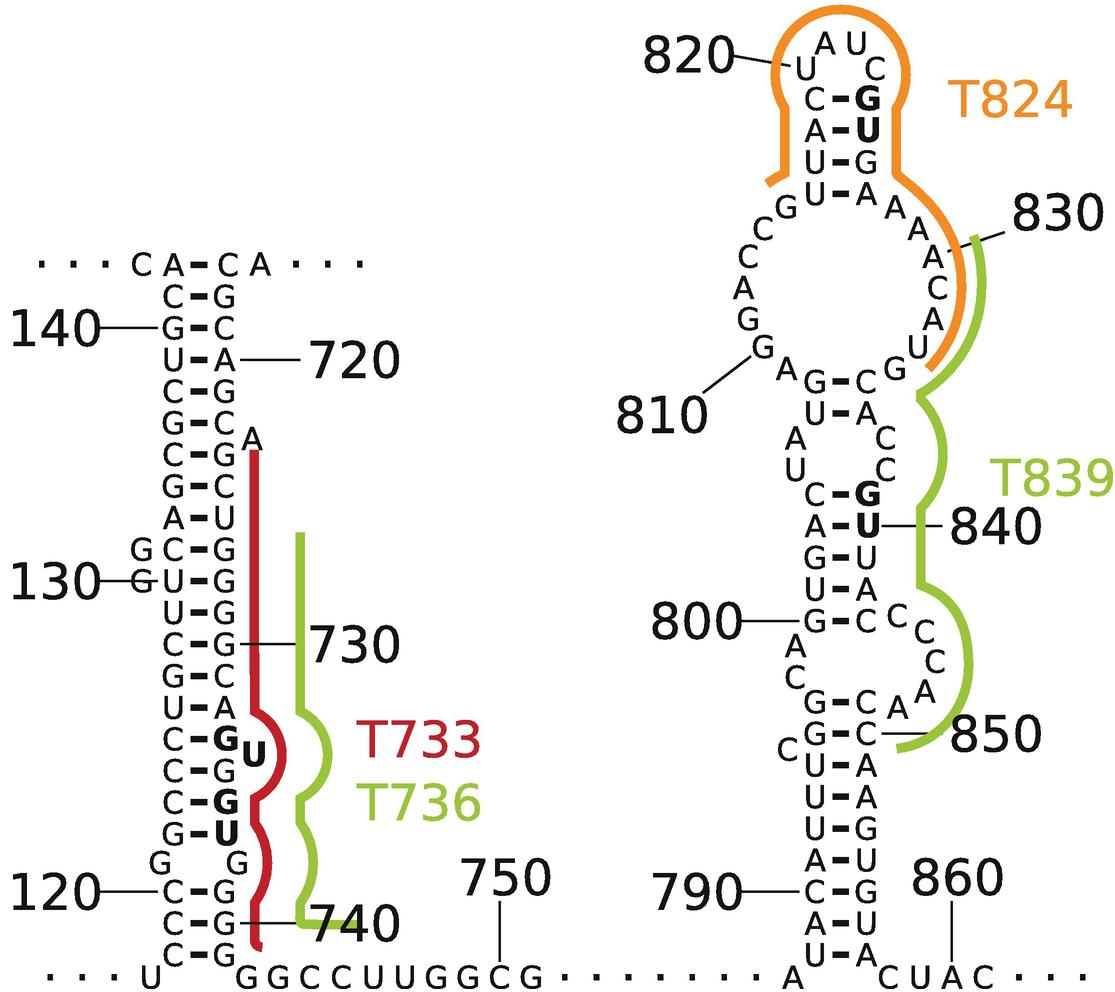
```

**Fig. 2** Prediction of a secondary structure with minimal free energy and basepairing probability of partition folding using **RNAfold**. Line 1: Convert the sequence file **NM\_000311.gb** from GenBank into Vienna format (**-f20**) and name the resulting file **PrP.vie** (**-oPrP.vie**). Line 3–4: A Vienna-formatted sequence file contains a sequence header line starting with “>” followed by a single line containing the sequence. Line 5: **RNAfold** is used to predict the thermodynamically optimal secondary structure, which is the structure with a minimum of free energy (mfe), and the probability of each canonical basepair to be part of any possible structure as a matrix plot by partition folding reading the sequence from file **PrP.vie** and writing its output to the file **PrP\_rnafold.log**. The given options **-p** force partition folding, and **--noLP** disallows lonely basepairs that have no neighboring and stacking pairs. Line 6: Show the content of **RNAfold**’s output file. Line 7: Header of sequence file **PrP.vie**. Line 8: Sequence; note that **RNAfold** automatically converts T into U. Line 9: Mfe structure in bracket notation followed by its free energy ( $\Delta G^0$ ) in kcal/mol. This notation uses the letters “.”, “(”, and “)” to mark bases as unpaired, and up- or downstream paired, respectively. Line 10: Coarse representation of the pair probabilities in pseudo-bracket notation followed by the ensemble free energy in kcal/mol. This notation uses the letters “.”, “,”, “|”, “{”, “}”, “(”, and “)” to mark bases as essentially unpaired, weakly paired, strongly paired without 5’ or 3’ preference, weakly upstream or downstream paired, or strongly up- or downstream paired, respectively. The ensemble free energy is lower (more negative) than that of the mfe structure with the increasing number of structures in the partition function  $Q$  ( $\Delta G^0$  (ensemble) =  $-RT \ln Q \leq \Delta G^0$  (mfe)). Line 11: Coarse representation of the centroid structure followed by its free energy and distance to the ensemble. The centroid structure has minimal base pair distance to all other secondary structures in the ensemble; that is, it contains all basepairs with  $p_{ij} > 0.5$ . Line 12, 13: This line is split here into two lines to fit on the page width; ensemble diversity  $\langle d \rangle = \sum_{i,j} p_{ij} \cdot (1 - p_{ij})$  [26, 27]. Line 14: **RNAPlfold** is used to predict locally stable

mark bases as unpaired, and up- or downstream paired, respectively. Line 10: Coarse representation of the pair probabilities in pseudo-bracket notation followed by the ensemble free energy in kcal/mol. This notation uses the letters “.”, “,”, “|”, “{”, “}”, “(”, and “)” to mark bases as essentially unpaired, weakly paired, strongly paired without 5’ or 3’ preference, weakly upstream or downstream paired, or strongly up- or downstream paired, respectively. The ensemble free energy is lower (more negative) than that of the mfe structure with the increasing number of structures in the partition function  $Q$  ( $\Delta G^0$  (ensemble) =  $-RT \ln Q \leq \Delta G^0$  (mfe)). Line 11: Coarse representation of the centroid structure followed by its free energy and distance to the ensemble. The centroid structure has minimal base pair distance to all other secondary structures in the ensemble; that is, it contains all basepairs with  $p_{ij} > 0.5$ . Line 12, 13:

This line is split here into two lines to fit on the page width; ensemble diversity  $\langle d \rangle = \sum_{i,j} p_{ij} \cdot (1 - p_{ij})$  [26, 27]. Line 14: **RNAPlfold** is used to predict locally stable

structures. The given options and parameters have the following meaning (see Fig. 5):  
 - --noLP disallows lonely basepairs that have no neighboring and stacking pairs; -W 80 restricts calculations to a sliding window  $W = 80$ ; -L 40 restricts basepairs to a maximum distance  $L = 40$  of nucleotides; -u 10 computes the probability that a stretch of 10 consecutive nucleotides is unpaired

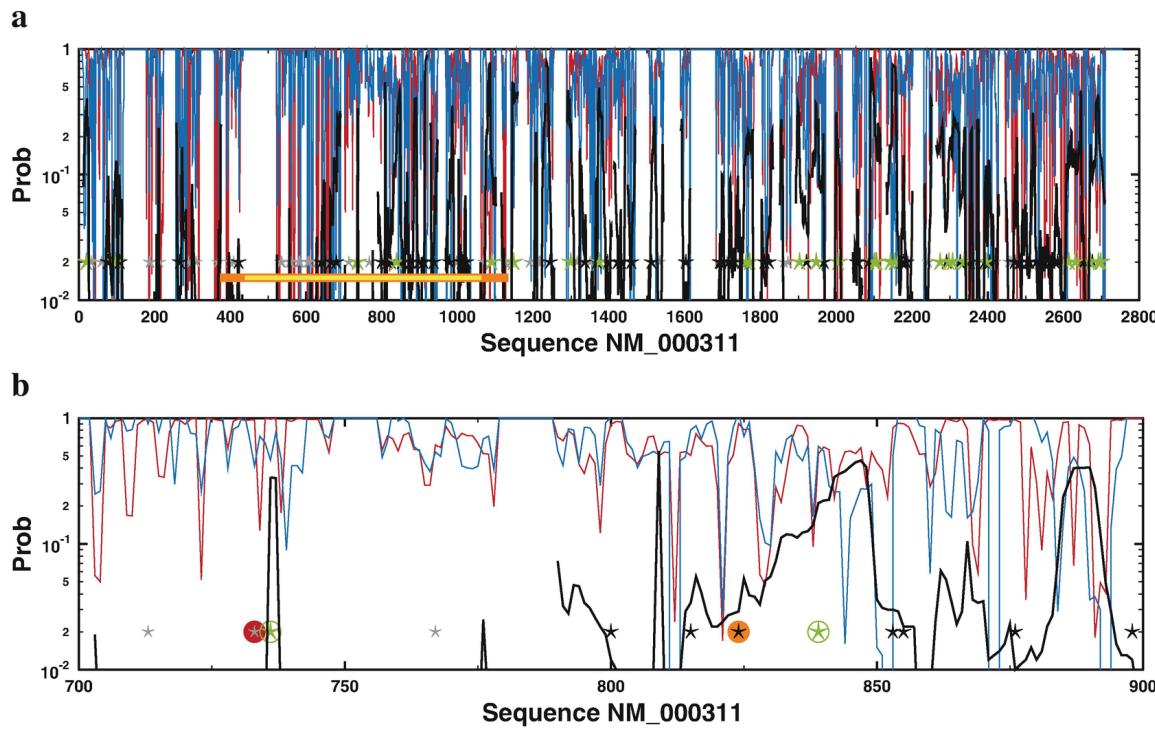


**Fig. 3** Part of PrP mfe structure. The mfe structure was predicted by RNAfold (Fig. 2). The two sections of this structure show the location of four selected target regions, named T733, T736, T824, and T839. The target regions are colored according to the accessibility of their central 10 nucleotides ( $u_{10}(T733) = 0.002$  (red),  $u_{10}(T736) = 0.338$  (green),  $u_{10}(T824) = 0.029$  (orange), and  $u_{10}(T839) = 0.211$  (green)) as predicted by RNAPlfold (Fig. 4)

The sequence download from GenBank already allows a user to choose FASTA format, which only differs from Vienna format by multiple lines of usually fixed length containing the sequence.

If the sequence is only available in other formats, one of the many sequence format conversion tools can be used, either locally installed or by web usage. For example, one can use the program `readseq` (see Table 1 and Fig. 2, line 1).

3. For a given sequence of length  $n$ , the program `RNAfold` [28, 29] predicts a thermodynamically optimal secondary structure, which is the structure of minimum free energy (mfe), and the partition function by which the basepairing probability  $p_{i,j}$  is calculated for any pair  $(i, j)$  in the sequence ( $1 \leq i < j \leq n$ ) (see Fig. 2, lines 5–13). A part of the mfe structure of PrP mRNA is shown in Fig. 3. The basepairing probabilities  $p_{i,j}$  are usually visualized as dots of area  $p_{i,j}$  in a matrix plot in PostScript format, which is automatically generated by `RNAfold`. Thus, the PostScript file contains lines



**Fig. 4** Accessibility of GU cleavage sites in the human PrP mRNA. The x-coordinate represents the position in the mRNA sequence (GenBank LOCUS NM\_000311; version 3), and the pairing probability is displayed on the y-axis. The global basepairing probability for each nucleotide (red line) was calculated by `RNAfold`. The local basepairing probability  $p_{\text{local}}$  for each nucleotide (blue line) was calculated by `RNAplfold` with maximal distance of pairing nucleotides  $L = 40$  and a length of sliding window  $W = 80$ . The black lines represent the probability  $u_{1,0}$  that a sequence

of length  $l = 10$  centered at this position is not basepaired. The global basepairing probability  $p_{\text{global}}$  (red lines) was calculated by RNAfold. GU sequences are represented by stars. Gray stars show GU sites predicted to be inaccessible with  $u_{10} < 0.01$ . Black stars show medium accessible GUs with  $u_{10} > 0.01$ , and green stars indicate the most accessible GUs with  $u_{10} > 0.1$ . The yellow bar indicates the open reading frame from position 373 to position 1134; the orange bar indicates the mature protein after post-translational modifications (439 to 1062). The graph in (a) shows the full mRNA sequence (positions 1–2755), while the graph in (b) shows a zoom to positions 700–900. The GUs selected as DNAzyme target sites are marked by circles.

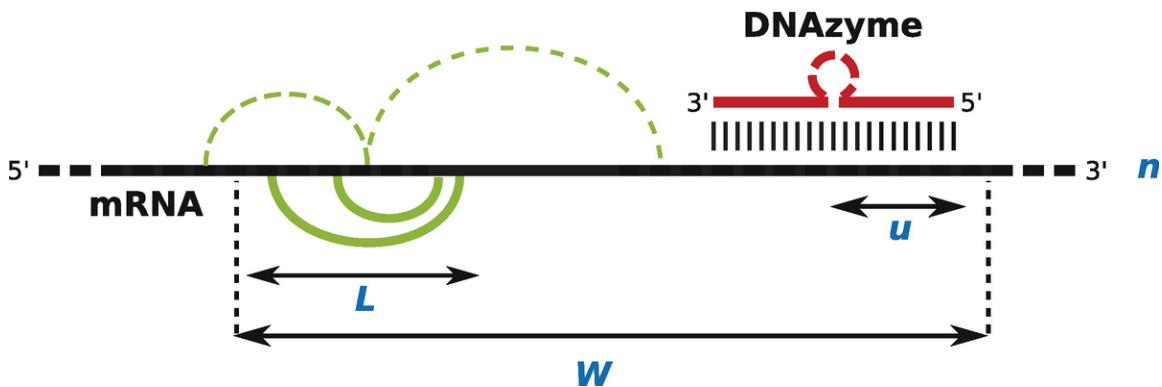
i j sqrt(p) ubox

for all pairs  $(i, j)$  and the square root of the corresponding basepairing probabilities  $\sqrt{p_{i,j}}$ , which can be extracted from the file to yield the pairing probability for each nucleotide  $i$ :

$$p_i = \sum_{i < j} p_{i,j} + \sum_{j < i} p_{j,i}$$

(see red line in Fig. 4).

4. The program **RNAPlfold** is able to predict locally stable secondary structures of an mRNA sequence that is thought as a near perfect substitute for prediction of kinetically optimal structures, which is not feasible for lengths of mRNAs [30, 31]. To predict the locally stable structures, **RNAPlfold** slides a window of length  $W$  nucleotides along the sequence (Fig. 5). Within this window, basepairs  $(i, j)$  are restricted to have a distance  $L \leq j - i$ . Thus, this procedure preferentially predicts structural elements consisting of stem loops with small hairpin loops, which are kinetically fast forming. The program produces two output files: one is a PostScript file with name `<header>_dp.ps` that shows a dotplot of basepair probabilities, which are averaged over all windows containing that basepair (blue line in Fig. 4); the other is a text file with name `<header>lunp`. The latter contains a matrix of size  $n \times x$ ; each line gives a position  $i$  followed by the probability  $u_1$  that the nucleotide at position  $i$  is unpaired, followed by the probability  $u_2$  that the nucleotide stretch  $[i - 1, i]$  is unpaired, etc. up to the probability  $u_x$  that a stretch  $[i - x + 1, i]$  is unpaired with length given by the program option “`-u x`” (black line in Fig. 4).



**Fig. 5** Scheme for RNAPlfold. The mRNA sequence is locally folded in a sliding window approach of length  $W$ . Basepairing is restricted to a maximum distance  $L$  of nucleotide pairs within a window  $W$ ; that is, basepairs depicted by solid green lines are possible, but not those depicted by dashed green lines. The length  $u$  represents the stretch of consecutive nts within a target site starting at its 3' end for which the accessibility is computed

RNAPlfold was shown to predict quite accurately the *in vivo* accessibility of target sites for interaction with siRNAs [32].

5. Given the basepairing probabilities for thermodynamic folding of the full mRNA sequence from **step 3** and the basepairing probabilities for local folding from **step 4**, it is easy to search for accessible cleavage sites on the mRNA and other requirements. In the example in Fig. 4, GU dinucleotides are marked by stars in three different categories with  $u_{10} < 0.01$  as mostly inaccessible (gray),  $0.01 \leq u_{10} \leq 0.1$  (black), and  $u_{10} > 0.1$  as accessible (green).
6. For any selected cleavage site (see **step 5**), the neighboring sequence regions available for complex formation by the DNAzyme are known. The length and sequence of these binding regions determine the stability of the mRNA:DNAzyme complex and thus influence catalytic turnover. Prediction of the complex stability, which is an RNA:DNA hybrid, needs the corresponding thermodynamic parameters and at least an entropic parameter for the DNAzyme's core loop. We choose the program **poland** with thermodynamic parameters for RNA:DNA hybrid stability [33], which are valid for 1 M NaCl concentration, and model the loop as a mismatch in the double strand.

Results from **poland** predictions for several selected PrP regions are shown in Table 2. As a standard we choose binding-arm lengths as

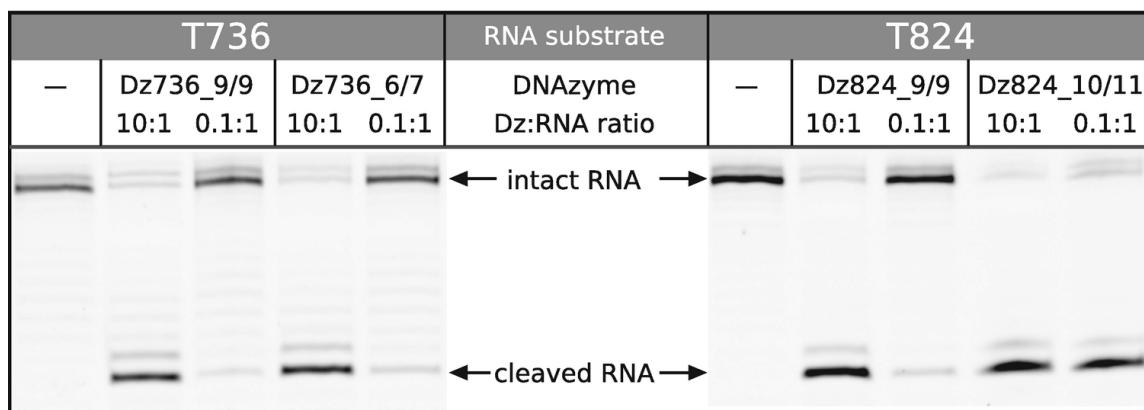
are shown in Table 2. As a standard, we choose binding-mutants as 9+9 basepairs and named these accordingly Dz<pos>\_9/9 cleaving at positions 733, 736, 824, and 839 of the PrP mRNA; the corresponding RNA oligonucleotide targets were named T<pos> (Fig. 3). The midpoint temperature  $T_m$  of the transition, at which 50% of the DNAzyme:RNA complex is denatured, and the degree of denaturation  $\Theta$  at a certain temperature are representatives of uncleaved and cleaved complex stability. Note that the absolute temperatures are predicted for 1 M NaCl conditions, which are not usually present under in vitro or in vivo conditions; therefore, only the relative temperatures are of relevance.

**Table 2** Denaturation temperatures  $T_m$  and degrees of denaturation  $\Theta$  at 37 °C for 10–23 DNAzyme:RNA complexes. Values were predicted by poland using thermodynamic parameters for RNA:DNA hybrids [33] with  $c_{\text{DNAzyme:RNA}} = 1 \times 10^{-6}$  M. In case of the “full-length” complex, the core loop was set to a mismatch in the RNA at position of the G of the GU cleavage site. The RNA targets are named T<Pos>\_L1/L2 with the position pos of the cleavage site in the PrP sequence (see Figs. 3 and 4) and lengths L1 and L2 of the 5' and 3' binding sequences, respectively, bound by the corresponding DNAzyme. For sequences, see Fig. 3 and [34]

10–23 DNAzyme complex with	RNA substrate		5' reaction product		3' reaction product	
	$T_m$ /°C	$\Theta_{37}$ /°C	$T_m$ /°C	$\Theta_{37}$ /°C	$T_m$ /°C	$\Theta_{37}$ /°C
T733_9/9	74.25	0.03	63.25	0.01	60.25	0.02
T736_9/9	77.25	0.02	62.75	0.05	69.25	0.01
T736_6/7	53.25	0.03	43.25	0.24	40.75	0.38
T824_9/9	35.75	0.69	26.25	0.99	24.75	0.98
T824_10/11	46.75	0.04	39.75	0.34	32.75	0.83
T839_9/9	53.75	0.08	46.75	0.12	32.25	0.81

The most stable Dz736\_9/9 as well as the least stable Dz824\_9/9 cleaved their corresponding short RNA oligonucleotide but were only weakly active under multiple turnover conditions (Fig. 6). We reasoned that the high stability of complex Dz736\_9/9:T736 even after cleavage leads to a low degree of dissociation hindering multiple turnover on

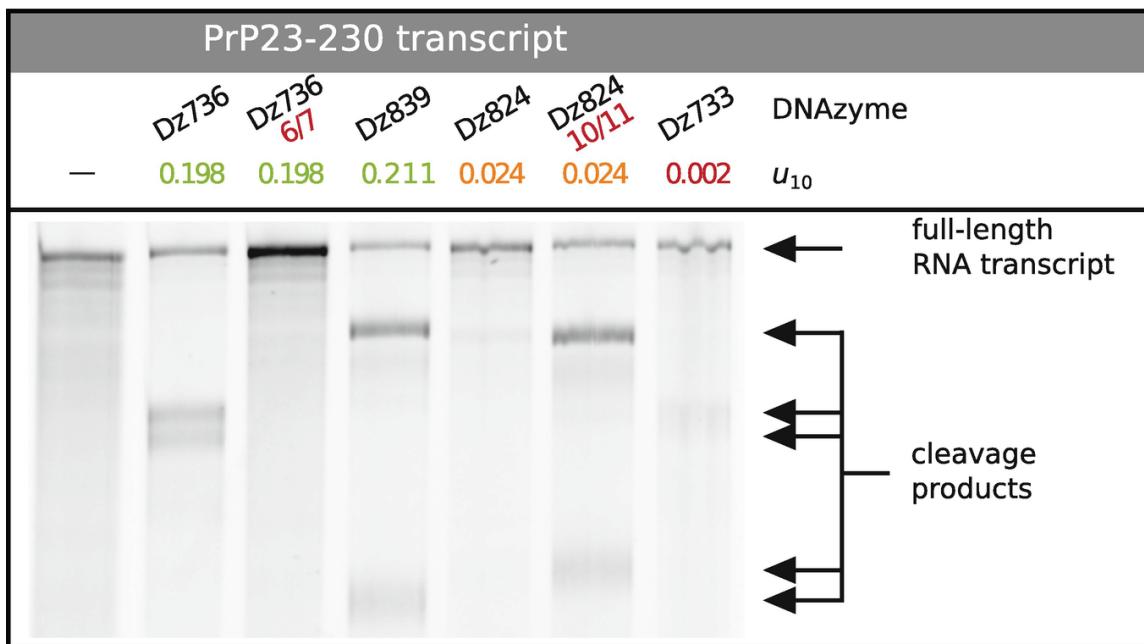
leads to a low degree of association hindering multiple turnover, on the contrary, the low stability of complex Dz824\_9/9:T824 might disfavor (re-)association of this DNAzyme with its target. This seems to be correct because shortening the binding arms of Dz736\_9/9 to 6+7 basepairs decreased its complex stability and enhanced its activity slightly under multiple turnover conditions; lengthening the binding arms of Dz824\_9/9 to 10+11 basepairs leads to increased complex stability and high activity under multiple turnover conditions (Fig. 6).



**Fig. 6** Single and multiple turnover behavior of DNAzyme variants. The RNA substrates T736 and T824 were incubated with Dz736 and Dz824 variants with shortened or lengthened substrate binding arms in a 10:1 or a 0.1:1 enzyme-to-substrate molar ratio (4  $\mu$ M and 0.04  $\mu$ M DNAzyme, respectively; 0.4  $\mu$ M RNA; 50 mM Tris·HCl, pH 7.5; 10 mM MgCl<sub>2</sub>). After cleavage, samples were separated by 18% denaturing polyacrylamide gel electrophoresis (PAGE; 19:1 (v/v) acrylamide:bisacrylamide with 7 M urea). During 3 h at 37 °C, T736 was cleaved by both Dz736 variants, but complete cleavage was only observed with a tenfold molar excess of DNAzyme. Multiple turnover cleavages could not be achieved. T824 was cleaved by both Dz824 variants. While Dz824\_9/9 did not display any multiple turnover behavior, Dz824\_10/11 cleaved a tenfold molar excess of substrate to completion. T, target; Dz, 10–23 DNAzyme

The activity of these DNAzymes during in vitro transcription of the full-length PrP mRNA (Fig. 7) was again in line with the RNApolfold (**step 4**) and poland predictions (**step 6**): Dz733\_9/9 is inactive due the low accessibility of the target, Dz736\_9/9 is active as predicted despite having its cleavage site located only three nucleotides upstream of the inactive Dz733\_9/9, Dz824\_10/11 is active but not Dz824\_9/9, and activity of Dz839\_9/9 is highest in agreement with the high accessibility of the target and intermediate stability of

Dz839\_9/9:target.



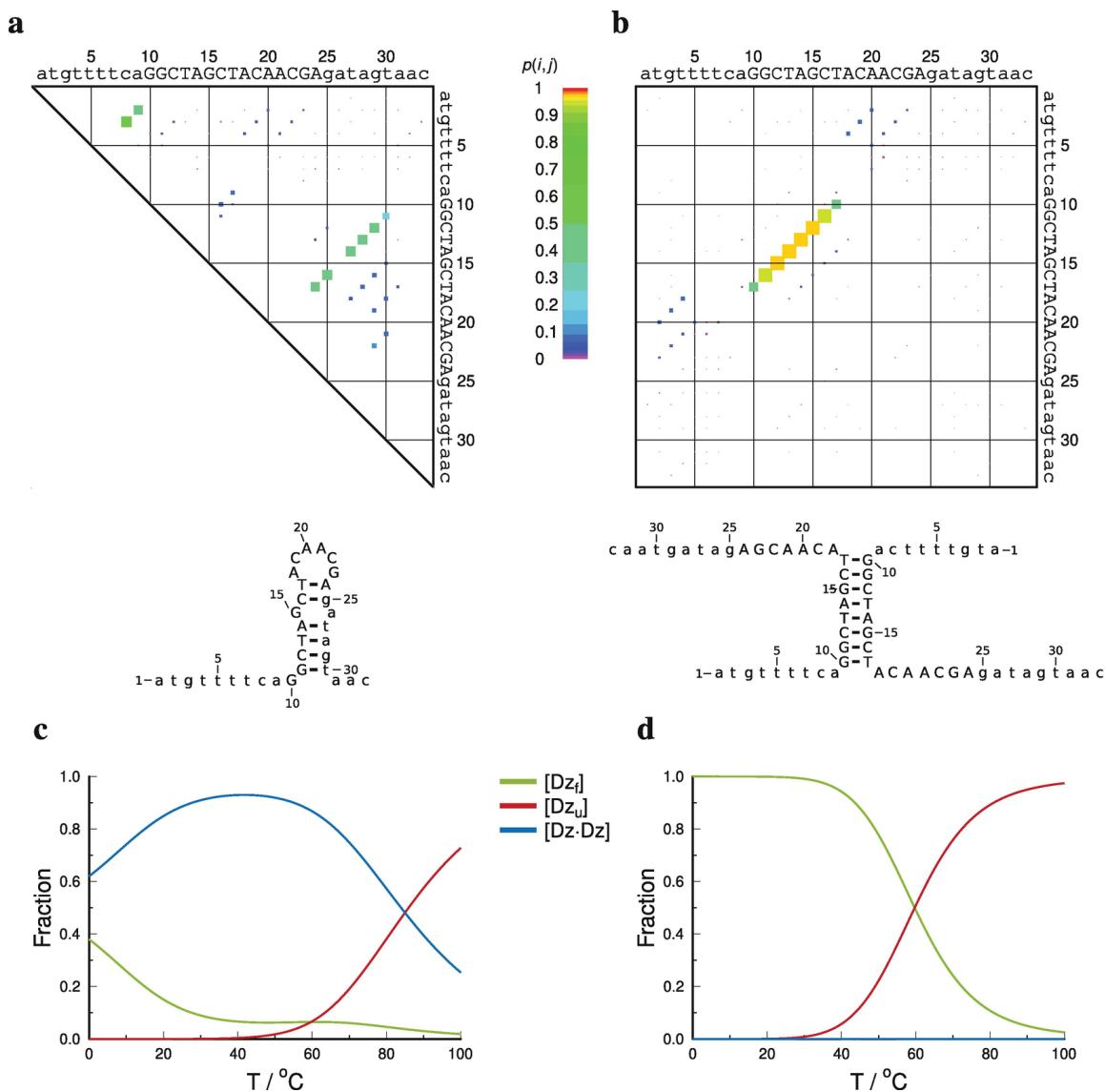
**Fig. 7** DNAzyme activity during RNA in vitro transcription. Human PrP23-230 RNA was produced via in vitro transcription in the presence of 10  $\mu$ M of either Dz736, Dz736\_6/7, Dz839, Dz824, Dz824\_10/11, or Dz733. The reaction was carried out for 1 h at 37 °C. Afterward, the products were purified, fluorescently labeled via click chemistry, and separated via 5% PAGE; for further details, see Note 2. The DNAzymes cleaved the RNA transcripts with observed efficacies that are consistent with the predicted target site accessibility  $u_{10}$  (Fig. 4): green, highly accessible; orange, medium accessible; red, inaccessible

7. Any DNAzyme should not have a stable structure neither as monomer nor as homodimer; this would disfavor complex formation between DNAzyme's binding arms and the target. Appropriate structure predictions are possible with Mfold or RNAfold for an intramolecular, monomeric structure and DINAMelt or RNACofold for a dimeric structure (see Table 1). Note that DNA parameters have to be used for the DNAzymes.

As an example, a part of the output using DINAMelt for DZ824\_9+9 is shown in Fig. 8. The free energy of the monomer mfe structure (Fig. 8a) is  $\Delta G = -0.11$  kcal/mol; a structure with two additional basepairs (positions 2–3:9–8) has  $\Delta G = 0.3$  kcal/mol.

Table 1 lists the DNAzyme sequences and their target sites.

that is, the DNAzyme has only a very low tendency to form intermolecular basepairs. The free energy of the dimer mfe structure (Fig. 8b) is  $\Delta G^0 = -6.4$  kcal/mol and that of the ensemble  $\Delta G_{37^\circ C} = -9.2$  kcal/mol calculated as difference of the ensemble free energies at 0 °C and 150 °C. This degree of dimer stability cannot be diminished because of the partial palindromic sequence of the 10–23 DNAzyme, which gives rise to the helix in the dimeric structure (Fig. 8b). This is, however, not a problem: the fraction of dimer is close to zero at relevant, micromolar (or lower) concentrations of the 10–23 DNAzyme (Fig. 8d).



**Fig. 8** Structures and stabilities of monomeric and dimeric Dz824\_9+9. In the

dotplots of monomer (a) and dimer (b), the probability of a basepair is shown; that is,

the area of a dot at position  $(i, j)$  is proportional to its basepair probability  $p(i, j)$ . In addition, dots are colored according to  $p(i, j)$  (for colors see central bar). Basepairing probabilities were predicted by **DINAMelt**; mfe structures were predicted by **Mfold** for the monomer and **DINAMelt**'s “Two State melting (hybridization),” respectively. Both tools used DNA parameters. Plots of relative concentrations of folded monomer ( $Dz_f$ , green), unfolded monomer ( $Dz_u$ , red), and dimer ( $Dz \cdot Dz$ , blue) predicted by **DINAMelt** using DNA parameters at 150 mM NaCl, 2 mM MgCl<sub>2</sub>, and total concentration [DNAzyme] = 1 M (c) and [DNAzyme] =  $1 \times 10^{-6}$  M (d), respectively. Dotplots and concentration plots were redrawn from text output of **DINAMelt** using **gle** (Table 1); the structure schemes were redrawn using **R2R** [22]

As a summary of this last consideration, secondary structure of a DNAzyme should only be of relevance in rare cases. For example, only a (full) complementarity of one of the binding arms to the DNAzyme's core sequence might lead to unwanted stability of a monomer structure.

---

## 4 Notes

1. We wrote a **perl** script named **prob.pl** (Table 1) that should ease the execution of **steps 3, 4, and 5** of the Methods section. For execution, the script needs **perl** and a local installation of **RNAfold** and **RNAp1fold**, which are part of the Vienna package (see Table 1).

The program (see Fig. 9, lines 1–26) reads sequence files in FASTA, Vienna, GenBank, and EMBL format, respectively; files may be compressed by **gzip**. In a sequence, “T” is equivalent to “U”; characters other than “A,” “C,” “G,” “T,” “U” are replaced by “N”. The RNA site required for cleavage is given via option **-m**; the option allows a user to use (simple) regular expressions like “[AG][UC]” for RY. The option **-o** accepts an even set of numbers, which is used to color open reading frames (see yellow bars in Fig. 4). The option **-H** takes one word, which is printed above the graphics output; multiple words have to be bracketed by quotes. All other options will be handed over to **RNAfold** and **RNAp1fold**, respectively.

```

1 $ prob.pl -h
2 Usage: prob.pl [OPTIONS]
3     Reads a sequence file with an RNA sequence; supported formats are
4         FASTA, Vienna, GenBank, and EMBL.
5     Determines pairing probabilities at a given temperature via
6         RNAfold and RNAPlfold.
7     Extracts from RNA*fold's output the pairing probabilities, and
8     plot it via GLE.

10    OPTIONS:
11    -f <gene.vie[.gz]> Input file
12                    (lines starting with [!%#] are not read)
13                    T is equivalent to U; [^AUCG]->N
14    [-m "regex"] cleavage site; allows for sequence or regexp
15                    (only containing square bracket pairs []);
16                    default: "GU"
17    [-o <#,#,...>] Start, end of ORF
18    [-t <#>] Temperature in degree centigrade (20 <= t <= 50)
19    [-p <gene.prob>] Output file with pairing probabilities
20    [-g <gene_RNAfold.gle>] Output file (GLE script for fold data)
21    [-u <#>] Mean prob of unpaired regions (l=1 to l=u)
22                    for RNAPlfold (default 10)
23    [-W <#>] Window size for RNAPlfold (default 80)
24    [-L <#>] Separation of a base pair for RNAPlfold (default 40)
25    [-k] Keep the intermediate files of RNA*fold (default: no)
26    [-H <header>] Printed above the graphics (ONE word!)

28 $ prob.pl -f NM_000311_3.gb
29 $ more NM_000311_3.dat
30 ...
31 ! pp[i] = pairing prob from partition function
32 ! mm[i] = nts paired in mfe structure
33 ! sp[i] = -Sum pp[i] * ln(pp[i]) = measure of well-definedness
34 ! pfL[i] = pairing prob from local partition function
35 ! spL[i] = -Sum ppL[i] * ln(ppL[i]) = measure of well-definedness
36 ! GU      = position of GU in sequence
37 ! GU>.01 = position of GU in sequence with pp[i]>0.01
38 ! GU>.1  = position of GU in sequence with pp[i]>0.1
39 ! lunp   = mean probability that regions of length 10 centered at i are unpaired
40 ! i pp    pp    mm    sp    pfL   pfL   spL   GU    GU>.01 GU>.1  lunp  seq
41 1 1.0   0.003 0.000 0.004 1.0   0.000 0.000  *    *    *    *    A
42 2 1.0   0.696 1.000 0.522 1.0   0.018 0.026  *    *    *    *    U
43 3 1.0   0.814 1.000 0.451 1.0   0.054 0.076  *    *    *    *    U
44 ...
45 732 0.965 0.965 1.000 0.170 0.713 0.713 0.517  *    *    *    0.002 A
46 733 0.973 0.973 1.000 0.140 0.416 0.416 0.453 0.020  *    *    *    0.002 G
47 734 0.128 0.128 0.000 0.172 0.718 0.718 0.515  *    *    *    0.002 U
48 735 0.988 0.988 1.000 0.078 0.635 0.635 0.531  *    *    *    0.003 G
49 736 0.983 0.983 1.000 0.098 0.464 0.464 0.482 0.020 0.020 0.020 0.338 G
50 737 0.993 0.993 1.000 0.049 0.763 0.763 0.492  *    *    *    0.333 U
51 ...
52 2755 1.0   0.002 0.000 0.002 1.0   0.000 0.000  *    *    *    *    A
54 $ gle NM_000311_3.gle

```

**Fig. 9** `prob.pl`. Help page, program execution, and resulting data file

Execution of `prob.pl` (Fig. 9, line 28) with sequence file `NM_000311_3.gb` produces two output files. The first contains a matrix of data (for example, see Fig. 9, lines 30–52), and the second is a `gle` script. Execution of the `gle` script (Fig. 9, line 54) produces a graphical representation of the data (Fig. 4a). The data matrix consists of  $n$  rows for a sequence of length  $n$  and 13 columns of following content:

- 1: Sequence position  $1 \leq i \leq n$ .
  - 2–
  - 3:  $p_i$  from the global prediction by `RNAfold`; in column 2, values are valid if  $(i - 10) \leq x \leq (i + 11)$  with  $x$  the position of a cleavage site; see red line in Fig. 4.
  - 4: Values are 1 or 0 if the nucleotide  $i$  is paired or unpaired, respectively, in the mfe structure predicted by `RNAfold`.
  - 5: Well-definedness or positional entropy  $S(i) = -\sum_j p_{ij} \cdot \log_2(p_{ij})$  of globally predicted pairs [35, 36].
  - 6–
  - 7:  $p_i$  from the local prediction by `RNAPlfold`; in column 6, values are valid if  $(i - 10) \leq x \leq (i + 11)$  with  $x$  the position of a cleavage site; see blue line in Fig. 4.
  - 8: Well-definedness or positional entropy of locally predicted pairs.
  - 9–
  - 11: If a value is given but not a star, the position is a cleavage site with low accessibility, accessibility  $u_{10} > 0.01$ , or  $u_{10} > 0.1$ , respectively; this is used to plot the stars in Fig. 4.
  - 12: Mean probability  $u_l$  that regions of length  $l$  centered at  $i$  is unpaired; see black line in Fig. 4.
  - 13: Nucleotide at position  $i$ .
2. For activity assays of DNAzymes during in vitro transcription, samples were prepared in 40 mM Tris·HCl pH 7.9 with 10 mM dithiothreitol, 10

mM NaCl, 6 mM MgCl<sub>2</sub>, 2 mM spermidine, 1 U/μl RNase inhibitor, 20 ng/μl plasmid DNA (linearized huPrP23-230 in pET11a), 1 mM each of ATP, GTP, and CTP as well as 0.75 mM UTP and 0.25 mM 5-Azido-UTP, 0.25 μM of DNAzyme denatured before, and 15 U/μl T7 RNA polymerase. The final reaction volume in these experiments was 10 μl. In vitro transcription was carried out for 1 h at 37 °C in the dark. Afterward, for purification with gel filtration columns (illustra MicroSpin G-25 columns), the samples had to be adjusted to a volume of 25 μl by adding H<sub>2</sub>O.

Labeling was carried out by adding 0.2 mM dibenzylcyclooctyne-sulfo-Cy5, which results in a covalent attachment of the fluorescent label to the RNA via copper-free “click chemistry.” The labeling reaction was incubated for 1 h at 37 °C and shaking at 300 rpm. Afterward, one volume of 2× RNA loading buffer (94% formamide, 25 mM EDTA pH 8.0, 0.02% Bromophenol blue (w/v), 0.02% Xylene cyanol (w/v)) was added to each sample followed by denaturation at 96 °C for 15 min. The samples were separated via 5% PAGE (19:1 (v/v) acrylamide:bisacrylamide) with 7 M urea and analyzed by detecting Cy5 fluorescence.

## Acknowledgements

We thank Prof. em. Dr. Dr. h. c. Detlev Riesner and Dr. Manuel Etzkorn for support and discussions.

---

## References

1. Geyer C, Sen D (1997) Evidence for the metal-cofactor independence of an RNA phosphodiester-cleaving DNA enzyme. *Chem Biol* 4:579–593. [https://doi.org/10.1016/S1074-5521\(97\)90244-1](https://doi.org/10.1016/S1074-5521(97)90244-1)
2. Torabi S, Wu P, McGhee C, Chen L, Hwang K, Zheng N, Cheng J, Lu Y (2015) In vitro selection of a sodium-specific DNAzyme and its application in intracellular sensing. *Proc Natl Acad Sci U S A* 112:5903–5908. <https://doi.org/10.1073/pnas.1420361112>  
[Crossref]
3. Santoro S, Joyce G (1998) Mechanism and utility of an RNA-cleaving DNA enzyme. *Biochemistry* 37:13330–13342. <https://doi.org/10.1021/bi9812221>

[Crossref]

4. Breaker R, Emilsson G, Lazarev D, Nakamura S, Puskarz I, Roth A, Sudarsan N (2003) A common speed limit for RNA-cleaving ribozymes and deoxyribozymes. *RNA* 9:949–957. <https://doi.org/10.1261/rna.5670703>  
[Crossref]
5. Parker D, Xiao Y, Aguilar J, Silverman S (2013) DNA catalysis of a normally disfavored RNA hydrolysis reaction. *J Am Chem Soc* 135:8472–8475. <https://doi.org/10.1021/ja4032488>  
[Crossref]
6. Nowakowski J, Shim PJ, Prasad GS, Stout CD, Joyce GF (1999) Crystal structure of an 82-nucleotide RNA–DNA complex formed by the 10–23 DNA enzyme. *Nat Struct Biol* 6(2):151–156. <https://doi.org/10.1038/5839>  
[Crossref]
7. Workman C, Krogh A (1999) No evidence that mRNAs have lower folding free energies than random sequences with the same dinucleotide distribution. *Nucleic Acids Res* 27:4816–4822. <https://doi.org/10.1093/nar/27.24.4816>  
[Crossref]
8. Rivas E, Eddy S (2000) Secondary structure alone is generally not statistically significant for the detection of noncoding RNAs. *Bioinformatics* 16:583–605. <https://doi.org/10.1093/bioinformatics/16.7.583>  
[Crossref]
9. Wada T, Becskei A (2017) Impact of methods on the measurement of mRNA turnover. *Int J Mol Sci* 18:2723. <https://doi.org/10.3390/ijms18122723>  
[Crossref]
10. Gilbert D (2003) Sequence file format conversion with command-line readseq. *Curr Protoc Bioinformatics Appendix 1:Appendix 1E*. <https://doi.org/10.1002/0471250953.bi01es00>
11. Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey A, Potter S, Finn R, Lopez R (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* 47:W636–W641. <https://doi.org/10.1093/nar/gkz268>  
[Crossref]
12. Gruber A, Lorenz R, Bernhart S, Neuböck R, Hofacker I (2008) The Vienna RNA websuite. *Nucleic Acids Res* 36:W70–W74. <https://doi.org/10.1093/nar/gkn188>  
[Crossref]
13. Hofacker I, Fontana W, Stadler P, Bonhoeffer S, Tacker M, Schuster P (1994) Fast

- folding and comparison of RNA structures. *Monatsh Chem* 125:167–188. <https://doi.org/10.1007/BF00818163>  
[Crossref]
14. Bernhart SH, Hofacker IL, Stadler PF (2006) Local RNA base pairing probabilities in large sequences. *Bioinformatics* 22(5):614–615. <http://bioinformatics.oxfordjournals.org/cgi/reprint/22/5/614.pdf>  
[Crossref]
  15. Bernhart S, Tafer H, Mückstein U, Flamm C, Stadler P, Hofacker I (2006) Partition function and base pairing probabilities of RNA heterodimers. 1:3. <https://doi.org/10.1186/1748-7188-1-3>
  16. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31:3406–3415. <https://doi.org/10.1093/nar/gkg595>  
[Crossref]
  17. Dimitrov R, Zuker M (2004) Prediction of hybridization and melting for double-stranded nucleic acids. *Biophys J* 87:215–226. <https://doi.org/10.1529/biophysj.103.020743>  
[Crossref]
  18. Markham N, Zuker M (2005) DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res* 33:W577–W581. <https://doi.org/10.1093/nar/gki591>  
[Crossref]
  19. Reuter J, Mathews D (2010) RNAsstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* 11:129. <https://doi.org/10.1186/1471-2105-11-129>  
[Crossref]
  20. Poland D (1974) Recursion relation generation of probability profiles for specific-sequence macromolecules with long-range correlations. *Biopolymers* 13:1859–1871. <https://doi.org/10.1002/bip.1974.360130916>  
[Crossref]
  21. Steger G (1994) Thermal denaturation of double-stranded nucleic acids: prediction of temperatures critical for gradient gel electrophoresis and polymerase chain reaction. *Nucleic Acids Res* 22:2760–2768. <https://doi.org/10.1093/nar/22.14.2760>  
[Crossref]
  22. Weinberg Z, Breaker R (2011) R2R – software to speed the depiction of aesthetic consensus RNA secondary structures. *BMC Bioinformatics* 12:3. <https://doi.org/10.1186/1471-2105-12-3>

[Crossref]

23. Benson D, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman D, Ostell J, Sayers E (2012) GenBank. Nucleic Acids Res 41:D36–D42. <https://doi.org/10.1093/nar/gks1195>  
[Crossref]
24. Amid C, Alako B, Balavenkataraman K, Burdett T, Burgin J, Fan J, Harrison P, Holt S, Hussein A, Ivanov E, Jayathilaka S, Kay S, Keane T, Leinonen R, Liu X, Martinez-Villacorta J, Milano A, Pakseresht A, Rahman N, Rajan J, Reddy K, Richards E, Smirnov D, Sokolov A, Vijayaraja S, Cochrane G (2020) The European nucleotide archive in 2019. Nucleic Acids Res 48:D70–D76. <https://doi.org/10.1093/nar/gkz1063>  
[Crossref]
25. Ogasawara O, Kodama Y, Mashima J, Kosuge T, Fujisawa T (2020) DDBJ database updates and computational infrastructure enhancement. Nucleic Acids Res 48:D45–D50. <https://doi.org/10.1093/nar/gkz982>  
[Crossref]
26. Wuchty S, Fontana W, Hofacker I, Schuster P (1999) Complete suboptimal folding of RNA and the stability of secondary structures. Biopolymers 49:145–165. [https://doi.org/10.1002/\(SICI\)1097-0282\(199902\)49:2%3C145::AID-BIP4%3E3.0.CO;2-G](https://doi.org/10.1002/(SICI)1097-0282(199902)49:2%3C145::AID-BIP4%3E3.0.CO;2-G)
27. Freyhult E, Gardner P, Moulton V (2005) A comparison of RNA folding measures. BMC Bioinformatics 6:241. <https://doi.org/10.1186/1471-2105-6-241>  
[Crossref]
28. Hofacker I (2003) Vienna RNA secondary structure server. Nucleic Acids Res 31:3429–3431. <https://doi.org/10.1093/nar/gkg599>  
[Crossref]
29. Steger G (2014) Secondary structure prediction. In: Hartmann R, Bindereif A, Schön A, Westhof E (eds) Handbook of RNA biochemistry. Weinheim: Wiley-VCH, pp 549–577
30. Schmitz M, Steger G (1996) Description of RNA folding by “simulated annealing”. J Mol Biol 255:254–266. <https://doi.org/10.1006/jmbi.1996.0021>  
[Crossref]
31. Flamm C, Fontana W, Hofacker I, Schuster P (2000) RNA folding at elementary step resolution. RNA 6:325–338. <https://doi.org/10.1017/S1355838200992161>  
[Crossref]
32. Tafer H, Ameres S, Obernosterer G, Gebeshuber C, Schroeder R, Martinez J, Hofacker I (2008) The impact of target site accessibility on the design of effective siRNAs. Nat

Biotechnol 26:578–583. <https://doi.org/10.1038/nbt1404>  
[Crossref]

33. Sugimoto N, Nakano S, Katoh M, Matsumura A, Nakamura H, Ohmichi T, Yoneyama M, Sasaki M (1995) Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. Biochemistry 34:11211–11216. <https://doi.org/10.1021/bi00035a029>  
[Crossref]
34. Victor J, Steger G, Riesner D (2017) Inability of DNAzymes to cleave RNA in vivo is due to limited Mg<sup>2+</sup> concentration in cells. Eur Biophys J 47:333–343. <https://doi.org/10.1007/s00249-017-1270-2>  
[Crossref]
35. Morgan S, Higgs P (1998) Barrier heights between ground states in a model of RNA secondary structure. J Phys A: Math Gen 31:3153–3170. <https://doi.org/10.1088/0305-4470/31/14/005>  
[Crossref]
36. Lorenz W, Clote P (2011) Computing the partition function for kinetically trapped RNA secondary structures. PLoS One 6:e16178. <https://doi.org/10.1371/journal.pone.0016178>  
[Crossref]

## 5. Fluorescence-Based Kinetic Measurements for RNA-Cleaving DNAzymes

Hannah Rosenbach<sup>1</sup>✉ and Gerhard Steger<sup>1</sup>✉

(1) Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany

✉ **Hannah Rosenbach**

Email: [hannah.rosenbach@hhu.de](mailto:hannah.rosenbach@hhu.de)

✉ **Gerhard Steger (Corresponding author)**

Email: [steger@biophys.uni-duesseldorf.de](mailto:steger@biophys.uni-duesseldorf.de)

### Abstract

Studying the catalytic behavior of biocatalysts under different conditions including temperature, buffer conditions, and cofactor concentrations is an important tool to understand their reaction mechanism. We describe two protocols that allow for the investigation of the catalysis of RNA-cleaving DNAzymes. The techniques include the use of FRET-labeled RNA substrates for studying the RNA-cleavage reaction in real-time under high throughput as well as RNA substrates labeled with a fluorescein molecule at the 5' end for gel-based assays. Both methods allow for an accurate determination of rate constants given a reaction model.

**Key words** Data analysis – Förster resonance energy transfer (FRET) – Observed rate constant ( $k_{\text{obs}}$ ) – Kinetic measurements – RNA cleavage

---

## 1 Introduction

Studying the kinetic properties of enzymes under different conditions is a prerequisite to understand the mechanism of chemical and biological reactions. The RNA-cleaving, Mg<sup>2+</sup>-dependent 10–23 DNAzyme is the most prominent DNAzyme and has been studied for more than two decades [1], but the mechanism by which the enzyme cleaves its RNA substrate remains still unknown. Although the 10–23 DNAzyme efficiently cleaves its RNA substrate *in vitro*, its *in vivo* activity is significantly lower. As reasons for this poor *in vivo* activity, a low concentration of intracellularly accessible Mg<sup>2+</sup> has been suggested [2] as well as high concentrations of monovalent metal ions such as Na<sup>+</sup> or K<sup>+</sup>, which compete for the same binding sites as the Mg<sup>2+</sup> ions without inducing catalytic activity [3].

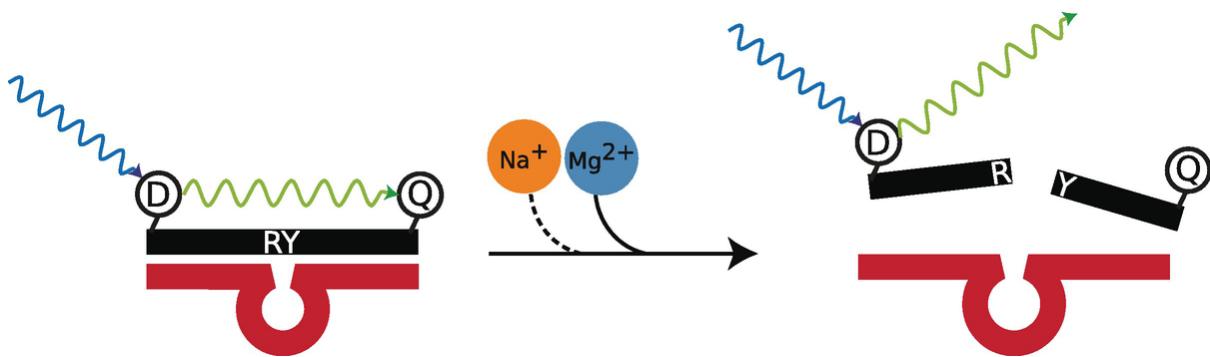
Initial studies on the 10–23 DNAzyme and its kinetic properties have been performed using <sup>32</sup>P-labeled RNA substrates [4]. The radioactive label as a standard tool to detect nucleic acids, however, has a short shelf life, needs additional safety measures, and has high disposal costs. Thus, great efforts have been made to develop alternative labeling techniques for nucleic acids [5], including fluorescence labeling. RNA-cleaving DNAzymes with fluorescence labels have been initially designed as catalytic beacons for the detection of metal ions in environmental samples [6]. The development of these biosensors involved the integration of a signal transducer or reporter molecule, which transforms the recognition event to a measurable signal. In the last two decades, catalytic beacons have been studied for the development of different signaling mechanisms involving non-labeled [7] or labeled [8] fluorescent, electrochemical [9], or colorimetric [10] sensors.

Among the variety of optical sensing methods, fluorescence-based techniques are most widely used. They provide high selectivity, sensitivity, and reproducibility, a broad linear range of detection, and an easy operational procedure [11]. The fluorescence signal is monitored as a function of time, and the rate increase is proportional to the concentration of the analyte.

Incorporation of fluorophores into the DNAzyme, however, is essential to obtain the desired properties, since nucleic acids lack intrinsic fluorescence. Fluorescence-based biosensors can be divided into labeled and non-labeled sensors. Labeled fluorescence-based sensors have the

fluorophore molecule covalently attached to the DNAzyme. For example, the initially reported  $\text{Pb}^{2+}$ -detecting DNAzyme-based biosensor had a fluorophore attached to one end of the substrate and a corresponding fluorescence quencher molecule attached to the DNAzyme [6]. In the absence of the metal ion, the substrate is bound to the DNAzyme and the fluorescence signal is quenched due to close proximity of fluorophore and quencher molecule. This principle is known as Förster resonance energy transfer (FRET). Thus, substrate cleavage in the presence of the analyte ( $\text{Pb}^{2+}$ ) can be measured by an increasing fluorescence signal.

The principle of fluorescence-labeled DNAzyme:substrate pairs has further been adapted for the measurement of kinetic activities of RNA-cleaving DNAzymes [2, 3, 12]. In order to reduce background fluorescence deriving from non-hybridized substrate molecules, several improvements have been made to optimize the signal-to-noise ratio: (i) an additional quencher molecule can be attached at the other end of the substrate [13], (ii) both the fluorophore and the quencher molecule can be linked to the two ends [14], or (iii) next to the cleavage site of the RNA substrate [15]. Based on the variety of fluorescence-labeled catalytic beacons, we designed a setup for the fluorescence-based measurements of RNA-cleaving DNAzymes using an RNA substrate, which is labeled with a fluorescein (6-FAM) molecule as a donor at the 5' end and a Black Hole Quencher 1 (BHQ-1) as a quencher at the 3' end (Fig. 1). This setup allows for a reliable and reproducible measurement of the cleavage reaction rate in the presence of different buffers and metal ions, as well as under varying temperatures in high throughput [3]. In this protocol, we describe how to carry out the assay in high throughput, and we provide a detailed description of the data analysis. Furthermore, we describe a gel-based assay with fluorescein-labeled RNA substrates that can complement the FRET-based kinetics and underlines the reproducibility of the FRET-based kinetic assay.



**Fig. 1** Scheme of fluorescence-based measurements of RNA-cleaving DNAzymes. A DNAzyme (red) cleaves an RNA substrate (black) binding both  $\text{Na}^+$  and  $\text{Mg}^{2+}$  ions, but only  $\text{Mg}^{2+}$  functioning as a catalytic cofactor. The RNA substrate is labeled with 6-FAM as donor (D) at the 5' end and BHQ-1 as quencher (Q) at the 3' end. In the intact RNA, prior to cleavage between, for example, a purine R and a pyrimidine Y, the donor is quenched due to close neighborhood of the quencher; in the cleaved RNA the donor shows its undisturbed fluorescence [16]

## 2 Materials

Prepare all buffers and stock solutions using deionized water ( $18 \text{ M}\Omega\text{cm}$ ) from a water purification system with a filter of  $0.2 \mu\text{m}$ .

### 2.1 Activity Assay with FRET-Labeled RNA Substrates

1. 4  $\mu\text{M}$  DNAzyme in  $\text{H}_2\text{O}$ .
2. 4  $\mu\text{M}$  FRET-labeled RNA substrate in  $\text{H}_2\text{O}$ . The RNA substrate is labeled with a 6-FAM at the 5' end and BHQ-1 at the 3' end (*see Note 1*).
3. 0.5 M Tris–HCl pH 7.5 stock solution; use this stock solution to prepare 50 mM Tris–HCl pH 7.5.
4. 4 M NaCl in 50 mM Tris–HCl pH 7.5; use this stock solution to prepare 80, 240, 320, and 400 mM NaCl in 50 mM Tris–HCl pH 7.5.
5. 1 M  $\text{MgCl}_2$  stock solution; use this stock solution to prepare 8 mM  $\text{MgCl}_2$  solution in 50 mM Tris–HCl pH 7.5.

6. 8 mM ethylenediaminetetraacetic acid (EDTA).
7. 1.5 mL reaction tubes.
8. 15 mL reaction tube.
9. 384 Well non-binding microplate (e.g., Greiner Bio-One).
10. Plate reader with temperature control, reagent injector, and excitation and emission wavelength of 484 and 530 nm, respectively (e.g., CLARIOstar, BMG Labtech).
11. Sealing tape for microplate (e.g., Polyolefin Acrylate, Thermo Scientific).
12. Heating block for 1.5 mL reaction tubes (37–73 °C).
13. Tabletop centrifuge.
14. Pipets (2–200 µL).
15. Pipet tips with filter.

## 2.2 Gel-Based Activity Assays with Fluorescein-Labeled RNA Substrates

1. 4 µM DNAzyme in H<sub>2</sub>O.
2. 4 µM Fluorescein-labeled RNA substrate in H<sub>2</sub>O.
3. 0.5 M Tris–HCl pH 7.5.
4. 1 M NaCl in water.
- 5.

20 mM MgCl<sub>2</sub> in water.

6. 10 mM EDTA.
7. 1.5 mL reaction tubes.
8. Heating block for 1.5 mL reaction tubes (37–73 °C).
9. 18% polyacrylamide solution for denaturing polyacrylamide gel electrophoresis (PAGE):  
7 M urea  
3 mL Tris-buffered EDTA (TBE) (10×)  
13.5 mL 40% acrylamide/bisacrylamide (19:1)  
ad 30 mL H<sub>2</sub>O  
0.1% (v/v) tetramethylethylenediamine (TEMED)  
0.1% (w/v) ammonium persulfate (APS)
10. Tabletop centrifuge.
11. 2× RNA loading dye:  
9.4 mL formamide  
0.5 M EDTA  
2 mg bromophenol blue  
2 mg xylene cyanol  
ad 10 mL H<sub>2</sub>O
12. 10× Tris-borate-EDTA (TBE) buffer stock solution: 890 mM Tris, 890 mM boric acid, 20 mM EDTA (pH 8.0), autoclaved.
13. Appropriate gel apparatus.
14. Suitable documentation system for gel visualization.
15. Pipets (1 μL–300 μL).

16. Pipet tips with filter.

## 2.3 Data Analysis

Any graphics program [17] that is able to plot given data and to fit the data to a given equation will be sufficient. These programs include Microsoft Excel, MATLAB, Origin, to name only three. Anyway, python with lmfit [18] and matplotlib [19, 20] libraries will do a perfect job for those with experience in programming. The plots shown in this chapter are produced with gle [21].

---

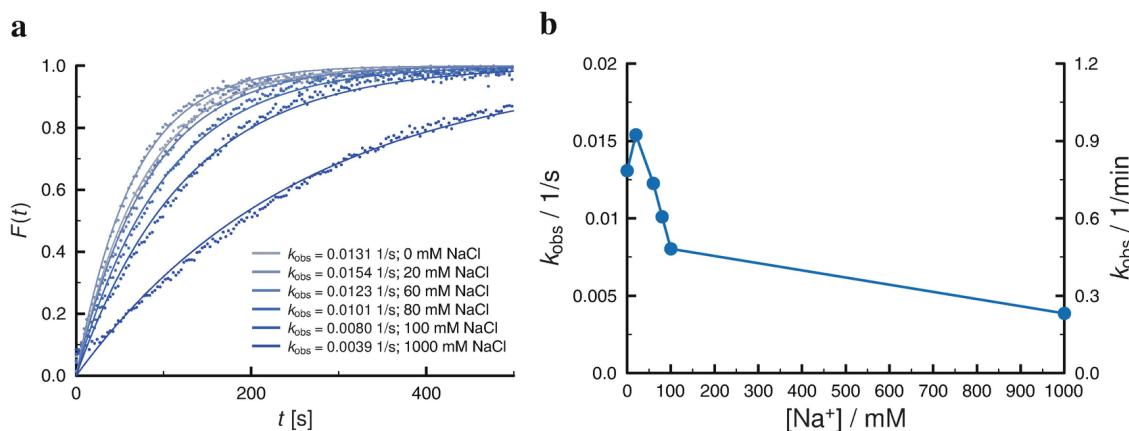
## 3 Methods

### 3.1 Activity Assay with FRET-Labeled RNA Substrates

The protocol is designed for measuring the activity of the RNA-cleaving 10–23 DNAzyme in the presence of 2 mM MgCl<sub>2</sub> under seven different conditions (NaCl concentrations ranging from 0 to 1000 mM in 50 mM Tris–HCl pH 7.5), including one control experiment without the DNAzyme.

The described experimental setup can be used to analyze for example the influence of monovalent metal ions such as Na<sup>+</sup> on the Mg<sup>2+</sup>-induced RNA cleavage. Figure 2 summarizes the results using this protocol, indicating that the observed rate constant ( $k_{\text{obs}}$ ) of the RNA-cleavage reaction is reduced with increasing NaCl concentrations [3]:

1. Preheat the heating block to 73 °C.



**Fig. 2** Activity assay with FRET-labeled RNA substrates indicating the influence of  $\text{Na}^+$  levels on DNAzyme catalysis. a: Normalized fluorescence intensities (dots) resulting from the FRET-based activity assays at 2 mM  $\text{Mg}^{2+}$  and different  $\text{Na}^+$  concentrations (see legend). The lines are fits of the respective data (Fig. 4). b: Rate constants ( $k_{\text{obs}}$ ) as obtained in a for the DNAzyme catalyzed RNA cleavage in the presence of different concentrations of  $\text{Na}^+$

2. Prepare a master mix DNAzyme:RNA solution—at least 35  $\mu\text{L}$  for the seven samples—containing the following:

50 mM Tris–HCl pH 7.5  
0.8  $\mu\text{M}$  DNAzyme  
0.8  $\mu\text{M}$  RNA substrate  
0.8 mM EDTA

3. Prepare a master mix RNA control solution—for better handling, prepare 20  $\mu\text{L}$ —without the DNAzyme (see Note 2):

50 mM Tris–HCl pH 7.5  
0.8  $\mu\text{M}$  RNA substrate  
0.8 mM EDTA

4. Incubate the solutions for 5 min at 73 °C.
5. Spin down the solutions briefly and incubate at room temperature for 10 min (see Note 3).
6. Dilute the solutions in an 1:4 ratio with 50 mM Tris–HCl pH 7.5 to obtain the following:

DNAzyme:RNA mix:

50 mM Tris–HCl pH 7.5  
0.2  $\mu\text{M}$  DNAzyme  
0.2  $\mu\text{M}$  RNA substrate  
0.2 mM EDTA

RNA control solution:

50 mM Tris–HCl pH 7.5  
0.2  $\mu\text{M}$  RNA substrate

0.2 mM EDTA

7. Preheat the plate reader to 37 °C.
8. In a 15 mL reaction tube, prepare 5 mL 8 mM MgCl<sub>2</sub> in 50 mM Tris–HCl pH 7.5 from the stock solutions.
9. Pipet 20 µL of the RNA control solution into the first well of a 384 well microplate, and mix with 20 µL 50 mM Tris–HCl pH 7.5 to a final volume of 40 µL.
10. Pipet 20 µL of the DNAzyme:RNA mix into the next wells of the same row and mix with 10 µL of either 50 mM Tris–HCl pH 7.5 or one of the NaCl solutions in 50 mM Tris–HCl.
11. Seal the plate with the sealing tape.
12. Place the sealed microplate and the tube containing the MgCl<sub>2</sub> solution in the plate reader and incubate for 30 min at 37 °C (*see Note 4*).
13. Wash the reagent injector with H<sub>2</sub>O, followed by a washing step with the 8 mM MgCl<sub>2</sub> solution in 50 mM Tris–HCl pH 7.5. Make sure that the tube of the reagent injector is fixed in the MgCl<sub>2</sub> solution.
14. Remove the sealing tape.
15. Set the injection volume for the MgCl<sub>2</sub> solution to 10 µL (*see Note 6*).
16. Set parameters at the plate reader as follows:

Excitation wavelength 484(20) nm

Emission wavelength 530(20) nm

Gain 1000

Focal height 6.8 mm

Top optics

The number of cycles 1000

Cycle time 4 s  
Pump speed 300 µL/s

17. Start the measurement.
18. When the measurement is finished, save the data in an appropriate format (see Subheading 3.3, step 1).

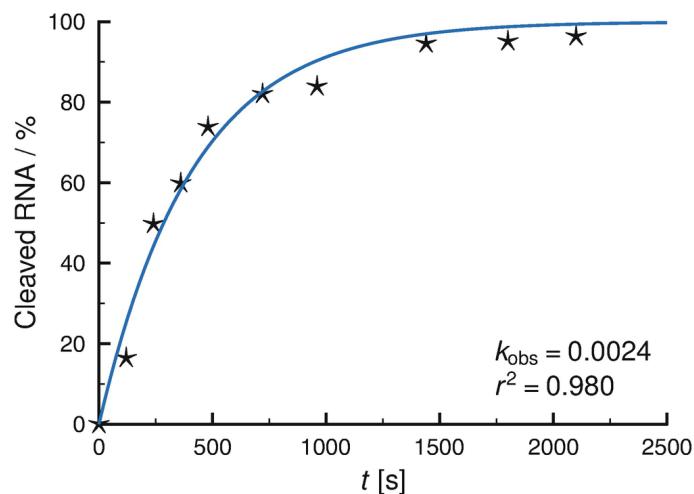
## 3.2 Gel-Based Activity Assays with Fluorescein-Labeled RNA Substrates

The protocol is designed for time-resolved kinetic measurements of the RNA-cleaving 10–23 DNAzyme with ten time points in the presence of 2 mM MgCl<sub>2</sub> and 1 M NaCl. For an example on this gel-based assay, see Fig. 3.

1. Preheat the heating block to 73 °C.

**a**

	-	x	x	x	x	x	x	x	x	x	x
DNAzyme [0.4 μM]	-	x	x	x	x	x	x	x	x	x	x
RNA [0.4 μM]	x	x	x	x	x	x	x	x	x	x	x
MgCl <sub>2</sub> [2 mM]	x	x	x	x	x	x	x	x	x	x	x
NaCl [1000 mM]	x	x	x	x	x	x	x	x	x	x	x
time [s]	2100	0	120	240	360	480	720	960	1440	1800	2100
RNA cleaved [%]	-	0	16	50	60	74	82	84	95	95	96

**b**

**Fig. 3** Gel-based activity assay. Concentration of Mg<sup>2+</sup>: 2 mM; temperature: 37 °C; 1000 mM NaCl; 0.4 μM RNA substrate; 0.4 μM DNAzyme. a: Denaturing gel visualizing the fluorescein-labeled RNA fragment. The amount of cleaved or uncleaved RNA was determined using a gel documentation system. b: The rate constant determined by fitting of data (black stars) to

$\frac{[\text{RNA}_{\text{cleaved}}]}{[\text{RNA}_{\text{total}}]} = 1 - \exp(-k_{\text{obs}} \cdot t)$  (blue line) is  $k_{\text{obs}} = 0.0028$ , which is very close to the rate constant determined in the FRET assay,  $k_{\text{obs}} = 0.0039$  (Fig. 2) with otherwise identical experimental conditions

2. Prepare a master mix DNAzyme:RNA solution—at least 99 μL for the ten samples—containing the following:

- 50 mM Tris–HCl pH 7.5
- 0.4 μM DNAzyme
- 0.4 μM RNA substrate
- 0.4 mM EDTA

1 M NaCl

3. Prepare the RNA control solution—at least 9  $\mu$ L—without the DNAzyme:
  - 50 mM Tris–HCl pH 7.5
  - 0.4  $\mu$ M RNA substrate
  - 0.4 mM EDTA
  - 1 M NaCl
4. Incubate the solutions for 5 min at 73 °C.
5. Spin down the solution briefly and incubate at room temperature for 10 min (*see Note 3*).
6. Aliquot the master mix DNAzyme:RNA solution to obtain 10 reaction tubes with each a volume of 9  $\mu$ L.
7. Add 1  $\mu$ L of the 20 mM MgCl<sub>2</sub> solution to each tube (including the tube that contains the RNA control solution). Spin down briefly and incubate at 37 °C.
8. After 0, 120, 240, 360, 480, 720, 960, 1440, 1800, and 2100 s, take one tube, put it on ice, and immediately quench the reaction by adding 10  $\mu$ L of the 2× RNA loading dye. Incubate the RNA control solution for 2100 s before adding the 2× RNA loading dye.
9. Assemble the gel apparatus.
10. Prepare the polyacrylamide solution and prepare the gel. Allow the gel to polymerize for 40 min (*see Note 7*).
11. Pre-run the gel for 20 min at 20 W (*see Note 7*).
12. Load the samples (20  $\mu$ L each) on the gel. Run the gel for 1.5 h at 20 W (*see Note 7*).
13. Visualize the fluorescein-labeled RNA fragments by fluorescence detection using a gel documentation system.

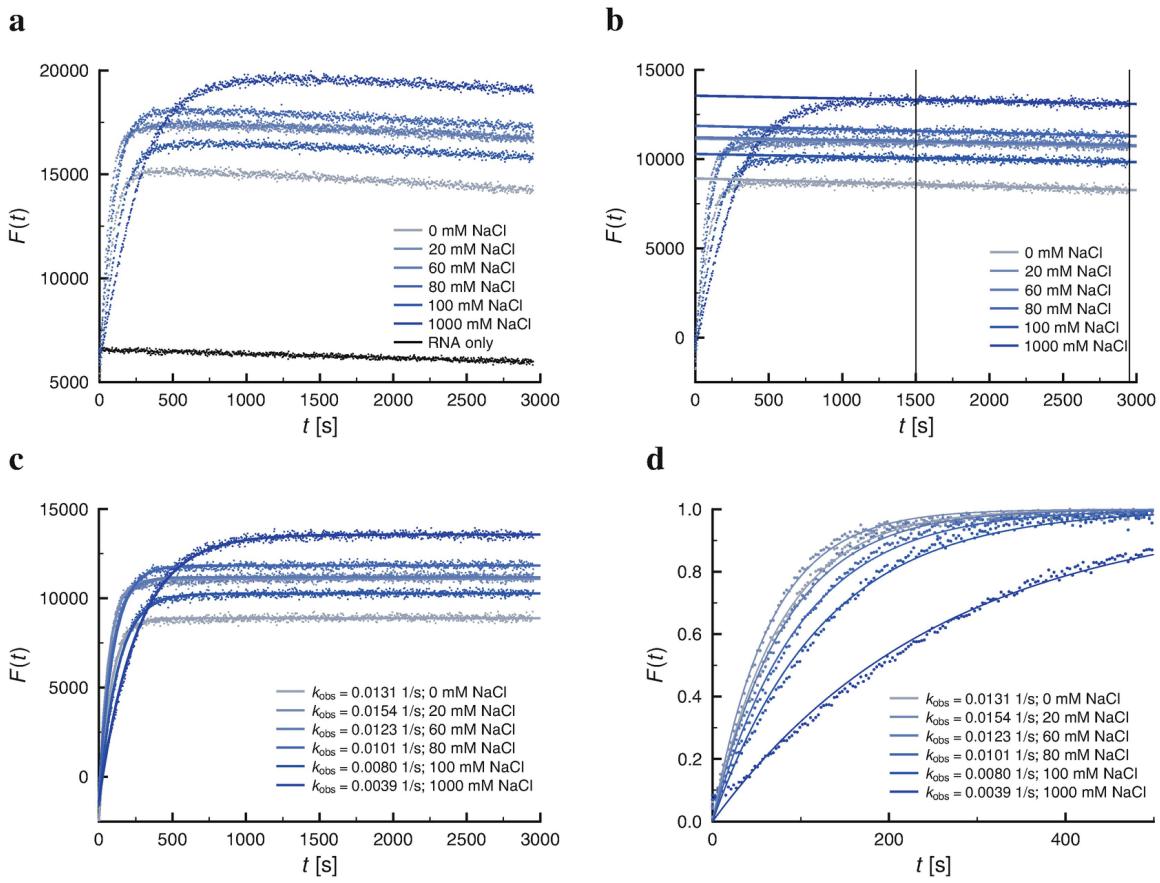
### 3.3 Data Analysis

The analysis is described here for a FRET-based assay (*see* Subheading 3.1) of DNAzyme cleavage activity in dependence on time for two different NaCl concentrations. Three measurements were performed for each NaCl concentration:

1. Store the fluorescence data of the activity assay in tabular form; that is, one column contains the time values and each other column contains fluorescence values in dependence on time.

The CSV (comma-separated values) file format is preferred, because it should be readable by most data analysis and graphics programs. Despite the file name, values have not to be separated by commas; other delimiters are also used quite often.

2. Plot the raw data (Fig. 4a). Take care to select sufficient ranges of time (x-axis) and fluorescence (y-axis) to show all data. Select different colors for the different data (*see* Note 8); here we chose shades of blue for data with the different NaCl concentrations and black for data of only RNA without DNAzyme.



**Fig. 4** Evaluation of FRET-based activity assay. Experimental conditions: 0.1  $\mu$ M RNA substrate, 0.1  $\mu$ M DNAzyme, 2 mM Mg<sup>2+</sup>, 37 °C (see Subheading 3.1). a: Raw FRET curves. Bluish dots show RNA fluorescence in the presence of different NaCl concentrations (see legend); black dots show RNA fluorescence in the absence of DNAzyme (RNA only). b: FRET data after subtraction of RNA-only data. The straight lines are from linear fits of data in the range  $1500 \leq t \leq 2952$  s (black vertical lines). The linear fit is thought to compensate for a fluorescence decrease with increasing time of light exposure. c: FRET data after subtraction of RNA-only data and the linear decrease in fluorescence. The lines are from fits of data to  $F(t) = F_{\max} (1 - \exp(-k_{\text{obs}} \cdot t)) + F_{\min}$ . Resulting  $k_{\text{obs}}$  values are given in the legend. d: FRET data after normalization to the respective  $F_{\min}$  and  $F_{\max}$  values from c. The length of the time axis was shortened to 500 s to show more clearly the quality of the fit. This figure is identical to Fig. 2a

3. Subtract the negative control data from the positive data (Fig. 4b). In the example data, the negative control data are from fluorescence-labeled RNA without DNAzyme. Correct the y-axis range accordingly. In our hands, the negative control results in a straight line with some

noise. If so, this subtraction can be omitted.

4. Correct the data for the usual fluorescence decrease with increasing time of light exposure. That is, select an appropriate range of the data and fit the range by a linear least squares fit  $F_{\text{decrease}}(t) = \text{slope} \cdot t + \text{offset}$ . Draw the corresponding straight lines to verify the correct range selection (Fig. 4b).
5. Subtract the straight lines from the data (Fig. 4c). Correct the y-axis range accordingly.
6. The corrected data should conform to an irreversible first-order reaction; that is, the reaction rate depends linearly on the RNA concentration

$$-\frac{d [\text{RNA}]}{dt} = \frac{d [\text{RNA}_{\text{cleaved}}]}{dt} = k_{\text{obs}} \cdot [\text{RNA}](t).$$

Integration and taking into account that the concentration of cleaved RNA is directly proportional to the measured fluorescence results in

$$F(t) = F_{\text{max}} (1 - \exp(-k_{\text{obs}} \cdot t)) + F_{\text{min}}. \quad (1)$$

Fit the corrected data to Eq. 1, and plot the values from the fitted equation (Fig. 4c).

7. Normalize the data using  $F_{\text{max}}$  and  $F_{\text{min}}$  from Eq. (1). Thus, plot the normalized experimental fluorescence data

$$F_{\text{norm}}(t) = \frac{F(t) - F_{\text{min}}}{F_{\text{max}}}$$

and the fitted values

$$F_{\text{norm}}(t) = 1 - \exp(-k_{\text{obs}} \cdot t)$$

(Fig. 4d).

---

## 4 Notes

1. Of course, also other fluorophore and quencher molecules are possible.

Select the FRET pair so that the absorption spectrum of the quencher overlaps with the emission spectrum of the reporter/fluorophore. For the chosen pair FAM:BHQ-1 also static quenching needs to be considered. FAM and BHQ-1 can bind together to form a non-fluorescent intramolecular dimer. Both FRET quenching and static quenching can occur together [22].

2. This RNA-only sample is important to control the stability of the RNA target during the measurement. The data obtained from this sample further serve as background signal in the data analysis.
3. This step allows the DNAzyme and the RNA target to anneal and to form a cleavage competent state. With this setup, it is possible to focus on the cleavage reaction without any effects of the association step.
4. The incubation is required to prevent fluctuations in the fluorescence signal due to temperature inhomogeneity.
5. The settings need to be adjusted for different experimental setups. The described settings are optimized for our experimental parameters, including volume, FRET pair, and the nucleic acid concentration. It is wise to adjust the focal height prior to each measurement.
6. The injected volume has to be large enough, and the pump speed has to be fast enough to mix the solutions properly.
7. We used a gel of 20 cm × 21 cm size.
8. Take care to choose colors that allow a reader to distinguish between the different colors. Complementary colors like red and green have high contrast for people with normal eye sight but fail for color-blind people. For further details and a color check of figures, see [23–25].

## Acknowledgements

We thank Dr. Julian Victor for the initial design of the experimental setup and Prof. Dr. Ingrid Span for support. This work was supported by the Chemical Industry Fund (Hoe 700080 to H.R.) and the German Academic Scholarship Foundation (to H.R.).

---

## References

1. Santoro SW, Joyce GF (1997) A general purpose RNA-cleaving DNA enzyme. *Proc Natl Acad Sci U S A* 94(9):4262–4266. <https://doi.org/10.1073/pnas.94.9.4262> [Crossref]
2. Victor J, Steger G, Riesner D (2018) Inability of DNAzymes to cleave RNA in vivo is due to limited Mg<sup>2+</sup> concentration in cells. *Eur Biophys J* 47(4):333–343. <https://doi.org/10.1007/s00249-017-1270-2> [Crossref]
3. Rosenbach H, Borggräfe J, Victor J, Wuebben C, Schiemann O, Hoyer W, Steger G, Etzkorn M, Span I (2020) Influence of monovalent metal ions on metal binding and catalytic activity of the 10–23 DNAzyme. *Biol Chem* 402:99–111. <https://doi.org/10.1515/hsz-2020-0207> [Crossref]
4. Santoro SW, Joyce GF (1998) Mechanism and utility of an RNA-cleaving DNA enzyme. *Biochemistry* 37(38):13330–13342. <https://doi.org/10.1021/bi9812221> [Crossref]
5. Mansfield ES, Worley JM, McKenzie SE, Surrey S, Rappaport E, Fortina P (1995) Nucleic acid detection using non-radioactive labelling methods. *Mol Cell Probes* 9(3):145–156. <https://doi.org/10.1006/mcpr.1995.0023> [Crossref]
6. Li J, Lu Y (2000) A highly sensitive and selective catalytic DNA biosensor for lead ions. *J Am Chem Soc* 122(42):10466–10467. <https://doi.org/10.1021/ja0021316> [Crossref]
7. Fan X, Li H, Zhao J, Lin F, Zhang L, Zhang Y, Yao S (2012) A novel label-free fluorescent sensor for the detection of potassium ion based on DNAzyme. *Talanta* 89:57–62. <https://doi.org/10.1016/j.talanta.2011.11.056> [Crossref]
8. Torabi SF, Wu P, McGhee CE, Chen L, Hwang K, Zheng N, Cheng J, Lu Y (2015) In vitro selection of a sodium-specific DNAzyme and its application in intracellular sensing. *Proc Natl Acad Sci U S A* 112(19):5903–5908. <https://doi.org/10.1073/pnas.1420361112> [Crossref]
9. Shen L, Chen Z, Li Y, He S, Xie S, Xu X, Liang Z, Meng X, Li Q, Zhu Z et al (2008) Electrochemical DNAzyme sensor for lead based on amplification of DNA–Au bio-

- bar codes. *Anal Chem* 80(16):6323–6328. <https://doi.org/10.1021/ac800601y> [Crossref]
10. Tian Y, Wang Y, Xu Y, Liu Y, Li D, Fan C (2015) A highly sensitive chemiluminescence sensor for detecting mercury (II) ions: a combination of exonuclease III-aided signal amplification and graphene oxide-assisted background reduction. *Sci China Chem* 58(3):514–518. <https://doi.org/10.1007/s11426-014-5258-9> [Crossref]
  11. Kumar S, Jain S, Dilbaghi N, Ahluwalia AS, Hassan AA, Kim KH (2019) Advanced selection methodologies for DNAzymes in sensing and healthcare applications. *Trends Biochem Sci* 44(3):190–213. <https://doi.org/10.1016/j.tibs.2018.11.001> [Crossref]
  12. Kahan-Hanum M, Douek Y, Adar R, Shapiro E (2013) A library of programmable DNAzymes that operate in a cellular environment. *Sci Rep* 3(1):1–6. <https://doi.org/10.1038/srep01535> [Crossref]
  13. Liu J, Lu Y (2003) Improving fluorescent DNAzyme biosensors by combining inter- and intramolecular quenchers. *Anal Chem* 75(23):6666–6672. <https://doi.org/10.1021/ac034924r> [Crossref]
  14. Zhang XB, Wang Z, Xing H, Xiang Y, Lu Y (2010) Catalytic and molecular beacons for amplified detection of metal ions and organic molecules with high sensitivity. *Anal Chem* 82(12):5005–5011. <https://doi.org/10.1021/ac1009047> [Crossref]
  15. Chiuman W, Li Y (2007) Efficient signaling platforms built from a small catalytic DNA and doubly labeled fluorogenic substrates. *Nucleic Acids Res* 35(2):401–405. <https://doi.org/10.1093/nar/gkl1056> [Crossref]
  16. Marras S, Kramer F, Tyagi S (2002) Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes. *Nucleic Acids Res* 30:e122. <https://doi.org/10.1093/nar/gnf121> [Crossref]
  17. Wikipedia (2020) List of information graphics software. [https://en.wikipedia.org/wiki/List\\_of\\_information\\_graphics\\_software](https://en.wikipedia.org/wiki/List_of_information_graphics_software). Accessed at November 18, 2021
  18. Newville M, Otten R, Nelson A, Ingargiola A, Stensitzki T, Allan D, Fox A, Carter F, Michat, Pustakhod D, Ram Y, Glenn, Deil C, Stuermer, Beelen A, Frost O, Zobrist N,

Pasquevich G, Hansen ALR, Spillane T, Caldwell S, Polloreno A, andrewhannum, Zimmermann J, Borreguero J, Fraine J, deep 42-thought, Maier BF, Gamari B, Almarza A (2020) lmfit/lmfit-py 1.0.1. <https://zenodo.org/record/3814709#.YA04v9bTUIY>. Accessed at November 18, 2021

19. Hunter JD (2019) Matplotlib v3.3.3. <https://matplotlib.org/>. Accessed at November 18, 2021
20. Hunter JD (2007) Matplotlib: a 2D graphics environment. Comput Sci Eng 9(3):90–95. <https://doi.org/10.1109/MCSE.2007.55>  
[Crossref]
21. Pugmire C, Rohde A, Struyf A, Abbott L, LaBella V, Mundt S (1990) Graphics Layout Engine (GLE). <https://glx.sourceforge.io/index.html>. Accessed at November 18, 2021
22. Johansson M, Fidder H, Dick D, Cook R (2002) Intramolecular dimers: a new strategy to fluorescence quenching in dual-labeled oligonucleotide probes. J Am Chem Soc 124:6950–6956. <https://doi.org/10.1021/ja025678o>  
[Crossref]
23. Dougherty B, Wade A (2015) Vischeck. <http://www.vischeck.com/>. Accessed at 19 Jan 2021
24. Ito K, Okabe M (2002) How to make figures and presentations that are friendly to color blind people. [jfly.iam.u-tokyo.ac.jp/html/manuals/pdf/color\\_blind.pdf](http://jfly.iam.u-tokyo.ac.jp/html/manuals/pdf/color_blind.pdf). Accessed at 19 Jan 2021
25. Wickline M (2001) Coblis – color blindness simulator. <https://www.color-blindness.com/coblis-color-blindness-simulator/>. Accessed at 19 Jan 2021

## 6. Stability and Activity of the 10–23 DNAzyme Under Molecular Crowding Conditions

Nina Kirchgässler<sup>1</sup>, Hannah Rosenbach<sup>1</sup> and Ingrid Span<sup>1</sup> 

(1) Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany

 **Ingrid Span**  
Email: [ingrid.span@hhu.de](mailto:ingrid.span@hhu.de)

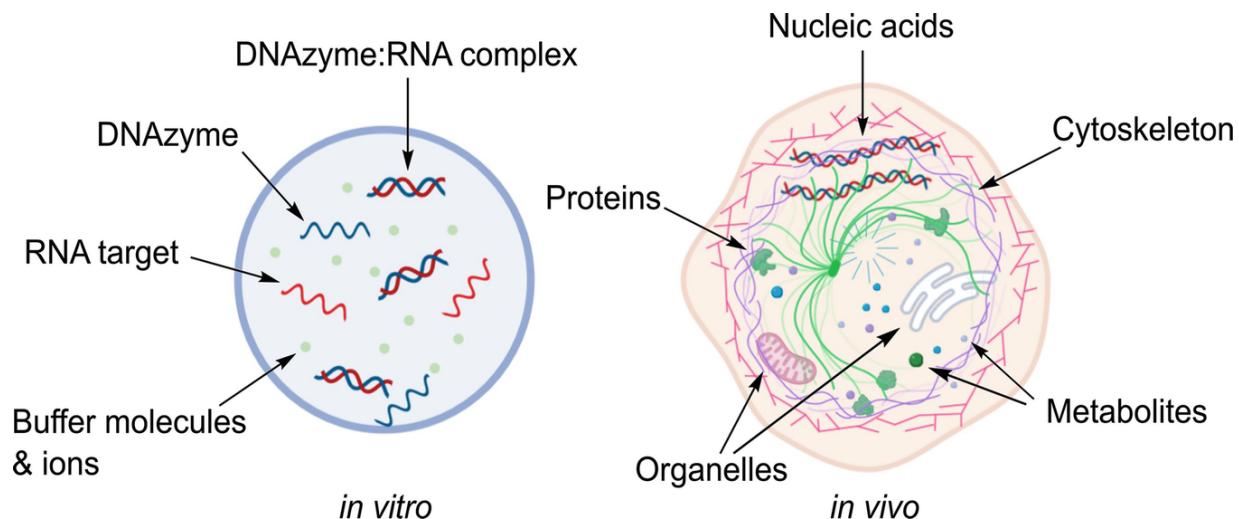
### Abstract

DNAzymes are biocatalysts that have been selected in vitro and their function inside cells (in vivo) is extremely low. Thus, almost all studies have been carried out in diluted solutions (in vitro). The cellular presence of molecules such as amino acids, polypeptides, alcohols, and sugars introduces forces that modify the kinetics and thermodynamics of DNAzyme-mediated catalysis. The crowded intracellular environment referred to as molecular crowding can be mimicked by adding high concentrations of natural or synthetic macromolecules to the reaction conditions. Here, we investigate the activity of the 10–23 DNAzyme and the stability of the DNAzyme:RNA complex under molecular crowding conditions. Therefore, we use a Förster resonance energy transfer (FRET)-based activity assay in combination with denaturing urea polyacrylamide gel electrophoresis and circular dichroism (CD) spectroscopy.

**Key words** 10–23 DNAzyme – Deoxyribozymes – Circular dichroism (CD) spectroscopy – Denaturation temperature – Förster resonance energy

## 1 Introduction

Deoxyribozymes (DNAzymes) are single-stranded, catalytically active DNA sequences. They were isolated using *in vitro* selection and they are capable of accelerating a broad spectrum of chemical reactions [1–3]. Although DNAzymes are highly active *in vitro*, only poor *in vivo* activity has been reported so far [4, 5]. The conditions in biochemical experiments (*in vitro*) and inside cells (*in vivo*) vary strongly (Fig. 1) and could be the reason for the discrepancy between *in vitro* activity and performance in cells. While *in vitro* experiments are performed in highly diluted solutions and under strictly controlled conditions, the *in vivo* environment is dominated by a diversity of macromolecules such as nucleic acids, proteins, organelles, osmolytes, and metabolites, creating crowded and heterogeneous conditions [6, 7]. The concentration of cellular molecules (ranging from 50 to 400 g/L) not only allows for crosstalk between them, but also affects solution properties, increases volume exclusion effects, unspecific interactions, and steric repulsion [8, 9]. The crowded intracellular environment referred to as molecular crowding affects the structure of biomolecules, their interactions, thermodynamics and kinetics of chemical reactions [10, 11].



**Fig. 1** Schematic illustration of the conditions of biochemical experiments (*in vitro*) and inside the cell (*in vivo*). *In vitro* experiments (left panel) are carried out in diluted, homogeneous solutions under controlled conditions. The environment *in vivo* (right panel)

is highly condensed due to the presence of cellular organelles and a diversity of macromolecules, such as proteins, sugars, osmolytes, and nucleic acids, with concentrations between 50 and 400 g/L.

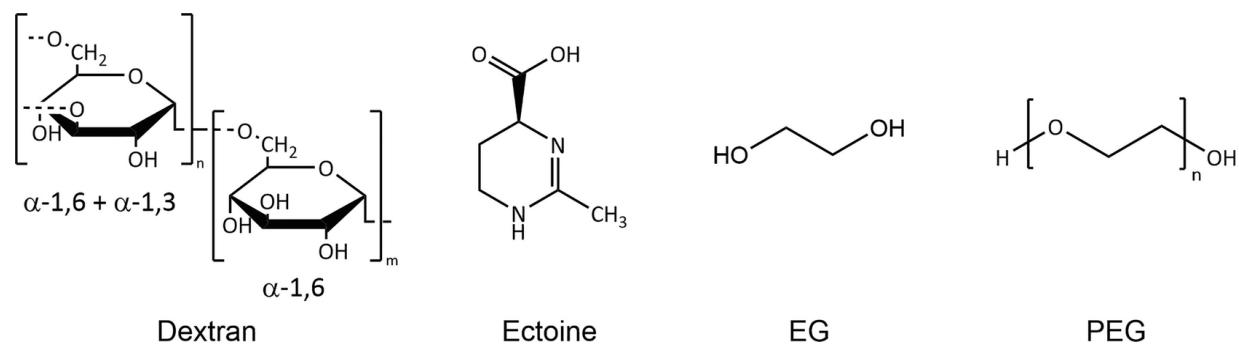
The addition of high concentrations of natural and synthetic macromolecules, such as dextran and polyethylene glycol (PEG), to the reaction mixture enables molecular crowding to be mimicked *in vitro* [11–14]. In general, agents preferably used for molecular crowding analysis *in vitro* provide a high solubility in water, do not cause a degradation or precipitation of the molecule of interest and do not show higher binding affinities than water [15, 16]. The activity of DNAzymes and the stability of the DNAzyme:RNA complex under such conditions *in vitro* can be analyzed, e.g., by fluorescence-based activity assays and circular dichroism (CD) spectroscopy. Activity assays using fluorescence-labeled RNA targets in combination with denaturing polyacrylamide gel electrophoresis (PAGE) allow to distinguish between uncleaved and cleaved molecules, and potentially also degradation products. Addition of 6 to 8 M of urea to the polyacrylamide gel enables the denaturation of secondary and tertiary DNA and RNA structures, thus providing nucleic acid separation by size. This setup can be used to analyze the activity and specificity of DNAzymes as well as the stability of its RNA target. The resolution of this method is very high and enables the detection of length differences of one nucleotide [17, 18].

DNAzyme activity can also be measured by Förster resonance energy transfer (FRET) assays. These measurements require a double-labeled RNA target containing a fluorophore at one end and a quencher molecule at the other end. Here, metal ion-dependent DNAzyme-induced RNA cleavage leads to an increase in fluorescence allowing for time-resolved measurements and calculation of the observed rate constant  $k_{\text{obs}}$  [19].

CD spectroscopy is a very sensitive method to obtain information on the secondary structure and to determine the denaturation temperature of biomolecules. It exploits the chirality of molecules and the resulting difference in the absorption of right and left polarized light (ellipticity) [20]. The axial chirality of nucleic acids can be used to analyze structural conformation [21]. An additional implementation of temperature gradients allows an analysis of the denaturation behavior, since denaturation of

nucleic acids results in an increase in absorbance close to 260 nm with increasing temperature [22, 23].

Here, we use different high molecular weight polymers and an osmolyte (Fig. 2) to simulate different degrees of volume exclusion effects. Hereby, we selected polymers of high molecular weight ranging from 3 to 20 kDa, as well as the polysaccharide dextran T500 (500 kDa). In addition, DNAzyme activity has been monitored in the presence of ethylene glycol (EG) and the osmolyte ectoine covering low molecular substances that also define the cellular environment and change solution properties [16]. Ectoine naturally accumulates in the cytoplasm of halophilic microorganisms aiming to protect cells from osmotic stress and to stabilize biomolecules [24, 25].



**Fig. 2** Chemical structures of agents commonly used to mimic molecular crowding conditions

Here, we provide a protocol for the analysis of the catalytic activity of the 10–23 DNAzyme and the denaturation behavior of the DNAzyme:RNA complex under different molecular crowding concentrations using a variety of crowding agents.

## 2 Materials

Prepare all solutions using ultrapure water (obtained with a resistivity of 18.2 MΩ cm at 25 °C). Unless indicated otherwise, prepare and store all reagents at room temperature.

### 2.1 Analysis of Construct Stability and Specificity

1. Reaction buffer: 500 mM Tris(hydroxymethyl)aminomethane (Tris)–HCl pH 7.5 (see Notes 1 and 2).

2. 4 M NaCl in water.
3. 8 mM Ethylenediaminetetraacetic acid (EDTA) in water.
4. 10 mM MgCl<sub>2</sub> in water.
5. Thermocycler.
6. Tabletop centrifuge.
7. Crowding agents.
  - 50% (w/v) PEG 3000, 3350, 4000, 6000, 8000, and 20,000 in Tris–HCl pH 7.5.
  - 50% (w/v) EG in Tris–HCl pH 7.5.
  - 50% (w/v) Ectoine in Tris–HCl pH 7.5.
  - 50% (w/v) Dextran T500 in Tris–HCl pH 7.5.
8. DNAzyme and RNA targets.
  - 100 µM 10–23 DNAzyme (Dz) in water.
  - 100 µM simply labeled RNA target, containing 6-carboxyfluorescein (6-FAM) at the 5'-end (T\_6-FAM), in water.
  - 100 µM dual labeled RNA target, containing 6-FAM at the 5'-end and Black Hole Quencher 1 (BHQ-1) at the 3'-end (T\_FRET), in water.
  - 100 µM unlabeled RNA target (T) in water.
  - 100 µM stabilized RNA target (T\_2'F) in water; the 2' hydroxyl group (2'-OH) of the unpaired nucleotide is substituted by a fluorine atom to prevent RNA cleavage.All oligonucleotides can be stored at –20 °C (*see Notes 3 and 4*).

## 2.2 Urea Polyacrylamide Gel (18%)

1. 10× Tris-borate-EDTA (TBE) buffer: 0.89 M Tris, 0.89 M boric acid,

20 mM EDTA pH 8.0 (autoclaved). Prepare 1 L solution. Use water for dilution to 1× TBE buffer.

2. 40% acrylamide/bis-acrylamide solution (19:1) (store at 4 °C).
3. Urea (Molecular Biology Grade).
4. 10% Ammonium persulfate (APS) in water (store at –20 °C).
5. *N,N,N',N'-Tetramethyl-ethylenediamine* (TEMED) (store at 4 °C).
6. 2× RNA gel loading dye: 9.4 mL formamide, 500 mM EDTA, 2 mg bromophenol blue, and 2 mg xylene cyanol (store at –20 °C).
7. Casting frame and glass plates (inner and outer plate) to obtain gels, two spacer stripes, and a 14-well comb (*see Note 5*).
8. Vertical electrophoresis cell and power supply.
9. Metal plate and clamps (*see Note 6*).
10. Nucleic acid gel staining agent.
11. Rocker.
12. Gel documentation system.

## 2.3 CD Spectroscopy

1. Reaction buffer: 500 mM Tris–HCl pH 7.5.
2. 4 M NaCl in water.
3. 8 mM EDTA in water.
4. 100 µM Dz in water (*see Note 4*).

5. 100  $\mu$ M RNA target (T or T\_2'F) in water (*see Note 4*).
  6. Thermocycler.
  7. Tabletop centrifuge.
  8. Cuvette: quartz cuvette, 1 mm path length, for use with CD spectrometer.
  9. CD spectrometer with a temperature control system, suitable to run temperature gradients.
- 

### 3 Methods

#### 3.1 Activity Assay for Stability Analysis

A control sample consisting of only RNA and a sample consisting of the DNAzyme:RNA complex have been analyzed for each condition.

1. Set up a reaction mixture in 50 mM Tris–HCl pH 7.5 with a final volume of 10  $\mu$ L in a 1.5 mL reaction tube and add 0.4  $\mu$ M Dz, 0.4  $\mu$ M RNA target (T\_6-FAM or T), and a crowding agent (*see Note 7*). Mix and incubate for 5 min at 73 °C (*see Notes 8 and 9*).
2. Centrifuge the sample for 4 s at 2000  $\times g$  to ensure that the entire solution is at the bottom of the tube and incubate for 10 min at room temperature.
3. Add 1  $\mu$ L of 10 mM MgCl<sub>2</sub> to the reaction mixture and incubate for 3 h at 37 °C (*see Note 10*).
4. Centrifuge the sample for 4 s at 2000  $\times g$  and add 10  $\mu$ L of 2× RNA gel loading dye. Mix and incubate for 10 min at 95 °C, then centrifuge again for 4 s at 2000  $\times g$ .

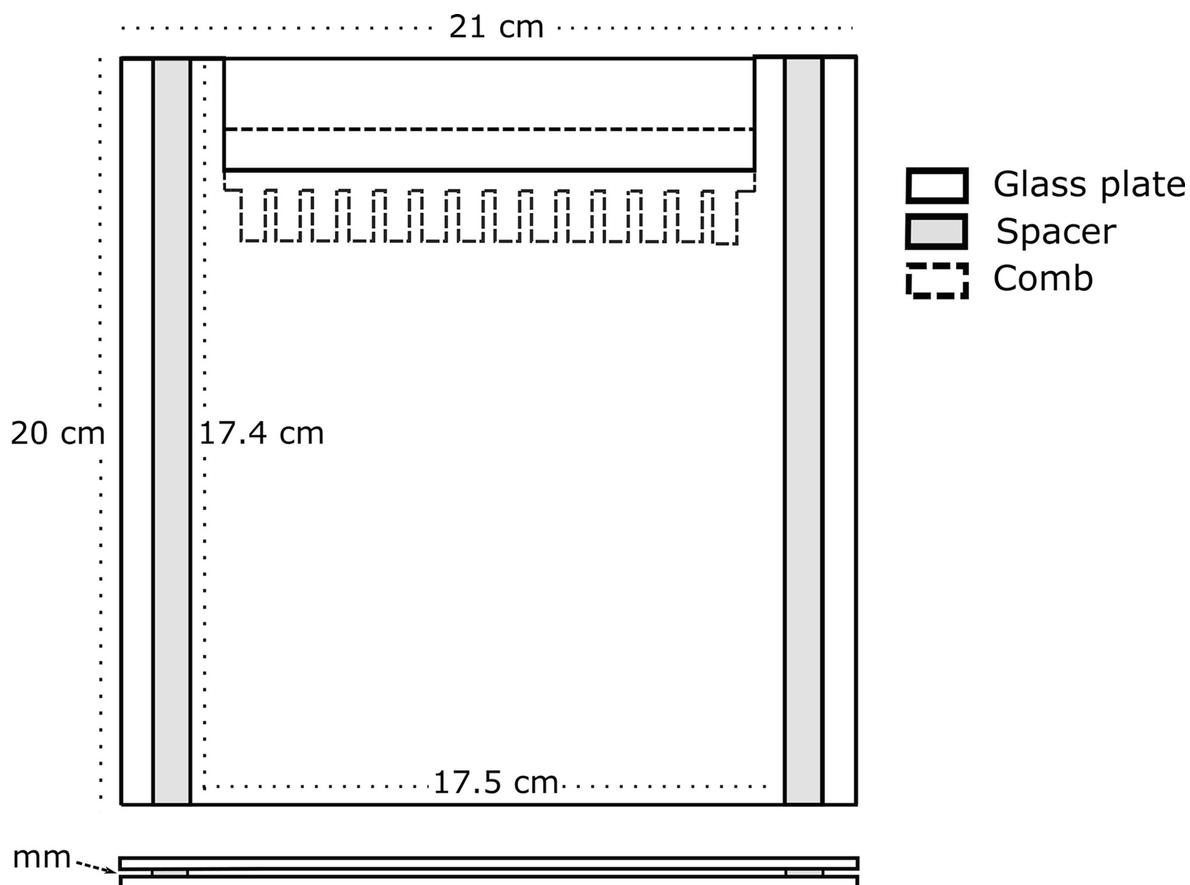
#### 3.2 18% Urea-PAGE

1. During the 3 h incubation (*see Subheading 3.1, step 3*), prepare the

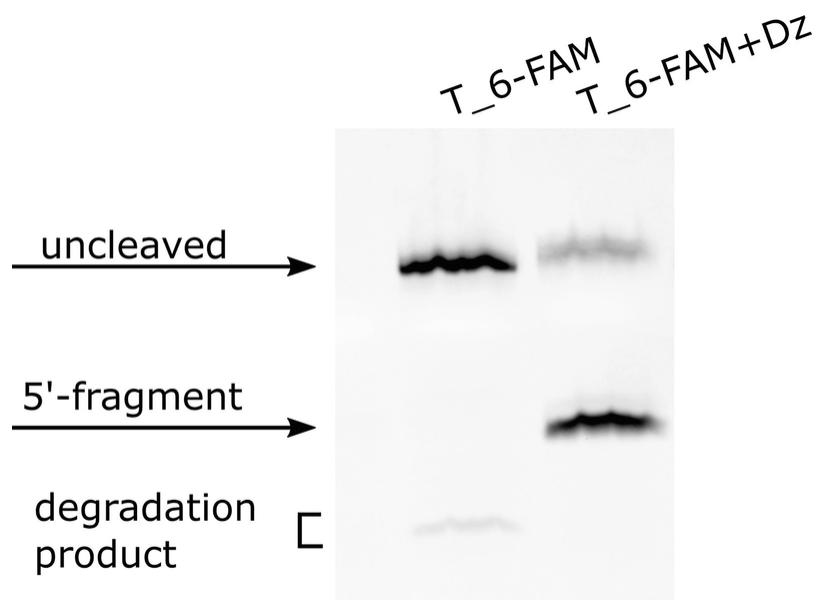
18% polyacrylamide gel. First, put a stirring bar in a 50 mL glass cylinder. Weigh 7 M of urea (12.61 g) and transfer it to the cylinder. Add 13.5 mL of the 40% acrylamide/bis-acrylamide solution (19:1) and 3 mL 10× TBE buffer. Heat the reaction mixture for a few seconds in a microwave until it becomes slightly warm. Seal the cylinder and stir until the urea is completely dissolved. Add water to a total volume of 30 mL. Add 30 µL (0.1%, v/v) TEMED and 300 µL (0.1%, w/v) APS, then mix shortly. Transfer the solution to a prepared casting frame and insert a 14-well comb immediately without introducing air bubbles (Fig. 3). Wait until the gel is polymerized (approximately 1 h at room temperature).

2. After polymerization, transfer the gel to a vertical electrophoresis system. Place a metal plate in front of the gel and fix the setup using clamps.
3. Fill the buffer reservoirs with 1× TBE buffer without allowing the buffer to touch the metal plate.
4. Remove the comb and flush the wells with 1× TBE buffer.
5. Pre-run the gel for 30 min at 20 W to reduce temperature inhomogeneity throughout the gel. Flush the wells again with buffer prior to sample loading.
6. Add 20 µL of sample per gel lane (*see Note 11*).
7. Run the gel for 50 min at 20 W until the dye bromophenol blue has passed the first third of the gel.
8. For experiments with samples containing the fluorescence-labeled RNA target T\_6-FAM, place the gel directly (do not remove the glass plates, since it poses the risk of damaging the gel) into the gel documentation system. Fluorescence bands become visible by excitation with blue light at 495 nm (Fig. 4). For experiments with non-fluorescent targets (T or T\_2'F, used for CD spectroscopy), carefully remove the gel from the glass plates with the help of a spatula or a similar tool. Carefully transfer the gel to a tray filled with 150 mL 1× TBE buffer and 50 µL nucleic acid gel staining agent and cover the

tray. Incubate on a rocker for 1 h at room temperature. After 1 h, place the gel into a gel documentation system. The stained nucleic acids become visible by excitation with UV light (302 nm).



**Fig. 3** Schematic representation of an 18% urea polyacrylamide gel. Two glass plates of  $21 \times 20$  cm in size are sealed and held together by clamps. Two spacer stripes ( $1.2 \times 20$  cm) are placed between these plates to obtain gels with  $17.5 \times 17.4$  cm in size and 1 mm in thickness. The comb contains 14 wells, each well with a filling capacity of 20  $\mu\text{L}$



**Fig. 4** Example of an urea-PAGE with samples consisting of RNA (T\_6-FAM) and DNAzyme:RNA complex (T\_6-FAM + Dz). The samples were incubated for 3 h at 37 °C prior to PAGE analysis, gel electrophoresis was carried out for 50 min, and the nucleic acids were visualized using fluorescence detection of the 6-carboxyfluorescein dye. While DNAzyme-mediated RNA cleavage results in two cleavage products, only the 5'-labeled fragment can be detected

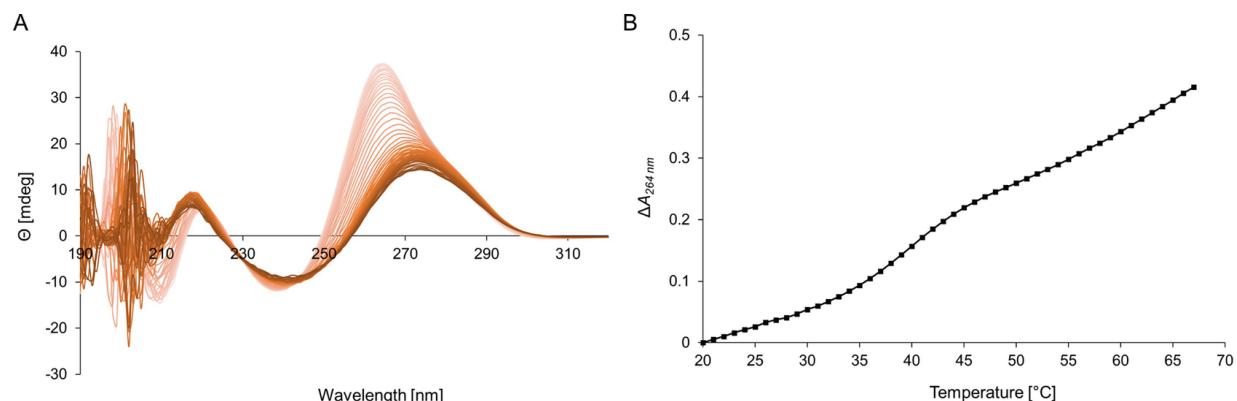
### 3.3 Determining the Denaturation Temperature Using CD Spectroscopy

All samples are prepared with a total volume of 250 µL.

1. Mix 25 µM Dz and 25 µM RNA target (T or T\_2'F) with 0.1 mM EDTA and incubate for 5 min at 73 °C (*see Note 8*).
2. Centrifuge the samples for 4 s at 2000 × g and incubate for 10 min at room temperature.
3. Transfer the entire sample into a cuvette (avoid air bubbles) and seal it (*see Notes 12 and 13*).
4. Record CD spectra from 190 nm to 320 nm in the temperature range of 20 °C to 73 °C. Record the sample's ellipticity, absorbance, and voltage applied to the photomultiplier. To provide a good signal to noise ratio at a complex concentration of 25 µM, perform measurements with ten accumulations. Set the scan rate to 100 nm/min with a heating rate of maximally 1 °C/min. Collect data each nm and

with a temperature interval of at least 1 °C.

5. Determine the denaturation temperature  $T_m$  of the DNAzyme:RNA complex in absence of molecular crowding agents by monitoring the CD signal at 264 nm. This wavelength was selected as the band in the CD spectrum (Fig. 5a) reveals the greatest impact of the temperature increase. Plot the absorbance difference at 264 nm ( $\Delta A_{264\text{nm}}$ ) with  $\Delta A = A_T - A_{20\text{ °C}}$  against the temperature (Fig. 5b).  $T_m$  refers to half height of the sigmoidal denaturation curve. For further analysis of denaturation curves, we refer to the literature [22, 23, 26].
6. After measurement, transfer the sample to a 1.5 mL reaction tube (store at –20 °C). For subsequent analysis by urea-PAGE (see Subheading 3.2), dilute the sample to a final concentration of 10 μM using 50 mM Tris pH 7.5. Use 10 μL of the sample and add 10 μL 2× RNA loading dye and follow the instructions (see Subheading 3.2).



**Fig. 5** Determining the denaturation temperature using CD spectroscopy. (a) CD spectra of the DNAzyme:RNA complex in absence of molecular crowding conditions collected in the temperature range from 20 to 73 °C. (b) Increase in absorbance at 264 nm ( $\Delta A_{264\text{nm}}$ ) with  $\Delta A = A_T - A_{20\text{ °C}}$  as a function of the temperature  $T$

## 4 Notes

1. Prior to buffer preparation, make sure to decontaminate all materials to minimize the risk of RNA degradation by RNases. All surfaces, including benchtops, pipettes, glassware, pH electrodes, and benchtop

instruments, should be treated with a commercially available RNase decontamination agent. Alternatively, glassware can be heated for at least 4 h at 180 °C prior to usage.

2. Reaction buffer (stock solution): 1 M Tris pH 7.5. Weigh 24.22 g Tris and transfer to a 400 mL glass beaker. Add 150 mL water to a 250 mL graduated cylinder and transfer to the glass beaker. Dissolve and transfer to the cylinder. Add water to a volume of 180 mL, mix, and adjust the pH with HCl. Fill up to 200 mL with water. For experimental approaches and dissolving of crowding agents, prepare dilutions of 500 mM and 50 mM Tris pH 7.5.
3. For activity assays, prepare one stock solution of 4 µM Dz and one stock solution of 4 µM RNA target. Use water to dilute the samples.
4. Store nucleic acid solutions on ice during the experiments and long-term at –20 °C.
5. Regarding gel preparation, glass plates of 21 × 20 cm were used to obtain gels of 17.5 × 17.4 cm in size and 1 mm in thickness, containing 30 mL of mixture.
6. The metal plate should cover two-thirds of the glass plate to ensure a consistent temperature throughout the gel.
7. The cellular environment was mimicked by using the following concentration ranges: 0 to 0.18 g/mL for all PEG variants and EG, 0 to 0.004 g/mL for dextran T500 and 0 to 0.15 g/mL for ectoine.
8. Mix the sample gently. **Attention:** with increasing PEG size the solution's viscosity increases notably. Adapt the speed of mixing to prevent air bubbles.
9. During sample preparation, consider the volume of MgCl<sub>2</sub> needed for catalysis. Metal cofactors are added immediately prior to the 3 h incubation at 37 °C.
10. Use a piece of aluminum foil or something similar to cover the samples, thus reducing photobleaching effects on the label's fluorescence.

11. The addition of 10 µL 2× RNA loading dye in the remaining lanes is optional.
  12. Prior to sample transfer, clean the CD cuvette with a special cleaning solution for optical equipment (in- and outside) and dry in a stream of nitrogen gas or air.
  13. CD spectroscopic measurements are performed without adding any metal cofactors to prevent RNA cleavage, thus enabling the analysis of the DNAzyme:RNA complex.
- 

## References

1. Xiang Y, Lu Y (2014) DNA as sensors and imaging agents for metal ions. Inorg Chem 53:1925–1942. <https://doi.org/10.1021/ic4019103>  
[Crossref][PubMed]
2. Baum DA, Silverman SK (2008) Deoxyribozymes: useful DNA catalysts in vitro and in vivo. Cell Mol Life Sci 65:2156–2174. <https://doi.org/10.1007/s00018-008-8029-y>  
[Crossref][PubMed][PubMedCentral]
3. Breaker RR, Joyce GF (1994) A DNA enzyme that cleaves RNA. Chem Biol 1:223–229. [https://doi.org/10.1016/1074-5521\(94\)90014-0](https://doi.org/10.1016/1074-5521(94)90014-0)  
[Crossref][PubMed]
4. Rosenbach H, Victor J, Etzkorn M, Steger G, Riesner D, Span I (2020) Molecular features and metal ions that influence 10-23 DNAzyme activity. Molecules 25:11–21. <https://doi.org/10.3390/molecules25133100>  
[Crossref]
5. Victor J, Steger G, Riesner D (2018) Inability of DNAzymes to cleave RNA in vivo is due to limited Mg<sup>2+</sup> concentration in cells. Eur Biophys J 47:333–343. <https://doi.org/10.1007/s00249-017-1270-2>  
[Crossref][PubMed]
6. Ellis RJ (2001) Macromolecular crowding: obvious but underappreciated. Trends Biochem Sci 26:597–604. [https://doi.org/10.1016/S0968-0004\(01\)01938-7](https://doi.org/10.1016/S0968-0004(01)01938-7)  
[Crossref][PubMed]
- 7.

- Zimmerman SB, Minton AP (1993) Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annu Rev Biophys Biomol Struct* 22:27–65  
[Crossref]
8. DasGupta S (2020) Molecular crowding and RNA catalysis. *Org Biomol Chem* 18:7724–7739. <https://doi.org/10.1039/d0ob01695k>  
[Crossref][PubMed]
  9. Zhou H-X, Rivas G, Minton AP (2008) Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu Rev Biophys* 37:375–397. <https://doi.org/10.1146/annurev.biophys.37.032807.125817>  
[Crossref][PubMed][PubMedCentral]
  10. Minton AP (2001) The influence of Macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *J Biol Chem* 276:10577–10580. <https://doi.org/10.1074/jbc.R100005200>  
[Crossref][PubMed]
  11. Paudel BP, Fiorini E, Börner R, Sigel RKO, Rueda DS (2018) Optimal molecular crowding accelerates group II intron folding and maximizes catalysis. *Proc Natl Acad Sci U S A* 115:11917–11922. <https://doi.org/10.1073/pnas.1806685115>  
[Crossref][PubMed][PubMedCentral]
  12. Paudel BP, Rueda D (2014) Molecular crowding accelerates ribozyme docking and catalysis. *J Am Chem Soc* 136:16700–16703. <https://doi.org/10.1021/ja5073146>  
[Crossref][PubMed][PubMedCentral]
  13. Azri A, Giamarchi P, Grohens Y, Olier R, Privat M (2012) Polyethylene glycol aggregates in water formed through hydrophobic helical structures. *J Colloid Interface Sci* 379:14–19. <https://doi.org/10.1016/j.jcis.2012.04.025>  
[Crossref][PubMed]
  14. Tyrrell J, Weeks KM, Pielak GJ (2015) Challenge of mimicking the influences of the cellular environment on RNA structure by PEG-induced Macromolecular crowding. *Biochemistry* 54:6447–6453. <https://doi.org/10.1021/acs.biochem.5b00767>  
[Crossref][PubMed]
  15. Fiorini E, Börner R, Sigel RKO (2015) Mimicking the in vivo environment - the effect of crowding on RNA and biomacromolecular folding and activity. *Chimia (Aarau)* 69:207–212. <https://doi.org/10.2533/chimia.2015.207>  
[Crossref]
  16. Nakano SI, Miyoshi D, Sugimoto N (2014) Effects of molecular crowding on the

- structures, interactions, and functions of nucleic acids. *Chem Rev* 114:2733–2758. <https://doi.org/10.1021/cr400113m>  
[Crossref][PubMed]
17. Summer H, Grämer R, Dröge P (2009) Denaturing urea polyacrylamide gel electrophoresis (urea PAGE). *J Vis Exp* 3–5. <https://doi.org/10.3791/1485>
  18. Albright LM, Slatko BE (2001) Denaturing polyacrylamide gel electrophoresis. *Curr Protoc Nucleic Acid Chem Appendix* 3:1–5. <https://doi.org/10.1002/0471142905.hga03fs00>  
[Crossref]
  19. Rosenbach H, Borggräfe J, Victor J, Wuebben C, Schiemann O (2021) Influence of monovalent metal ions on metal binding and catalytic activity of the 10–23 DNAzyme. *Biol Chem* 402:99–111. <https://doi.org/10.1515/hsz-2020-0207>  
[Crossref]
  20. Kypr J, Kejnovská I, Bednářová K, Vorlíčková M (2012) Circular dichroism spectroscopy of nucleic acids. *Compr Chiroptical Spectrosc* 2:575–586. <https://doi.org/10.1002/9781118120392.ch17>  
[Crossref]
  21. Kypr J, Kejnovská I, Renčík D, Vorlíčková M (2009) Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res* 37:1713–1725. <https://doi.org/10.1093/nar/gkp026>  
[Crossref][PubMed][PubMedCentral]
  22. Riesner D, Römer R (1973) Thermodynamics and kinetics of conformational transitions in oligonucleotides and tRNA. In: Duchesne J (ed) Structural studies on nucleic acids and other biopolymers. Academic Press, Cambridge, pp 237–318
  23. Schmitz M, Tinoco I (2000) Thermodynamics of formation of secondary structure in nucleic acids. In: Di Cera E (ed) Thermodynamics in biology. Oxford University Press, New York, pp 131–176
  24. Pastor JM, Salvador M, Argandoña M, Bernal V, Reina-Bueno M, Csonka LN, Iborra JL, Vargas C, Nieto JJ, Cánovas M (2010) Ectoines in cell stress protection: uses and biotechnological production. *Biotechnol Adv* 28:782–801. <https://doi.org/10.1016/j.biotechadv.2010.06.005>  
[Crossref][PubMed]
  25. Czech L, Hermann L, Stöveken N, Richter AA, Höppner A, Smits SHJ, Heider J, Bremer E (2018) Role of the Extremolytes Ectoine and Hydroxyectoine as stress

- protectants and nutrients: genetics, Phylogenomics, biochemistry, and structural analysis. *Genes (Basel)* 9:1–58. <https://doi.org/10.3390/genes9040177>  
[Crossref]
26. Serra MJ, Turner DH (1995) Predicting thermodynamic properties of RNA. *Methods Enzymol* 259:242–261. [https://doi.org/10.1016/0076-6879\(95\)59047-1](https://doi.org/10.1016/0076-6879(95)59047-1)  
[Crossref][PubMed]

## 7. Quantifying the Number and Affinity of Mn<sup>2+</sup>-Binding Sites with EPR Spectroscopy

Christine Wuebben<sup>1</sup> and Olav Schiemann<sup>1</sup>✉

(1) Institute of Physical and Theoretical Chemistry, Rheinische Friedrich-Wilhelms-University, Bonn, Germany

✉ Olav Schiemann  
Email: [schiemann@pc.uni-bonn.de](mailto:schiemann@pc.uni-bonn.de)

### Abstract

During the last decades, various functional oligonucleotides have been discovered including DNAzymes, ribozymes, and riboswitches. Their function is based on their ability to form and change their three-dimensional structure. Binding of divalent ions to specific binding pockets was found to be important for the global structure and function. Here, we present a protocol that allows counting the number of Mn<sup>2+</sup>-binding sites and to determine their dissociation constants by means of continuous wave X-band Electron Paramagnetic Resonance (EPR) spectroscopy. In this method, Mn<sup>2+</sup> is titrated into the oligonucleotide-containing sample and the intensity of the EPR spectrum is recorded. By comparison with a Mn<sup>2+</sup>-only reference sample, the binding isotherm can be constructed and fitted to binding models yielding the number and affinities of the binding sites. This method has been successfully applied to several functional oligonucleotides.

**Key words** Binding sites – Binding isotherms – Mn<sup>2+</sup> titration – Metal ions – EPR spectroscopy

---

## 1 Introduction

The interaction between metal ions and nucleic acids is highly important for the structure and function of oligonucleotides. On the one hand, metal ions are involved in the charge neutralization of the polyanionic backbone of oligonucleotides. On the other hand, some metal ions and in particular divalent metal ions like Mg<sup>2+</sup> are bound with high affinity at specific sites like metal ions in metalloproteins. These metal ions have been shown to be necessary for the correct folding and the function of the oligonucleotide. For example, for some ribozymes it is a high-affinity Mg<sup>2+</sup> ion at which the phosphodiester bonds are broken and formed. It is thus essential to have methods at hand that can provide insights into metal ion binding and that complement each other [1–6].

High resolution X-ray structures are excellent in revealing tightly bound metal ions but require growing crystals [7, 8]. Investigations in liquid solution can be done, e.g., by nuclear magnetic resonance (NMR) spectroscopy, which provides information about the dynamics of the nucleic acids and metal ion binding sites can be predicted. However, NMR is limited to sizes below ~100 kDa due to spectral overlap [6, 9, 10]. By contrast, metal ion-induced hydrolytic cleavage is size unrestricted and yields information on metal ions catalyzing phosphodiester cleavage, but requires <sup>32</sup>P-end-labeling [6, 10, 11]. Another solution method is isothermal titration calorimetry (ITC), which permits measuring thermodynamic parameters, affinity constants, and binding stoichiometries. However, high sample amounts are required [6, 10, 12]. Here, we present an in-solution method that is based on continuous wave (cw) X-band Electron Paramagnetic Resonance (EPR) spectroscopy [13–15]. Since the spin of unpaired electrons is detected in EPR, the binding of diamagnetic divalent ions like, e.g., Mg<sup>2+</sup> cannot be monitored directly. However, many studies have shown that Mg<sup>2+</sup> can be exchanged for the paramagnetic ion Mn<sup>2+</sup>, because they have similar physical properties like radius, charge, and hydration number. But one should keep in mind that Mg<sup>2+</sup> is a “hard ion” with a Pearson hardness of 32 and prefers indirect binding to nucleic acids

via hydrogen-bonded water molecules. In contrast, Mn<sup>2+</sup> has a Pearson hardness of 9 and is found to directly bind to a nitrogen atom of a nucleobase meaning the binding affinity of Mn<sup>2+</sup> is usually higher than that of Mg<sup>2+</sup> [10]. Thus, cwEPR measurements can be used to learn something about the binding of Mn<sup>2+</sup> itself or about Mg<sup>2+</sup>-binding sites with the limitation that this information was gained on the similar but not identical Mn<sup>2+</sup>.

In practice, Mn<sup>2+</sup> is titrated into the oligonucleotide-containing sample and the intensity of the Mn<sup>2+</sup> EPR spectrum is recorded. In an aqueous solution with a pH of roughly 7, Mn<sup>2+</sup> will predominately form the highly symmetric [Mn(H<sub>2</sub>O)<sub>6</sub>]<sup>2+</sup> ion, which gives rise to a cw X-band EPR spectrum of roughly six lines due to the coupling of the electron spin S = 5/2 to the nuclear spin I(<sup>55</sup>Mn) = 5/2. Binding to a biomolecule induces a ligand field asymmetry and the comparatively large size of the biomolecule reduces the molecular tumbling rate, which both lead to a strong line broadening such that the EPR signal of bound Mn<sup>2+</sup> is not detectable under these conditions [16, 17]. Thus, comparing the obtained EPR signal intensity with that of an aqueous Mn<sup>2+</sup>-only reference sample enables one to quantify free and bound Mn<sup>2+</sup> and to construct the binding isotherm. The binding isotherm can then be fitted to binding models yielding the number and affinity of binding sites. The method is size unrestricted with respect to the RNA and independent of diamagnetic salts, and has been successfully applied to quantify Mn<sup>2+</sup>-binding sites and affinities in various oligonucleotides including ribozymes and DNAzymes [18–27]. Finally, if one would like to resolve the structure of the binding site one freezes the sample to cryogenic temperatures under which condition an EPR signal can be recorded and one can then apply a large range of cw and pulsed hyperfine EPR spectroscopic methods, which provide structural information about 10 Å around the Mn<sup>2+</sup> ion [23, 24, 27–32].

---

## 2 Materials

### 2.1 EPR Spectrometer

1. CW X-band EPR spectrometer, e.g., EMXmicro EPR spectrometer

(Bruker).

2. Resonator, e.g., ER 4103TM Cylindrical Mode Resonator (Bruker).

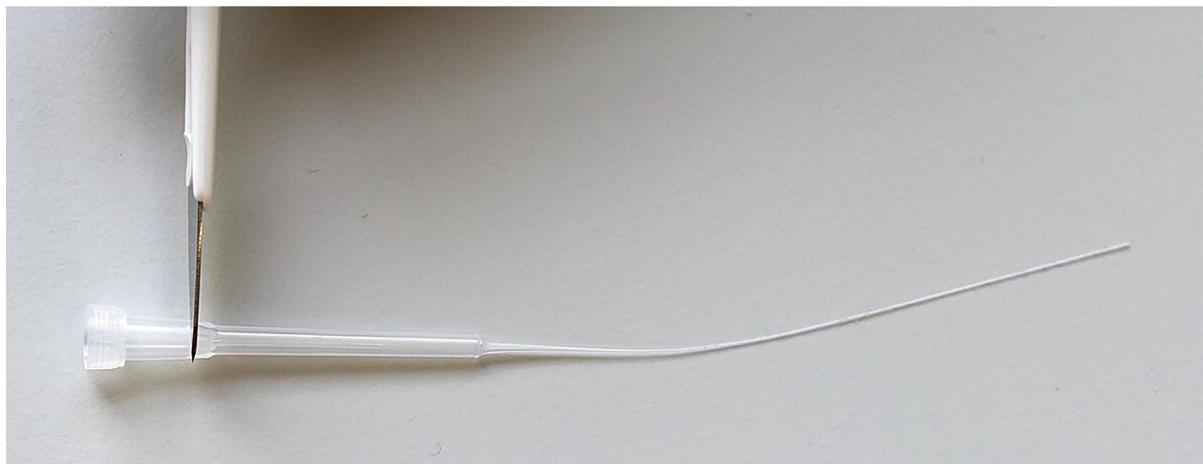
## 2.2 Sample Preparation

1. 500  $\mu$ L Flat cell (Rototec Spintec) (*see Fig. 1*).
2. Cleaning solution for cuvettes (Hellmanex III, Hellma).
3. Ultrapure Water.
4. Acetone (high purity).
5. At least 20 nmol oligonucleotide (e.g., DNAzyme:RNA-complex).
6. Buffer (*see Note 1*).
7. 1 M  $Mn^{2+}$  stock solution.
8. 10  $\mu$ L Microloader Tips (*see Fig. 2*).
9. 1 mL Tuberculin syringe (*see Fig. 2*).

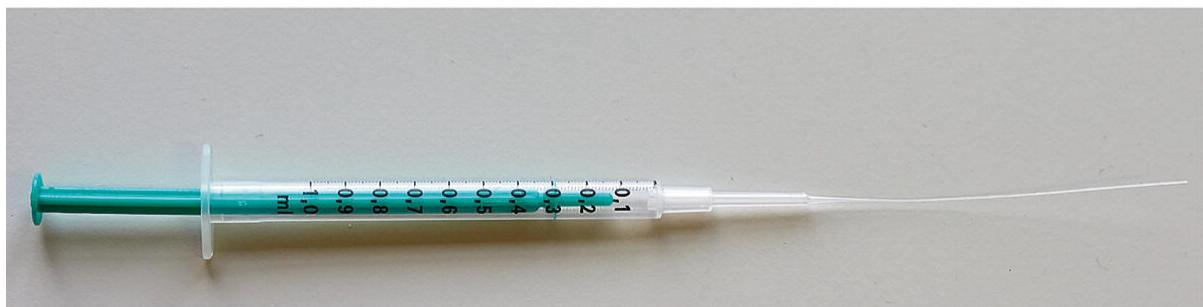


**Fig. 1** The flat cell

A



B



**Fig. 2** Preparation of the home-build syringe. (a) Approximately 4 mm of the tip groove is cut with a sterile scalpel. The end is shortly flamed to smoothen it. (b) Introducing the tip in a syringe leads to the home-build construction that enables proper sample mixing

---

### 3 Methods

#### 3.1 Starting the EMXmicro

1. Turn on the water cooling.
2. Switch on the spectrometer (power on the console and the magnet power supply).
3. Start the Xenon program.
4. Connect to the spectrometer (*see Note 2*).

5. Open the microwave bridge tuning dialog box.
6. Switch the microwave bridge into the Operate mode and wait for 1 h for the spectrometer to warm up (*see Notes 3 and 4*).

## 3.2 Measurement

### 3.2.1 Oligonucleotide Sample

The oligonucleotide measurement dictates the concentration range of the added Mn<sup>2+</sup>. Hence, we recommend measuring the oligonucleotide sample first and afterwards the reference sample to enable the detection in the same concentration range.

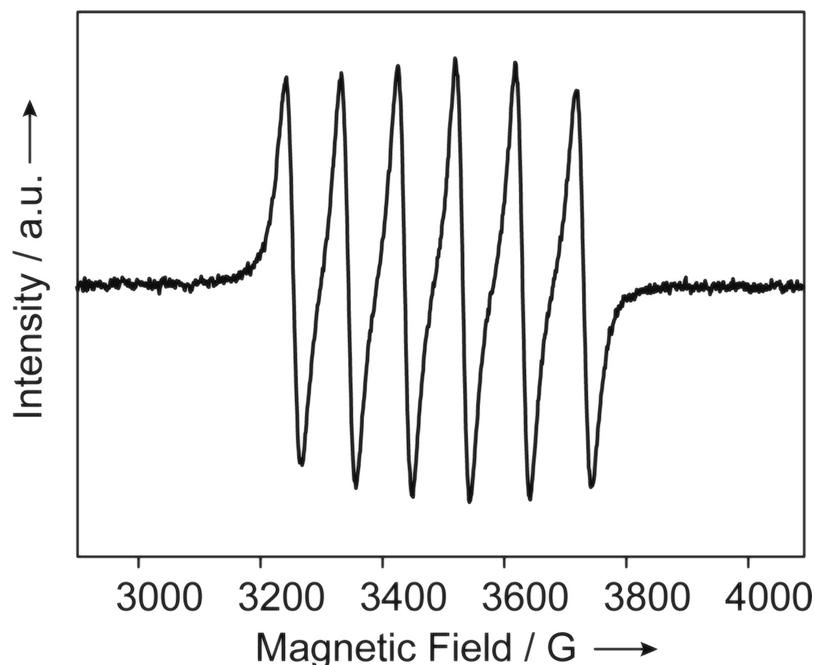
1. Incubate the flat cell in a cleaning solution for cuvettes (*see Note 5*).
2. Rinse the flat cell several times with RNase-free ultrapure water.
3. Wash the flat cell with pure acetone.
4. Blow the flat cell dry using, e.g., N<sub>2</sub> but be careful to not introduce dust.
5. Prepare a proper folded oligonucleotide sample with at least 40 µM oligonucleotide in 500 µL buffer (*see Note 6*).
6. Fill the oligonucleotide sample slowly into the flat cell using a standard 500 µL pipette (*see Note 7*).
7. Set the microwave bridge into the Tune mode and the attenuator to 60 dB.
8. Insert the flat cell into the resonator.
9. Tuning the spectrometer has the aim to make the dip as narrow and deep as possible. Set the attenuation to 20 dB and find the resonator

~~Keep as possible. Set the attenuation to 33 dB and tune the resonator dip by swiping the frequency. It may help to vary the position and orientation of the flat cell. For our setup, the dip was usually around~~

7.8 GHz. When the dip is found, it should be optimized by adjusting the orientation and position of the flat cell, the iris, and the frequency. Now you can also adjust the bias and the signal phase. Mark the position of the flat cell! Precisely the same position is required every time the flat cell is removed and again inserted, thus, after each titration data point (*see Note 8*).

10. Tune the resonator via the Auto Tune procedure (*see Note 9*).
11. After successful auto tuning, the spectrometer will switch to Operate mode. Go back into the Tune mode, adjust the attenuation to 33 dB, and measure the Q-value. We always found the Q-value to be  $3400 \pm 100$  (*see Note 10*).
12. Switch into Operate mode and set the measurement parameters (*see Note 11*). We have used a modulation frequency of 100 kHz, a modulation amplitude of 5 G and 1440 points in the field range of 2890–4090 G.
13. Record the spectrum, which should only be a baseline. Otherwise, clean the flat cell and check the sample solution for contaminations.
14. Save the spectrum.
15. Go into the Tune mode and increase the attenuation to 60 dB.
16. Remove the flat cell from the resonator.
17. Add 2  $\mu$ L of the Mn<sup>2+</sup> stock solution with a 10  $\mu$ L Microloader tip in combination with a 10  $\mu$ L pipette directly into the solution in the flat cell (*see Note 12*).
18. The whole solution is collected and mixed with a home-built syringe construction (*see Fig. 2*). Afterwards the homogenate solution is reinjected into the flat cell (*see Note 13*).

19. Repeat **steps 8–18** until enough titration points are recorded (see **Note 14** and Fig. 3 for a cw X-band  $\text{Mn}^{2+}$  EPR spectrum).
20. Recover the sample.
21. Clean the flat cell by rinsing with pure water, followed with acetone.
22. Blow the flat cell dry using, e.g.,  $\text{N}_2$ .



**Fig. 3**  $\text{Mn}^{2+}$  cw X-band EPR spectrum

### 3.2.2 Reference Sample

1. Fill 500  $\mu\text{L}$  of the buffer used for the oligonucleotide into the flat cell.
2. Conduct **steps 7–22** from Subheading 3.2.1 and record the same  $\text{Mn}^{2+}$  concentration range as used for the oligonucleotide sample.

## 3.3 Data Analysis

1. Doubly integrate each spectrum (see **Note 15**).

2.

The obtained values of the double integrals (DI) depend on the  $Mn^{2+}$  concentration that is free in solution ( $[Mn_{free}^{2+}]$ ). The values for DI are corrected by taking the Q-value (Q) and the microwave power (P) into account (see Note 16). The values of the corrected double integral ( $DI_{corr}$ ) are calculated by

$$DI_{corr} = \frac{DI}{Q \times \sqrt{P}}$$

3.

Using the  $DI_{corr}$  values of the reference sample and the oligonucleotide-containing sample one can calculate the concentration of free  $Mn^{2+}$  in the oligonucleotide-containing sample ( $[Mn_{free,sample}^{2+}]$ ). To do so, the  $DI_{corr,reference}$  is plotted against the concentration of  $Mn^{2+}$  in the reference ( $[Mn_{added,reference}^{2+}]$ ) (see Note 17) and the slope is determined ( $slope_{reference}$ ) (Fig. 4).

4.

$[Mn_{free,sample}^{2+}]$  at each titration point is obtained through the ratio of the double integral of the sample and the  $slope_{reference}$ .

$$[Mn_{free,sample}^{2+}] = \frac{DI_{corr,sample}}{slope_{reference}}$$

5.

The bound fraction of  $Mn^{2+}$  ( $[Mn_{bound,sample}^{2+}]$ ) is obtained by subtracting  $[Mn_{free,sample}^{2+}]$  from the added  $Mn^{2+}$  concentration ( $[Mn_{added,sample}^{2+}]$ ).

$$[Mn_{bound,sample}^{2+}] = [Mn_{added,sample}^{2+}] - [Mn_{free,sample}^{2+}]$$

6.

The binding isotherm is obtained by plotting the ratio between  $[Mn_{bound,sample}^{2+}]$  and oligo concentration [oligo] against the concentration of free  $Mn^{2+}$  in the oligo sample ( $[Mn_{sample}^{2+}]$ ) (see

Fig. 5).

7.

The binding isotherm is used to determine the number of  $Mn^{2+}$ -binding sites and their respective affinity constants  $K_D$  by fitting to a binding model.

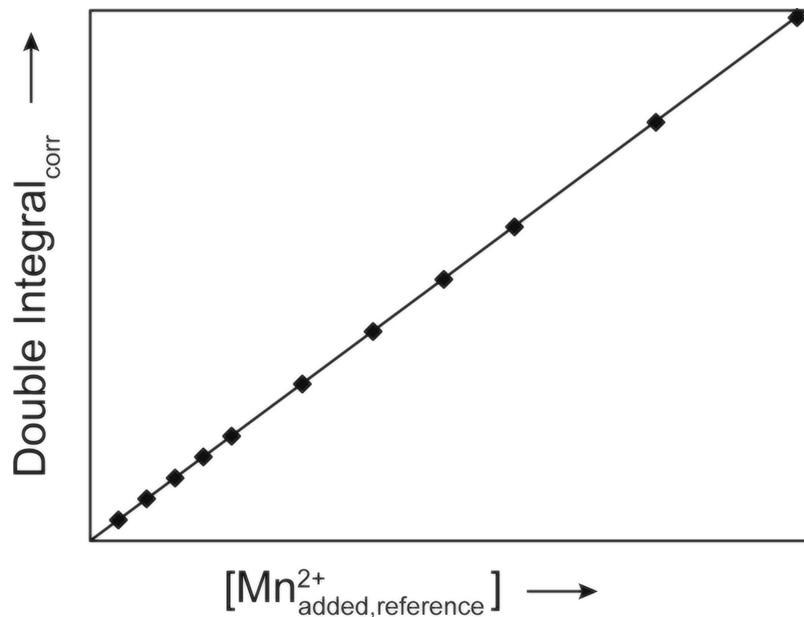
8.

A simple model for  $Mn^{2+}$  binding is to assume non-interacting binding sites. Each bound  $Mn^{2+}$  has  $n_i$  binding sites with dissociation constants  $K_{D,i}$ .

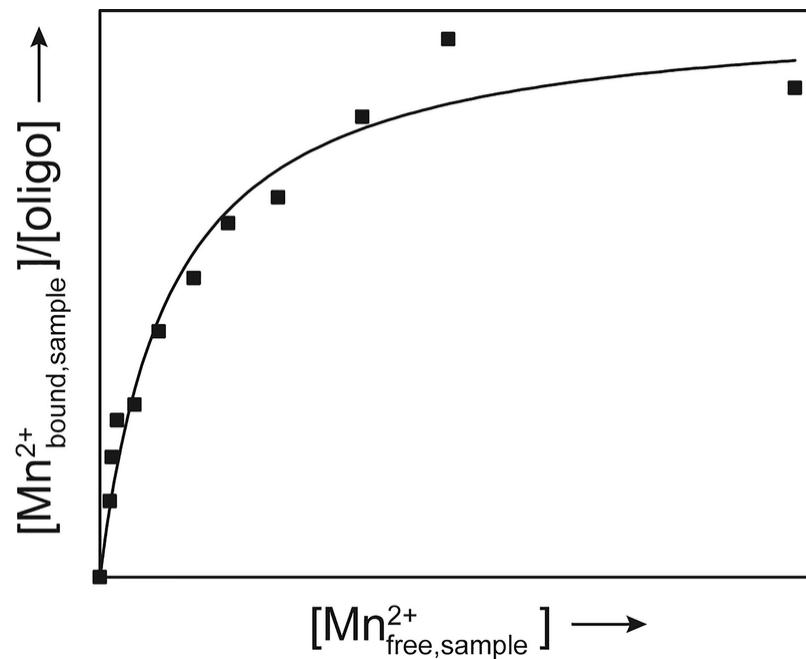
$$\frac{[Mn_{\text{bound,sample}}^{2+}]}{[\text{oligo}]} = \sum_{i=1}^j \frac{n_i \times [Mn_{\text{free,sample}}^{2+}]}{K_{D,i} + [Mn_{\text{free,sample}}^{2+}]}$$

Cooperative binding may be expressed by

$$\frac{[Mn_{\text{bound,sample}}^{2+}]}{[\text{oligo}]} = \sum_{i=1}^j \frac{[Mn_{\text{free,sample}}^{2+}]^{n_i}}{(K_{D,i} + [Mn_{\text{free,sample}}^{2+}])^{n_i}}$$



**Fig. 4**  $\text{Mn}^{2+}$  titration curve. The  $\text{DI}_{\text{corr,reference}}$  is plotted against  $[\text{Mn}^{2+}_{\text{added,reference}}]$ . Its slope is needed for the data analysis



**Fig. 5** Plot of a  $\text{Mn}^{2+}$  binding isotherm.  $[\text{Mn}^{2+}_{\text{bound,sample}}]/[\text{oligo}]$  is plotted against  $[\text{Mn}^{2+}_{\text{free,sample}}]$

## 4 Notes

1. Use a non-chelating buffer. Otherwise a competition might arise between the  $\text{Mn}^{2+}$  binding to the buffer vs. the oligonucleotide.
2. Make sure that the correct resonator is selected in the Xenon software.
3. Switch on the spectrometer at least 1 h before starting the experiments. If the spectrometer did not have ample time to warm up, the error of the measurements will be too large.
4. Make sure to use the “Operate” mode, because the spectrometer automatically switches into the “Standby” mode after a while, if “Tune” is used at this point.
5. We have seen that a flat cell that is not meticulously cleaned promotes that air bubbles are obtained when filling the flat cell. This will hinder

accurate data recording. Thus, we recommend incubating the flat cell

at least 2 h, best overnight, in 2% Hellmanex III cuvette cleaning solution.

6. It is crucial that the oligonucleotide is folded as homogeneous as possible. Differently folded conformations could result in different binding properties. Thus, a proper folded oligonucleotide is essential for good data quality.
7. Pipette the solution slowly into the flat cell from the top to avoid air bubbles.
8. If the dip is unstable, make sure that the spectrometer had proper time to warm up.
9. We have found that the dip is broad, but Auto Tuning is still possible! If not, adjust the dip manually and retry Auto Tuning.
10. The *Q*-value is required for quantitative measurements and can be used to judge whether the resonator is coupled correctly. For reproducibility, the *Q*-value should be the same for each measurement. If it is conspicuously different, repeat the tuning procedure.
11. If a similar spectrum was recorded before, data and measurement parameters can be reloaded. Make sure that “Normalize Acquisition” is switched on (Spectrometer Configuration → Signal Channel). By default, Xenon normalizes the data by a normalization constant taking the number of scans into account. Thus, higher number of scans are recommended in the lower Mn<sup>2+</sup> concentration regime. For time-compensation, at higher concentration the number of scans can then be decreased.
12. We performed the titrations without exchanging the tip to avoid different volumes due to manufacturing issues.
13. We have used a home-built syringe by mounting a 10 µL Microloader tip (Eppendorf) into a 1 mL Tuberculin syringe. Approximately 4 mm

of the tip's groove is cut with a sterile scalpel. The cut end of the tip is then shortly flamed to enable a smooth surface. The tip is then introduced into the syringe. 1 mL buffer is collected with these home-built syringe to check if it is leak-proof. Afterwards, the construction was used to fully collect the sample from the flat cell. The filled homo-built syringe is vortexed with the solution in it. The mixed solution is afterwards reinjected into the flat cell (Fig. 2).

14.

We recommend conducting data analysis parallel to the measurements. In this way, it can be ensured that the binding isotherm reaches the plateau.

15.

Double integration of the samples with low Mn<sup>2+</sup> concentration (<100 µM) was found to be reliable if performed as follows:

(a)

Single integration.

(b)

Baseline correction (polynomial of third order).

(c)

Single integration.

16.

The applied microwave power is provided in the respective DSC data file. The Q-value is only written therein if **step 11** during the data recording was followed.

17.

$\left[ \text{Mn}^{2+}_{\text{added,reference}} \right]$  is equal to the free Mn<sup>2+</sup> concentration in the reference measurement as the reference contains only buffer that does not bind Mn<sup>2+</sup>.

## Acknowledgments

We thank Hamed Alaei for taking the pictures of the equipment.

---

## References

1. Sigel A, Sigel H, RKO S (2012) Interplay between metal ions and nucleic acids. In: 1.

Aufl. Metal ions in life sciences, vol 10. Springer, Dordrecht

2. Sigel A, Sigel H, Sigel RKO (eds) (2011) Structural and catalytic roles of metal ions in RNA. Metal ions in life sciences, vol 9. De Gruyter, Berlin, Boston
3. Woodson SA (2005) Metal ions and RNA folding: a highly charged topic with a dynamic future. *Curr Opin Chem Biol* 9(2):104–109. <https://doi.org/10.1016/j.cbpa.2005.02.004>  
[Crossref][PubMed]
4. Müller J (2010) Functional metal ions in nucleic acids. *Metallomics* 2(5):318–327. <https://doi.org/10.1039/c000429d>  
[Crossref][PubMed]
5. Johnson-Buck AE, McDowell SE, Walter NG (2011) Metal ions: supporting actors in the playbook of small ribozymes. *Met Ions Life Sci* 9:175–196. <https://doi.org/10.1039/9781849732512-00175>  
[Crossref][PubMed][PubMedCentral]
6. Erat MC, Sigel RKO (2011) Methods to detect and characterize metal ion binding sites in RNA. In: Sigel A, Sigel H, Sigel RKO (eds) Structural and catalytic roles of metal ions in RNA, vol 9. De Gruyter, Berlin, Boston, pp 37–100  
[Crossref]
7. Handing KB, Niedzialekowska E, Shabalin IG et al (2018) Characterizing metal-binding sites in proteins with X-ray crystallography. *Nat Protoc* 13(5):1062–1090. <https://doi.org/10.1038/nprot.2018.018>  
[Crossref][PubMed][PubMedCentral]
8. Pley HW, Flaherty KM, McKay DB (1994) Three-dimensional structure of a hammerhead ribozyme. *Nature* 372:68–74. <https://doi.org/10.1038/372068a0>  
[Crossref][PubMed]
9. Marušić M, Schlagnitweit J, Petzold K (2019) RNA dynamics by NMR spectroscopy. *ChemBioChem* 20(21):2685–2710. <https://doi.org/10.1002/cbic.201900072>  
[Crossref][PubMed][PubMedCentral]
10. Pechlaner M, Sigel RKO (2012) Characterization of metal ion-nucleic acid interactions in solution. In: Sigel A, Sigel H, Sigel RKO (eds) Interplay between metal ions and nucleic acids, vol 10. Springer, Dordrecht, pp 1–42  
[Crossref]
11. Forconi M, Herschlag D (2009) Metal ion-based RNA cleavage as a structural probe. In: Biophysical, chemical, and functional probes of RNA structure, interactions and folding: part A, vol 468. Elsevier, Amsterdam, pp 91–106

[Crossref]

12. Salim NN, Feig AL (2009) Isothermal titration calorimetry of RNA. Methods 47(3):198–205. <https://doi.org/10.1016/j.ymeth.2008.09.003>  
[Crossref][PubMed]
13. Schweiger A, Jeschke G (2001) Principles of pulse electron paramagnetic resonance. Oxford Univ. Press, Oxford
14. Weil JA, Bolton JR (2007) Electron paramagnetic resonance: elementary theory and practical applications, 2nd edn. Wiley-Interscience, Hoboken, N.J
15. Goldfarb D, Stoll S (eds) (2018) EPR spectroscopy: fundamentals and methods. eMagRes books. Wiley, Chichester, West Sussex
16. Hunsicker-Wang L, Vogt M, DeRose VJ (2009) EPR methods to study specific metal-ion binding sites in RNA. Methods Enzymol 468:335–367. [https://doi.org/10.1016/S0076-6879\(09\)68016-2](https://doi.org/10.1016/S0076-6879(09)68016-2)  
[Crossref][PubMed]
17. Eaton GR, Barr DP, Eaton SS et al (2010) Quantitative EPR: a practitioners guide. Springer-Verlag Vienna, Vienna  
[Crossref]
18. Cohn M, Townsend J (1954) A study of Manganese complexes by paramagnetic resonance absorption. Nature 173:1090–1091. <https://doi.org/10.1038/1731090b0>  
[Crossref]
19. Danchin A, Guéron M (1970) Cooperative binding of manganese (II) to transfer RNA. Eur J Biochem 16:532–536. <https://doi.org/10.1111/j.1432-1033.1970.tb01113.x>  
[Crossref][PubMed]
20. Horton TE, Clardy DR, DeRose VJ (1998) Electron paramagnetic resonance spectroscopic measurement of Mn<sup>2+</sup> binding affinities to the hammerhead ribozyme and correlation with cleavage activity. Biochemistry 37(51):18094–18101. <https://doi.org/10.1021/bi981425p>  
[Crossref][PubMed]
21. Horton TE, DeRose VJ (2000) Cobalt Hexammine inhibition of the hammerhead ribozyme. Biochemistry 39(37):11408–11416. <https://doi.org/10.1021/bi001141g>  
[Crossref][PubMed]
22. Hunsicker LM, DeRose VJ (2000) Activities and relative affinities of divalent metals in unmodified and phosphorothioate-substituted hammerhead ribozymes. J Inorg Biochem 80(3-4):271–281. [https://doi.org/10.1016/S0162-0134\(00\)00079-9](https://doi.org/10.1016/S0162-0134(00)00079-9)

[Crossref][PubMed]

23. Kisseeleva N, Khvorova A, Westhof E et al (2005) Binding of manganese(II) to a tertiary stabilized hammerhead ribozyme as studied by electron paramagnetic resonance spectroscopy. *RNA* 11(1):1–6. <https://doi.org/10.1261/rna.7127105>  
[Crossref][PubMed][PubMedCentral]
24. Schiemann O, Fritscher J, Kisseeleva N et al (2003) Structural investigation of a high-affinity MnII binding site in the hammerhead ribozyme by EPR spectroscopy and DFT calculations. Effects of neomycin B on metal-ion binding. *Chembiochem* 4(10):1057–1065. <https://doi.org/10.1002/cbic.200300653>  
[Crossref][PubMed]
25. Wang KY, Gerena L, Swaminathan S et al (1995) Determination of the number and location of the manganese binding sites of DNA quadruplexes in solution by EPR and NMR. *Nucleic Acids Res* 23(5):844–848. <https://doi.org/10.1093/nar/23.5.844>  
[Crossref][PubMed][PubMedCentral]
26. Woody AY, Eaton SS, Osumi-Davis PA et al (1996) Asp537 and Asp812 in bacteriophage T7 RNA polymerase as metal ion-binding sites studied by EPR, flow-dialysis, and transcription. *Biochemistry* 35(1):144–152. <https://doi.org/10.1021/bi952037f>  
[Crossref][PubMed]
27. Kisseeleva N, Kraut S, Jäschke A, Schiemann O (2007) Characterizing multiple metal ion binding sites within a ribozyme by cadmium-induced EPR silencing. *HFSP J* 1(2):127–136. <https://doi.org/10.2976/1.2756332>  
[Crossref][PubMed][PubMedCentral]
28. Hoogstraten CG, Britt RD (2002) Water counting: quantitating the hydration level of paramagnetic metal ions bound to nucleotides and nucleic acids. *RNA* 8(2):252–260. <https://doi.org/10.1017/s1355838202014826>  
[Crossref][PubMed][PubMedCentral]
29. Morrissey SR, Horton TE, Grant CV et al (1999)  $Mn^{2+}$  –nitrogen interactions in RNA probed by electron spin–Echo envelope modulation spectroscopy: application to the hammerhead ribozyme. *J Am Chem Soc* 121(39):9215–9218. <https://doi.org/10.1021/ja9921571>  
[Crossref]
30. Morrissey SR, Horton TE, DeRose VJ (2000)  $Mn^{2+}$  sites in the hammerhead ribozyme investigated by EPR and continuous-wave Q-band ENDOR spectroscopies. *J Am Chem Soc* 122(14):3473–3481. <https://doi.org/10.1021/ja992989z>  
[Crossref]

31. Vogt M, Lahiri S, Hoogstraten CG et al (2006) Coordination environment of a site-bound metal ion in the hammerhead ribozyme determined by  $^{15}\text{N}$  and  $^2\text{H}$  ESEEM spectroscopy. *J Am Chem Soc* 128(51):16764–16770. <https://doi.org/10.1021/ja057035p>  
[Crossref][PubMed][PubMedCentral]
32. Schiemann O, Carmieli R, Goldfarb D (2007) W-band  $^{31}\text{P}$ -ENDOR on the high-affinity  $\text{Mn}^{2+}$  binding site in the minimal and tertiary stabilized hammerhead ribozymes. *Appl Magn Reson* 31(3–4):543–552. <https://doi.org/10.1007/BF03166601>  
[Crossref]

# **Part III**

## **Biophysical Characterization of DNAzymes**

## 8. Obtaining Crystals of Nucleic Acids in Complex with the Protein U1A Using the Soaking Method

Hannah Rosenbach<sup>1</sup>✉ and Ingrid Span<sup>1</sup>✉

(1) Institut für Physikalische Biologie, Heinrich-Heine-Universität  
Düsseldorf, Düsseldorf, Germany

✉ **Hannah Rosenbach**

Email: [hannah.rosenbach@hhu.de](mailto:hannah.rosenbach@hhu.de)

✉ **Ingrid Span (Corresponding author)**

Email: [ingrid.span@hhu.de](mailto:ingrid.span@hhu.de)

### Abstract

X-ray crystallography is one of the most prominent techniques for determining high-resolution structures of nucleic acids. The major challenges are to obtain well-diffracting single crystals and to solve the phase problem. The absence of structural information impedes the elucidation of the molecular details of biological processes. A particularly intriguing example is the RNA-cleavage catalyzed by the 10–23 deoxyribozyme (DNAzyme). This DNAzyme consists of a catalytic core that is flanked by two substrate binding arms, which can be designed to bind any RNA of interest. Structure elucidation of the 10–23 DNAzyme in a biologically relevant conformation faces three major challenges: (1) stabilization of the RNA substrate to capture the DNA:RNA complex in the pre-catalytic conformation, (2) prevention of the formation of an artificial duplex conformation due to a self-complementary sequence in the catalytic

core of the DNAzyme, and (3) the crystallization of nucleic acids with their uniform surfaces. Here, we provide a protocol for an innovative strategy facilitating the crystallization of protein:nucleic acid complexes using a soaking approach and discuss on how to apply this protocol for the structure elucidation of the 10–23 DNAzyme. For this purpose, we describe the purification procedure of an optimized variant of the RNA-binding protein U1A, the crystallization of this specific U1A variant, the soaking process with its specific RNA hairpin loop, and finally suggest a strategy for applying this procedure on the 10–23 DNAzyme in complex with its specific RNA target.

**Key words** X-ray crystallography – Crystallization – Soaking – Nucleic acids – 10–23 DNAzyme – Deoxyribozyme – U1 small nuclear ribonucleoprotein A (U1A)

---

## 1 Introduction

X-ray crystallography is an excellent method to obtain insights into the structure of nucleic acids on a molecular level as well as into structural relationships between nucleic acids and metal ions or other ligands. Structure elucidation of nucleic acids using X-ray crystallography, however, is often impeded by their regularly ordered and negatively charged phosphate backbone that often leads to crystals with a poor long-range order. Approaches to tackle this problem involve co-crystallization of the nucleic acid with a crystallization helper such as the human U1 small nuclear ribonucleoprotein A (U1A) that was previously used for the crystallization of several large RNAs including ribozymes and riboswitches [1, 2]. The first crystal structure of an RNA-cleaving DNAzyme, the so-called 8–17 DNAzyme [3], was obtained in the presence of the *African swine fever virus* DNA polymerase X (AsfvPolX) [4]. Proteins that serve as crystallization helpers bind to a specific nucleic acid motif. The U1A protein binds a specific RNA hairpin loop with high affinity [5]. The AsfvPolX associates with certain double-stranded DNA sequences [6].

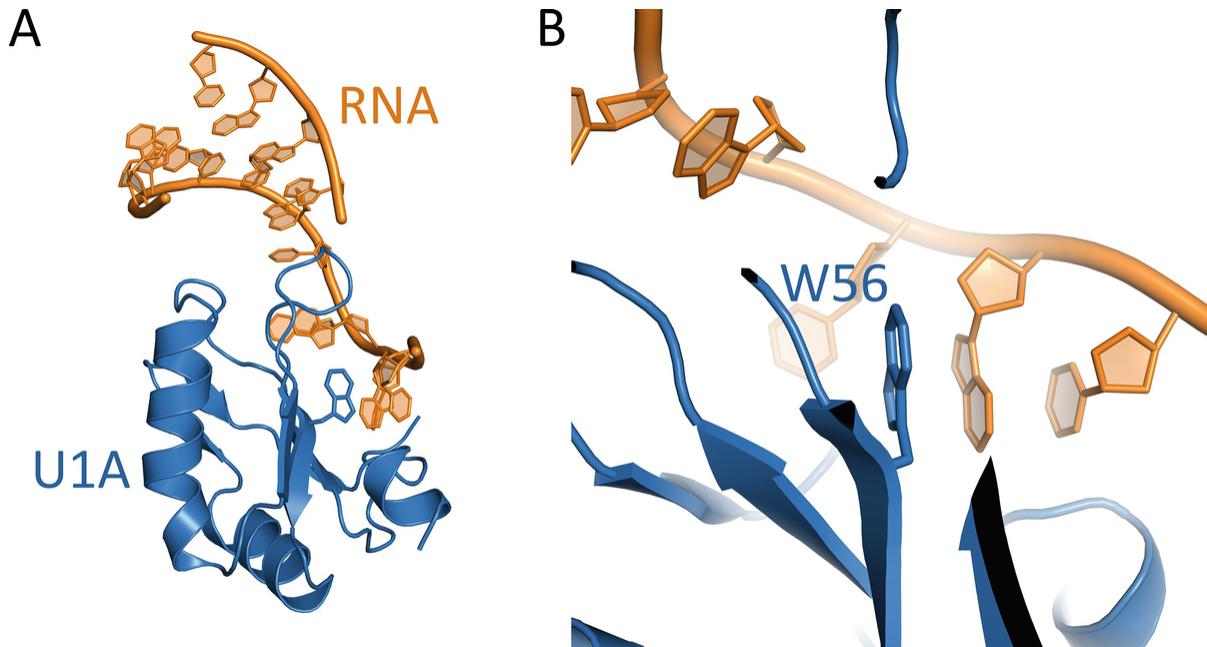
In contrast to nucleic acids, proteins consist of a larger variety of building blocks and their surface consists of many different amino acid side chains, which makes them very diverse and enables specific long-range interactions between different molecules. In addition, the proteins selected

as crystallization helpers have a compact fold facilitating the crystal packing. Furthermore, the co-crystallization with a known protein allows to determine the phases using molecular replacement.

Therefore, the formation of nucleic acid:protein complexes supports the formation of well-ordered crystals and contributes to solving the phase problem.

The RNA-binding domain (RBD) of the human U1A protein has been used for the crystallization of several different RNA sequences [7]. Therefore, a previously reported co-crystallization protocol with either wild-type or the double mutant U1AY31HQ36R U1A protein has been used [5, 8]. However, this approach has two major drawbacks: (1) the wild-type or the double mutant U1A proteins lack a tryptophan residue that is necessary for distinguishing between salt and protein crystals using fluorescence and (2) co-crystallization of the protein:nucleic acid complex requires the incubation of the biological sample under sometimes harsh crystallization conditions, such as high pH values or in the presence of high concentrations of heavy metals, which may support autocatalytic nucleic acid cleavage.

Here, we provide a protocol for the recombinant production of a tryptophan-containing variant of the U1A-RBD [9]. The introduced tryptophan residue enables the determination of protein concentration using electronic absorption spectroscopy and facilitates the identification of protein crystals by ultraviolet (UV) fluorescence. As the phenylalanine 56 that was exchanged by a tryptophan is located in the RNA-binding cavity of U1A (Fig. 1) [5, 10, 11], the binding to the RNA hairpin loop leads to a modulation of the fluorescence signal, which allows us to monitor the RNA-binding in the crystal. Furthermore, we provide instructions for performing soaking experiments with RNA molecules that specifically bind to the U1A-RBD into pre-formed U1A protein crystals and suggest strategies for the design of structural studies of the 10–23 DNAzyme.



**Fig. 1** Crystal structure of the F56W U1A variant. **(a)** Overall architecture of the U1A:RNA hairpin loop complex with the U1A protein shown in blue and the RNA shown in orange. A short sequence of the RNA that is not involved in binding to U1A is not well defined and missing in the structure. **(b)** Close-up view of the protein-RNA binding site highlighting the interaction of the tryptophan introduced at position 56 with a nucleic acid base

## 2 Materials

Prepare all buffers and stock solutions using deionized water from a water purification system with a filter of 0.2 µM. Protein-containing solution should be kept at 4 °C during the purification process. Isolated and purified U1A variants can be stored at –80 °C until they are needed.

### 2.1 Expression Plasmids and Gene Sequences

1.

- Expression plasmid pET16bTEV-U1AF56W (*see* Table 1 for gene sequences and Fig. 2 for plasmid map).

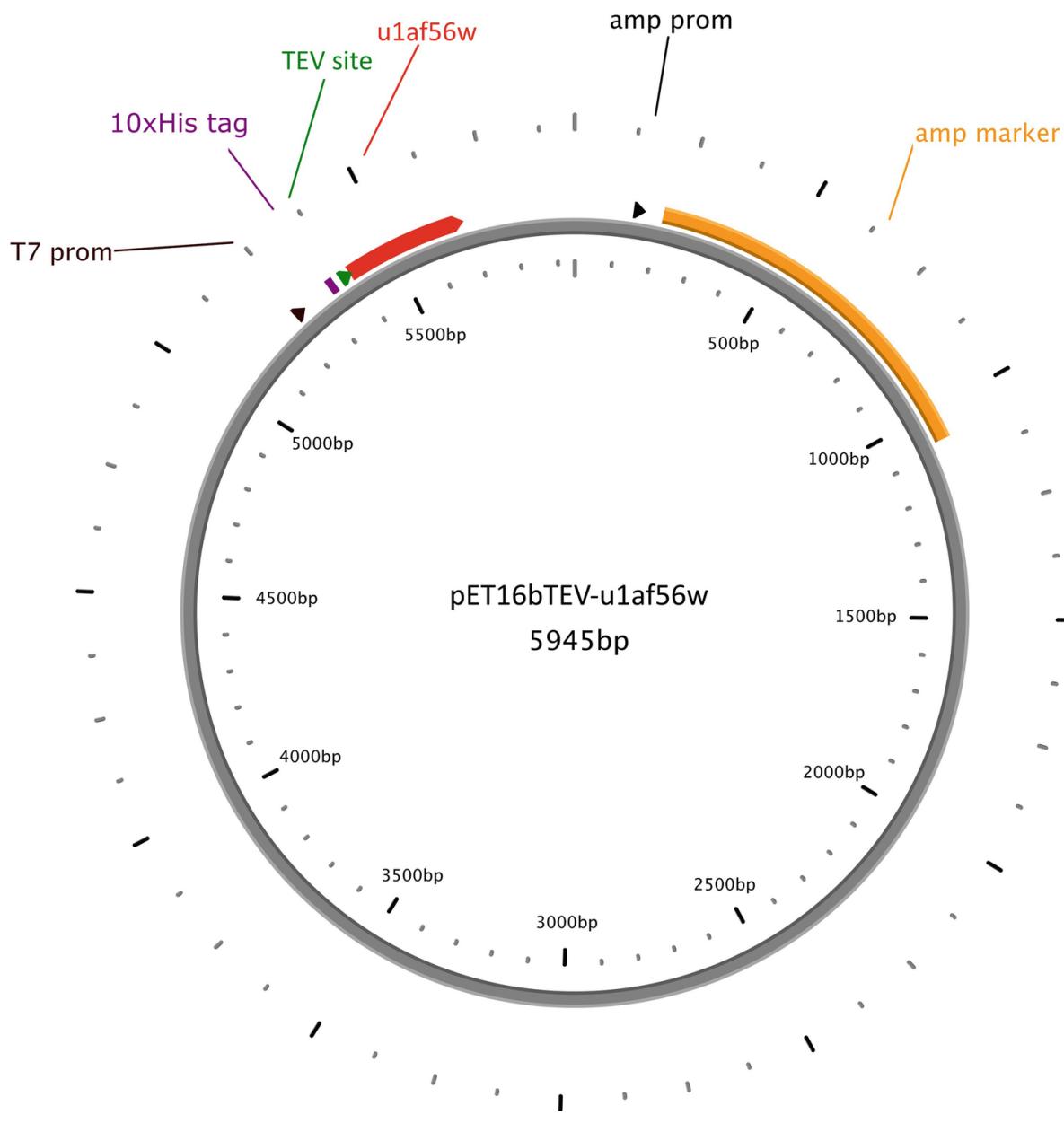
**Table 1** Sequences of the U1AF56W gene and the corresponding protein U1AF56W. In the gene sequence, the nucleotides encoding the poly-histidine tag are highlighted in purple, the nucleotides encoding the tobacco etch virus (TEV) recognition site are highlighted in green, the start codon is highlighted in blue, and the mutated codon encoding the amino

acid W56 is highlighted in red. In the protein sequence, the amino acid at position 56 is highlighted in red

---

Gene	CATCATCATCATCATCATCATCATAGCAGCGGCCATGAAAAC
sequence	CTGTATTTCAAGGGACATATGCAGTTCCGAGACCCGCCCTAACCA CACTATTATATCAACAAACCTCAATGAGAAGATCAAGAAGGATGAGCT AAAAAAAGTCCCTGCACGCCATCTTCTCCGGTTGCCAGATCCTGG ATATCCTGGTATCACGGAGCCTGAAGATGAGGGGCCAGGCCCTGGTC ATCTTCAAGGAGGTCAGCAGGCCACCAACGCCCTGCGCTCCATGCA GGGTTCCCTTCTATGACAAACCTATGCGTATCCAGTATGCCAAGAC CGACTCAGATATCATTGCCAAG ATGAAATAA
Protein	MAVPETRPNH TIYINNLNEK IKKDELKKSL HAIFSRFGQI LDILVSRSLK
sequence	MRGQA <b>W</b> VIFK EVSSATNALR SMQGFIFYDK PMRIQYAKTD SDIIAKMK

---



**Fig. 2** Map of the plasmid pET16bTEV-U1AF56W. The sequence encoding the U1AF56W protein (*U1AF56W*) is indicated in red, the sequence encoding the TEV cleavage site (TEV site) in green, and the sequence encoding the poly-histidine tag in purple. The T7 promotor region (T7 prom) and the ampicillin promotor region (amp prom) are shown in black, and the sequence encoding the ampicillin resistance (amp marker) is highlighted in yellow

## 2.2 Protein Preparation

1. Terrific Broth (TB) media.

2. 100 mg/L Ampicillin stock solution.
3. Biological incubator shaker.
4. Spectrophotometer.
5. 0.5 M Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) stock solution.
6. 2,2',2'',2'''-(Ethane-1,2-diyldinitrilo)tetraacetic acid (EDTA)-free protease inhibitor tablets.
7. High-speed centrifuge.
8. Sonification system for cell disruption (sonotrode).
9. Fast protein liquid chromatography (FPLC) system (*see Note 1*).
10. Diethylpyrocarbonate (DEPC)-treated H<sub>2</sub>O (*see Note 2*).
11. Nickel-nitrilotriacetic acid (Ni-NTA) column (5 mL bed volume).
12. Superdex-75 PG column (120 mL bed volume).
13. CHT-1 Hydroxyapatite column (5 mL bed volume).
14. Ultrafiltration device with a molecular weight cut-off of 3 kDa.
15. Lysis and wash buffer A: 50 mM 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES) pH 7.5, 100 mM KCl.
16. 5% Polyethylenimine (PEI) pH 7.5.
17. Magnetic stirrer.
18. 4.5 M Ammonium sulfate.

19. 0.5 M NaOH.
20. Buffer B: 50 mM HEPES pH 7.5, 100 mM KCl, 250 mM imidazole.
21. Buffer C: 10 mM potassium phosphate pH 7.5, 50 mM KCl.
22. Buffer D: 10 mM potassium phosphate pH 7.5, 50 mM KCl, 0.5 M ammonium sulfate.
23. Crystallization buffer: 25 mM Tris pH 7.5, 100 mM KCl.

## **2.3 Crystallization of U1A Variants F56W**

1. SWISSCI MRC 2 well crystallization plate (96-well plate format).
2. Sealing film for crystallization plates.
3. Protein solution: 1 mM protein in crystallization buffer.
4. Reservoir solutions for the crystallization of the U1A variant F56W:  
2.2 M ammonium sulfate and 0.2 mM tri-potassium citrate.
5. Microscope capable of ultraviolet imaging.

## **2.4 Soaking of RNA into Pre-formed Protein Crystals**

1. RNA solution: 1.5 mM in crystallization buffer.
2. Pre-formed protein crystals.
3. Cryoprotecting solution: 50% glycerol.

## **2.5 Electrophoretic Mobility Shift Assays (EMSA)**

1. 10× Tris-borate-EDTA (TBE) buffer stock solution: 890 mM Tris, 890 mM boric acid, 20 mM EDTA (pH 8.0), autoclaved.

2. 10% native polyacrylamide gel electrophoresis (PAGE), 30 mL: 3 mL 10× TBE buffer, 7.5 mL acrylamide/bisacrylamide (37.5:1), 0.1% (v/v) Triton X-100, 0.1% (v/v) tetramethylethylenediamine (TEMED), 0.1% (w/v) ammonium persulfate (APS).
  3. Fluorescent nucleic acid dye gel stain.
  4. Imaging system capable of detecting fluorescence.
- 

### 3 Methods

#### 3.1 Protein Preparation

1. Cultivate cultures of *E. coli* BL21(DE3), which contain the respective plasmid in Terrific Broth (TB) medium supplemented with ampicillin (100 µg/mL) at 37 °C and 160 rpm until the optical density at 600 nm (OD<sub>600</sub>) reaches a value of 2.
2. Induce gene expression by adding 0.5 mM IPTG to cell cultures.
3. Grow cells for approximately 18 h at 25 °C and 120 rpm.
4. Harvest cells by centrifugation at 6000 × *g* for 10 min and 4 °C. Resuspend cells in lysis buffer using a volume of 5 mL/g cell pellet. Add EDTA-free protease inhibitor cocktail tablets according to the manufacturer's instructions.
5. Sonicate the suspension for 20 min with an amplitude of 60% and a pulse of 1 s every 3 s using a suitable sonotrode.
6. Separate cell debris from the lysate by centrifugation for 45 min at 40,000 × *g*.
7. Fractionate the supernatant with 0.5% PEI, pH 7.5 to remove nucleic acids from the sample. Incubate the sample for 5 min on ice, vortex thoroughly, and centrifuge for 5 min at 15,000 × *g* and 4 °C.

8. Use the supernatant for precipitation with 75% (v/v) of a saturated ammonium sulfate solution to remove the protein from the solution. For this purpose, stir the protein solution on a magnetic stirrer at 4 °C. Resuspend the protein pellet in wash buffer.
9. Isolate the protein using a Ni-NTA column with a bed volume of 5 mL, which is equilibrated with wash buffer, using a chromatography system. Wash the column with 9 column volumes (CV) 25 mM imidazole in wash buffer. Elute the protein with 250 mM imidazole in wash buffer.
10. Pool fractions containing the protein and dialyze against 2 L of wash buffer for 2 h at 4 °C.
11. Treat the protein with TEV protease in a molar ratio of 1:100 in the presence of 0.5 mM EDTA and 1 mM dithiothreitol (DTT) for 16 h to remove the poly-histidine tag.
12. Remove TEV protease and poly-histidine tag using a Ni-NTA column, which is equilibrated with wash buffer containing 37.5 mM imidazole. Apply the protein to the column and collect the flow-through.
13. Concentrate the protein by ultracentrifugation using an ultrafiltration device with a molecular weight cut-off of 3 kDa.
14. For the next steps: wash the chromatography systems and columns with 0.5 M NaOH before usage and then flush with DEPC-treated H<sub>2</sub>O. Prepare all solutions with RNase-free water and buffers containing RNase-free grade chemicals.
15. Perform size-exclusion chromatography (SEC) using a Superdex-75 PG column pre-equilibrated with RNase-free wash buffer. Fractions containing the protein of interest are pooled and concentrated to a volume of 2.5 mL using an ultrafiltration device with a molecular weight cut-off of 3 kDa. Exchange the buffer of your sample to buffer C.

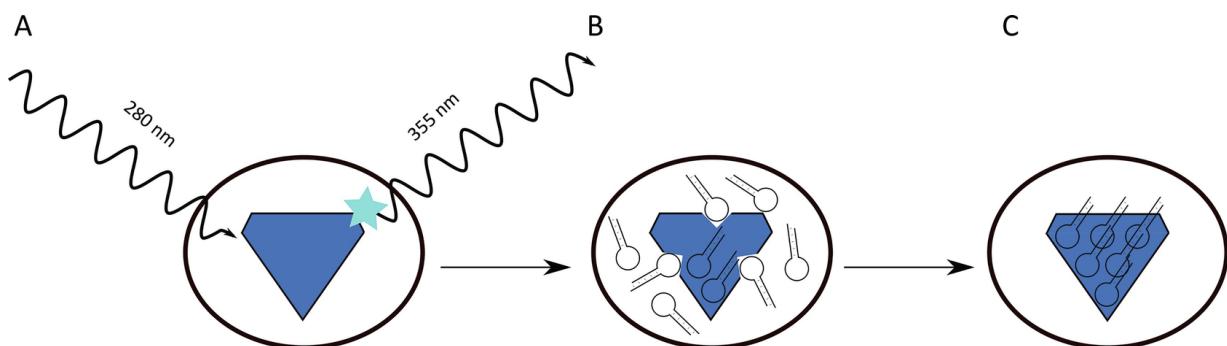
16. Use a CHT-1 hydroxyapatite column for the final purification step. Wash the CHT-1 column with 3 CV of 0.5 M NaOH. Then, equilibrate the column with 5 CV of buffer C. Apply the protein in buffer C to the column and run a 100 mL gradient to 100% buffer D. The protein elutes at a conductivity, which corresponds to an ammonium sulfate concentration of 250 mM.
17. Collect fractions that contain the target protein and exchange the buffer to crystallization buffer and store the protein at -80 °C.

### 3.2 Crystallization of U1A Variants

1. Prepare a protein solution with a concentration of 1 mM U1A variant F56W in crystallization buffer.
2. Place 50 µL precipitant in the reservoir well of the crystallization plate.
3. Place a 1 µL drop of the protein solution into the crystallization well.
4. Dilute the protein drop with 1 µL precipitant.
5. Seal the plate with sealing tape and incubate at 12 °C for 4 weeks.

### 3.3 Soaking of RNA into Pre-formed Protein Crystals

1. Mix 0.7 µL of a 1.5 mM RNA solution in wash buffer with an equal volume of the precipitant in the reservoir well, in which the protein crystal was grown, in the crystallization well of a 96-well plate. The reservoir remains empty.
2. Transfer a pre-formed protein crystal from its mother liquor into the drop containing the reservoir solution and the RNA (see Fig. 3).
3. Seal the plate and incubate the plate for 4 days at 12 °C.
4. Transfer the soaked crystal into a drop of cryoprotectant (50% glycerol). Flash-freeze the crystal and store it in liquid nitrogen.



**Fig. 3** Schematic representation of the soaking process. (a) The pre-formed protein crystal shows the typical tryptophan fluorescence when it is exposed to UV light. (b) The pre-formed protein crystal is transferred to a drop that contains the specific RNA sequence, which is capable of binding to the U1A protein. (c) The crystal is incubated under these conditions for 4 days at 12 °C, before the crystal is soaked in cryoprotectant, and flash-frozen in liquid nitrogen

### 3.4 Monitoring RNA-Binding to Pre-formed U1AF56W Crystals Using Tryptophan Fluorescence

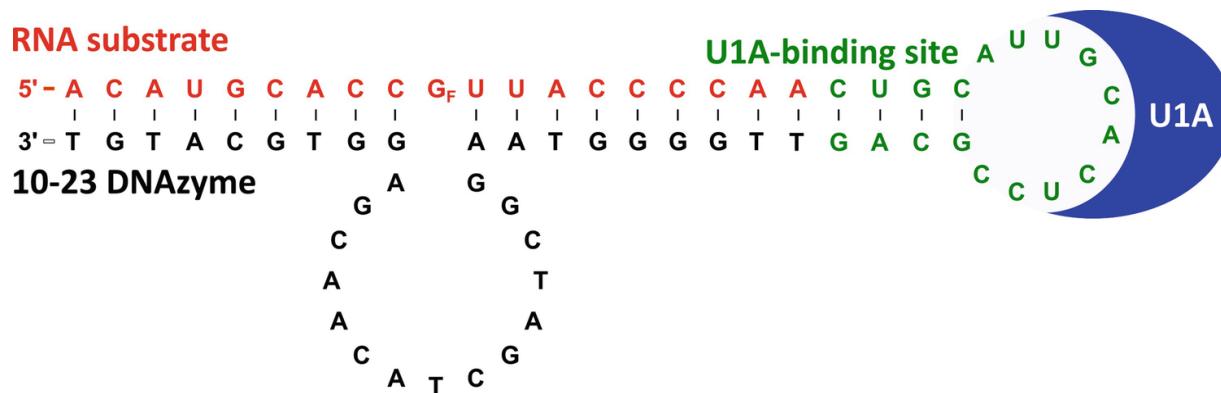
1. To monitor RNA-binding to the pre-formed U1AF56W crystal, transfer the soaked crystal back to its mother liquor prior to visualization (*see Note 3*).
2. Monitor the fluorescence in an appropriate imager that is capable of imaging crystal plates.

### 3.5 Design of a Suitable Biological Sample Comprising the 10–23 DNAzyme, its Specific RNA Target and the U1A Protein as a Crystallization Helper

1. Formation of the ternary complex (*see Fig. 4*) requires the presence of the U1A-specific RNA hairpin loop. Perform complex design using the software RNACofold [12] and RNAfold [13] with RNA default parameters 37 °C (*see Note 4*).
2. To crystallize the DNAzyme:RNA complex in a biologically relevant conformation, make sure that the RNA substrate is stabilized at its cleavage site and cannot be cleaved by the DNAzyme. For this

**Cleavage site** and cannot be cleaved by the DNAzyme. For this

purpose, we suggest substituting the 2'-OH group at the cleavage site by a fluoride, since it prevents transesterification while largely preserving the chemical properties (such as size and electronegativity) of the substituted OH group. Another possible modification may be substitution by a methoxy group.



**Fig. 4** Schematic representation of the ternary complex consisting of the 10–23 DNAzyme (black), the RNA substrate with the stabilized nucleotide at the cleavage site ( $\text{G}_\text{F}$ ) (red) that is elongated by the specific U1A-binding loop (green), as well as the U1A variant as a crystallization module (blue)

### 3.6 Sequential Complex Formation and its Verification by Electrophoretic Mobility Shift Assay

1. Dissolve DNAzyme and RNA target sequences in wash buffer. Mix equal amounts of DNAzyme and RNA target. Heat the sample to 73 °C for 5 min, followed by 20 min at room temperature to allow complex formation. Add 1 mM MgCl<sub>2</sub> and 2 mM spermidine.
2. Mix the DNAzyme:RNA complex with an equal amount of protein in wash buffer. Incubate the solution for 30 min at room temperature to allow formation of the ternary complex.
3. Mix the sample with DNA loading dye (see Note 5) and analyze 10 µM of the complex on a native PAGE for 3 h at 180 V (see Note 6).
4. For staining incubate the gel with a fluorescent nucleic acid dye gel stain in a ratio of 1:10,000 (see Note 7) for 1 h in 150 mL TBE buffer and visualize the gel with an imaging system.

---

## 4 Notes

1. Make sure that the chromatography apparatus tolerates exposure with 0.5 M NaOH. If not, consider other methods to clean the system properly.
2. DEPC is an alkylating agent that is used to destroy RNases. For this purpose, mix 250 µL DEPC with 1 L H<sub>2</sub>O and incubate for 10 min. Free DEPC can be destroyed by autoclaving the water for at least 30 min at 121 °C. During the autoclaving process DEPC is converted into ethanol and CO<sub>2</sub>. It is important that free DEPC is completely destroyed before the water comes in contact with the protein or RNA solution, since unreacted DEPC can bind to aromatic amino acids or nucleotides.
3. Due to the overlap of the absorption spectra of the RNA with the emission spectra of the tryptophan, it is crucial that the solution surrounding the crystal does not contain free RNA molecules to obtain a reliable result.
4. The following criteria should be considered: (1) the nucleic acids strands should have a low tendency for intramolecular folding, except for the formation of the hairpin loop, (2) the RNA and DNA, respectively, have a low tendency for intermolecular interactions, except for the formation of the desired DNAzyme:RNA complex, (3) formation of the DNAzyme:RNA complex with minimum free energy, (4) the hairpin loop is easily accessible for the U1A protein, (5) the hairpin loop is stable, and (6) for soaking experiments with the DNAzyme:RNA complex, the adjacent 3' end of the RNA should not be able to stack on the 5' end of the DNA strand to provide more flexibility, while for co-crystallization experiments, the 3' end of the RNA and the 5' end of the DNAzyme should be stacked to stiffen the conformation.
5. Make sure that the dye does not contain EDTA and denaturing agents such as urea or formamide.

6. For a gel with the following size: 20 cm × 21 cm.

7.

Of course, also other staining methods can be considered, e.g., silver staining.

---

## References

1. Cochrane JC, Lipchock SV, Strobel SA (2007) Structural investigation of the GlmS ribozyme bound to its catalytic cofactor. *Chem Biol* 14:97–105. <https://doi.org/10.1016/j.chembiol.2006.12.005>  
[Crossref][PubMed]
2. Ferré-D’Amaré AR, Zhou K, Doudna JA (1998) Crystal structure of a hepatitis delta virus ribozyme. *Nature* 395:567–574. <https://doi.org/10.1038/26912>  
[Crossref][PubMed]
3. Santoro SW, Joyce GF (1997) A general purpose RNA-cleaving DNA enzyme. *Proc Natl Acad Sci U S A* 94:4262–4266. <https://doi.org/10.1073/pnas.94.9.4262>  
[Crossref][PubMed][PubMedCentral]
4. Liu H, Yu X, Chen Y et al (2017) Crystal structure of an RNA-cleaving DNazyme. *Nat Commun* 8:1–10. <https://doi.org/10.1038/s41467-017-02203-x>  
[Crossref]
5. Oubridge C, Ito N, Evans PR et al (1994) Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* 372:432–438. <https://doi.org/10.1038/372432a0>  
[Crossref][PubMed]
6. Chen Y, Zhang J, Liu H et al (2017) Unique 5'-P recognition and basis for dG:dGTP misincorporation of ASFV DNA polymerase X. *PLoS Biol* 15:e1002599. <https://doi.org/10.1371/journal.pbio.1002599>  
[Crossref][PubMed][PubMedCentral]
7. Ferré-D’Amaré AR (2010) Use of the spliceosomal protein U1A to facilitate crystallization and structure determination of complex RNAs. *Methods* 52:159–167. <https://doi.org/10.1016/j.ymeth.2010.06.008>  
[Crossref][PubMed][PubMedCentral]
- 8.

- Oubridge C, Ito N, Teo CH et al (1995) Crystallisation of RNA-protein complexes II. The application of protein engineering for crystallisation of the U1A protein-RNA complex. *J Mol Biol* 249(2):409–423. <https://doi.org/10.1006/jmbi.1995.0306> [Crossref][PubMed]
9. Rosenbach H, Victor J, Borggräfe J, Biehl R, Steger G, Etzkorn M, Span I (2020) Expanding crystallization tools for nucleic acid complexes using U1A protein variants. *J Struct Biol* 210: 107480. <https://doi.org/10.1016/j.jsb.2020.107480>
  10. Kranz JK, Hall KB (1999) RNA recognition by the human U1A protein is mediated by a network of local cooperative interactions that create the optimal binding surface. *J Mol Biol* 285:215–231. <https://doi.org/10.1006/jmbi.1998.2296> [Crossref][PubMed]
  11. Law MJ, Chambers EJ, Katsamba PS et al (2005) Kinetic analysis of the role of the tyrosine 13, phenylalanine 56 and glutamine 54 network in the U1A/U1 hairpin II interaction. *Nucleic Acids Res* 33:2917–2928. <https://doi.org/10.1093/nar/gki602> [Crossref][PubMed][PubMedCentral]
  12. Hofacker IL, Fontana W, Stadler PF et al (1994) Fast folding and comparison of RNA secondary structures. *Monatshefte Chemie Chem Mon* 125:167–188. <https://doi.org/10.1007/BF00818163> [Crossref]
  13. Bernhart SH, Tafer H, Mückstein U et al (2006) Partition function and base pairing probabilities of RNA heterodimers. *Algorithms Mol Biol* 1(1):3. <https://doi.org/10.1186/1748-7188-1-3> [Crossref][PubMed][PubMedCentral]

## 9. Crystallization and Structural Determination of 8–17 DNAzyme

Hehua Liu<sup>1</sup>, Song Mao<sup>2</sup>, Jia Sheng<sup>2</sup>✉ and Jianhua Gan<sup>1</sup>✉

- (1) Shanghai Public Health Clinical Center, State Key Laboratory of Genetic Engineering, Collaborative Innovation Center of Genetics and Development, School of Life Sciences, Fudan University, Shanghai, China  
(2) Department of Chemistry and The RNA Institute, University at Albany, State University of New York, Albany, NY, USA

✉ Jia Sheng (Corresponding author)

Email: [jsheng@albany.edu](mailto:jsheng@albany.edu)

✉ Jianhua Gan (Corresponding author)

Email: [ganjh@fudan.edu.cn](mailto:ganjh@fudan.edu.cn)

### Abstract

DNAzymes are a group of DNA molecules that can catalyze various chemical reactions. Owing to their great application potentials, DNAzymes have received significant attention. However, due to their intrinsic difficulties in crystallization and structural determination, only very limited structural information of DNAzymes is available to date. Using co-crystallization with the *African Swine Fever Virus* Polymerase X (AsfvPolX) protein, we have recently solved a complex structure of the 8–17 DNAzyme, which represents the first structure of the catalytically active RNA-cleaving DNAzyme. In this chapter, we describe the detailed protocols including gene construction, AsfvPolX expression and purification, crystallization, structure determination, and in vitro cleavage assay. While the specific methods described herein were originally designed for the 8–17 DNAzyme, they can also be utilized to solve other DNAzyme structures.

**Key words** Crystallization – DNAzyme – 8–17 DNAzyme – RNA cleavage – Structure determination

---

### 1 Introduction

The 8–17 DNAzyme is one of the earliest discovered DNAzymes [1, 2]. It can cleave RNA in a sequence-specific manner and has been applied in many research areas, such as lead ion detection [3, 4], mRNA fragmentation for subsequent analysis [5], gene-silencing, and controlling viruses associated with human diseases [6]. Because of its great application potentials, 8–17 DNAzyme has attracted tremendous attention ever since its first discovery [7–9]. It was found that, besides RNA substrates, 8–17 DNAzyme can also cleave DNAs with single ribonucleotide at the cleavage site [1, 10]. However, due to its intrinsic difficulty in crystallization, no 8–17 DNAzyme crystal structure has been obtained for more than 20 years. To reveal the underlying basis for substrate recognition and degradation by this DNAzyme, we purified various DNA-binding proteins and performed co-crystallization trials. As a result, we determined a complex structure of 8–17 DNAzyme with AsfvPolX, representing the first structure of the catalytically active RNA-cleaving DNAzyme. The structure was refined at atomic resolution (2.55 Å) and it revealed a novel V-shape overall folding with many unexpected structural features of 8–17 DNAzyme [11].

The crystals were grown in the presence of AsfvPolX, the DNA Polymerase X encoded by the genome of *African Swine Fever Virus* (ASFV) [12]. AsfvPolX can be expressed in *E. coli* on a high level. It is very stable and can be readily purified to homogeneity. AsfvPolX is very unique and the smallest polymerase (174 amino acids in length) identified to date (see Note 1). The DNA-binding affinity of AsfvPolX is very high. The benefits of introducing the AsfvPolX into the complex system include: (1) it stabilizes the complex formed between 8–17 DNAzyme and the substrate analog, and significantly improves the molecular packing and crystallization success rate; (2) it can facilitate the structure determination process. In addition to crystallization, solving the phase problem is another major bottleneck in crystal structure determination. Since many high-resolution AsfvPolX

structures have been deposited in the protein database bank [13], they can serve as ideal search models in solving the phase problem by molecular replacement methods [14].

Here we describe the detailed procedures in solving the 8–17 DNAzyme complex structure, including gene construction, *AsfvPolX* expression and purification, crystallization, diffraction data collection, structure determination, and in vitro cleavage assays. While these protocols were designed for characterizing 8–17 DNAzyme structure, they can also be easily modified to study the structures of other DNAzymes.

---

## 2 Materials

All solutions are prepared using analytical grade reagents and ultrapure water, which is produced from deionized water with a sensitivity of 18 MΩ cm at 25 °C. Unless indicated otherwise, solutions and reagents are all prepared and stored at room temperature. All waste disposal regulations are strictly followed when disposing waste materials.

### 2.1 Gene Construction

1. The codon-optimized cDNA (5'-ATGCTGACCCTCATCCAGGGTAAAAAGATCGTTAACCAACCTGCGTTCTCGTCTGGCGTCGAATACA 3') of *AsfvPolX*.
2. The forward primer (5'-AAAGGATCCGGTGGTGGTATAGCTGACCCTCATCC-3') and reverse primer (5'-A/-
3. PrimeSTAR GXL DNA polymerase (1.25 U/ul).
4. 5× PrimeSTAR GXL buffer ( $Mg^{2+}$  plus).
5. Deoxynucleotide (dNTP) mixture (2.5 mM each).
6. The pET28a-Sumo vector, which encodes 6xHis-Sumo tag and Ulp1 protease recognizing sequence in the N-terminus.
7. 10× CutSmart buffer.
8. 10× T4 ligase buffer.
9. Endonucleases *Bam*HI and *Xho*I.
10. T4 DNA ligase.
11. LB-agar: 5 g yeast extract, 10 g tryptone, 10 g NaCl, and 15 g agar are dissolved in 800 mL water. Volume is adjusted to 1000 mL.
12. LB medium: 5 g yeast extract, 10 g tryptone, and 10 g NaCl are dissolved in 800 mL water. Volume is adjusted to 1000 mL.
13. 50 mg/mL Kanamycin is dissolved in ddH<sub>2</sub>O, filtered through a 0.22 μm syringe and stored at -20 °C.
14. Chemically competent bacterial cells of *E. coli* BL21(DE3).

### 2.2 Protein Expression

1. pET28a-Sumo vector containing the optimized cDNA of *AsfvPolX*.
2. LB medium: 5 g yeast extract, 10 g tryptone, and 10 g NaCl are dissolved in 800 mL water. Volume is adjusted to 1000 mL and autoclaved by high-pressure sterilizer at 121 °C for 30 min.
3. 50 mg/mL Kanamycin.

4. 0.2 M Isopropyl  $\beta$ -D-1-thiogalacto-pyranoside (IPTG) is dissolved in water, filtered through a 0.22  $\mu$ m syringe and stored at -20 °C.
5. Flasks with a volume of 2 L.

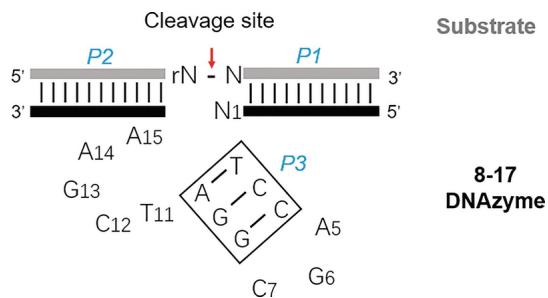
## 2.3 Protein Purification

1. Phosphate-buffered saline (PBS) buffer: 137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 2 mM KH<sub>2</sub>PO<sub>4</sub>.
2. Lysis and wash buffer: 20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 25 mM imidazole.
3. Cell disruption system.
4. Chromatography system.
5. Nickel-nitrilotriacetic acid (Ni-NTA)-column elution buffer: 20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 500 mM imidazole.
6. Ulp1 protease [15].
7. Dialysis buffer: 20 mM Tris-HCl, pH 8.0, 500 mM NaCl.
8. Dilution buffer: 20 mM Tris-HCl, pH 8.0, 2 mM Dithiothreitol (DTT).
9. HiTrap SP-column binding buffer: 20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 2 mM DTT.
10. HiTrap SP-column elution buffer: 20 mM Tris-HCl, pH 8.0, 1 M NaCl, 2 mM DTT.
11. Gel filtration buffer: 50 mM Tris-HCl, pH 7.5, 300 mM NaCl.
12. Ni-NTA column with a volume of 5 mL.
13. HiTrap SP column with a volume of 5 mL.
14. Dialysis tube: membrane 10,000 MWCO.
15. Ultrafiltration device: membrane 10,000 MWCO.
16. Hi 16/60 Superdex G75 column: bed volume 120 mL.
17. Sodium dodecyl sulfate (SDS) Polyacrylamide Gel Electrophoresis (PAGE) gel.
18. Equipment for SDS-PAGE.
19. Coomassie brilliant blue stain.

## 2.4 Crystallization of 8–17 DNAzyme Complex

1. 8–17 DNAzyme (5'-TGTAACGCAGGCCAGCGGCTCGAAATCTCTCGT-3') and the corresponding substrate analog (5'-ACGAGAGAGATGGGTGCGTTACA-3') are designed according to the sequences (Fig. 1) reported previously (see **Notes 2** and **3**) and dissolved in ddH<sub>2</sub>O (2 mM).
2. AsfvPolX protein (see Subheadings 3.1–3.3).

3. 20 mM PbCl<sub>2</sub> solution. Stored at -20 °C.
4. Gryphon crystallization robot system (Arts Robbin Instruments).
5. Crystallization kits: Crystal Screen, Index, PEG/Ion, Natrix (Hampton Research); Wizard I, Wizard II (Emerald Biosystems); JCSG+ Suite, JCSG Core I, II, III, and IV Suites (Qiagen).
6. Intelli-Plate 96-3 LVR crystallization plate (Hampton Research).
7. 24-Well vapor diffusion crystallization plates.
8. Transparent cover tape for crystallization plates.
9. Microscope.



**Fig. 1** Sequence and predicted secondary structure of 8–17 DNAzyme. 8–17 DNAzyme recognizes its substrate through base pairing at two arming regions, P1 and P2

## 2.5 Structure Determination

Software: HKL2000, HKL3000 [16], CCP4i [17] and PHENIX Suites [18], and Coot [19].

## 2.6 In Vitro Cleavage Assay

1. Native and mutated 8–17 DNAzymes {Native: 5'-TGTAACGCACTGCCAGCGGCTGAAATCTCTCTCGT-3'; 6mG6: 5'-TGTAACGCACTGCCA(6mG)CGGCTGAAATCTCTCTCGT-3'; 1mG13: 5'-TGTAACGCACTGCCAGCGGCTC(1mG)AAATCTCTCTCGT-3'}.
2. 5'-Carboxyfluorescein (FAM) labeled substrate.
3. 2× Reaction buffer: 800 mM KCl, 200 mM NaCl, 100 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)-NaOH pH 7.5, 0.5 mM PbCl<sub>2</sub> (see Note 4).
4. 8 M Urea 18% PAGE gel.
5. 5× Tris-buffer Ethylenediaminetetraacetic acid (EDTA) (TBE) buffer: 54 g Tris base, 27.5 g boric acid, and 20 mL 0.5 M EDTA (pH 8.0) are dissolved in ddH<sub>2</sub>O. Adjust volume to 1 L using ddH<sub>2</sub>O and store at room temperature.
6. Gel loading buffer: 90% formamide, 40 mM EDTA, 0.01% xylene cyanol.
7. Typhoon™ FLA 9000 Imaging Scanner (GE Healthcare).
8. GraphPad Prism 5 software.

---

### 3 Methods

All procedures should be carried out at room temperature unless otherwise stated.

#### 3.1 Construction and Transformation of AsfvPolX Expression Vector

1. Amplify the codon-optimized cDNA of AsfvPolX by PCR method. The forward and reverse primers contain *Bam*HI and *Xho*I restriction sites, respectively.
2. Digest the cDNA by *Bam*HI and *Xho*I at 37 °C for 2 h. Digest the pET28a-Sumo vector by *Bam*HI and *Xho*I at 37 °C for overnight. Digestion system: 43 µL DNA, 1 µL *Bam*HI, 1 µL *Xho*I, 5 µL 10× CutSmart buffer.
3. Separate the digested DNA and vector on an agarose gel and recover using the standard gel extraction kit and protocols.
4. Ligate recovered DNA and pET28a-Sumo vector. Ligation system: 100 ng DNA, 30 ng vector, 1 µL T4 ligase, 2 µL 10× T4 ligase buffer. Adjust volume to 20 µL using ddH<sub>2</sub>O. Mix and incubate at room temperate for 2 h.
5. Add 10 µL mixture of ligation product to 100 µL *E. coli* BL21(DE3) competent cells in a 1.5 mL reaction tube and incubate on ice for 30 min. Heat shock the cells for 45 s in a 45 °C water bath. Immediately, insert the tube into ice and chill for 5 min. Add 500 µL LB medium and incubate in 37 °C shaker for 1 h.
6. Spread the culture onto the LB plate containing Kanamycin (50 µg/mL), and incubate at 37 °C for 10–12 h.
7. Pick single colony and use colony to inoculate 500 µL LB medium containing Kanamycin (50 µg/mL). Incubate cultures at 37 °C and shake at 200 rpm for 4 h. Send halves of the cultures for sequencing.
8. Mix the second halves of the cultures with 30% glycerol and store in –80 °C freezer until usage.

#### 3.2 AsfvPolX Expression

1. Revive the frozen pET28a-Sumo-AsfvPolX expression strain in 25 mL LB medium containing Kanamycin (50 µg/mL). Shake at 225 rpm at 37 °C for overnight.
2. Use the 25-mL revived bacterium suspension to inoculate 1 L LB medium containing Kanamycin (50 µg/mL). Shake with 225 rpm at 37 °C.
3. When optical density at 600 nm (OD<sub>600</sub>) reaches 0.6, add 1/1000 volume of IPTG to the culture (final concentration of 0.2 mM). Lower the temperature to 18 °C and shake for additional 18 h.
4. Harvest the cells by centrifugation at 4 °C for 15 min, 2000 × g. Resuspend the pellets in 100 mL PBS buffer and centrifuge again. Store the pellets at –20 °C until usage.

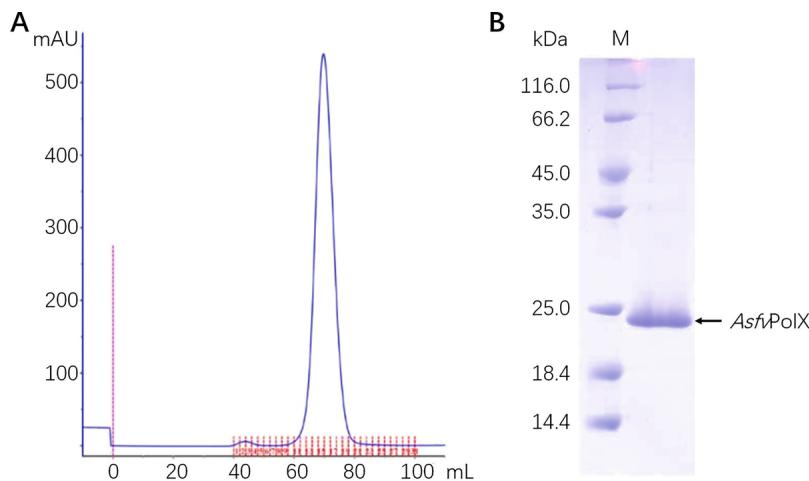
#### 3.3 AsfvPolX Purification

1. Resuspend the frozen cell pellets using six volumes of lysis and wash buffer. Lyse the cells under high pressure using a cell disruption system. Clarify the homogenate by centrifugation at 4 °C for 1 h at 14,500 × g.
2. Load the supernatant onto a Ni-NTA column (5 mL) equilibrated with lysis and wash buffer. Wash the column with 100 mL wash buffer, followed by 25 mL buffer composed of 15% Ni-NTA-column elution buffer and 85% wash buffer. Elute the His-Sumo-AsfvPolX protein using buffer containing 50% Ni-NTA-column elution buffer and 50% wash buffer.
3. Add Ulp1 protease to the fractions containing the desired fusion protein. Dialyze against dialysis buffer at 4 °C for 3 h using a dialysis tube with a molecular weight cut-off of 10 kDa.

4. Apply the sample to the Ni-NTA-column equilibrated with lysis and wash buffer, collect the flow through,
5. which contains the desired *AsfvPolX* protein.

Lower the salt concentration of the *AsfvPolX* sample to 150 mM by adding the dilution buffer. Apply the sample to HiTrap SP column (5 mL) equilibrated with HiTrap SP-column binding buffer. Wash the column with 50 mL HiTrap SP-column binding buffer, followed by 20 mL buffer containing 70% HiTrap SP-column binding buffer and 30% HiTrap SP-column elution buffer. Elute *AsfvPolX* protein with buffer composed of 50% HiTrap SP-column binding buffer and 50% HiTrap SP-column elution buffer.

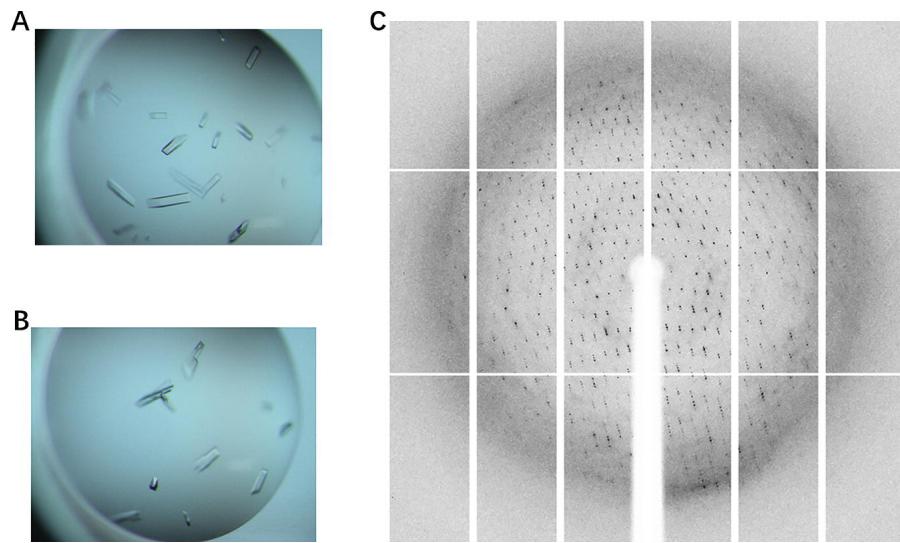
6. Concentrate the protein to 5 mL using an ultrafiltration device with a molecular weight cut-off of 10 kDa and further purify using a Hi 16/60 Superdex G75 column equilibrated with Gel Filtration Buffer (Fig. 2a). Analyze the purity of the proteins by a SDS-PAGE. Stain the gel using Coomassie brilliant blue stain (Fig. 2b).



**Fig. 2** Purification and analysis of *AsfvPolX* protein. **(A)** Gel filtration profile of *AsfvPolX*. **(B)** SDS-PAGE gel analysis of *AsfvPolX*. Protein molecular weight marker was labeled M on the gel. 8  $\mu$ L *AsfvPolX* sample was loaded on the gel and the result was visualized by Coomassie brilliant blue stain

### 3.4 Crystallization of 8–17 DNAzyme Complex

1. Mix 8–17 DNAzyme, substrate analog DNA, and *AsfvPolX* (in 50 mM Tris, pH 7.5, 300 mM NaCl). The concentrations of the DNAzyme, substrate analog, and protein are 0.3, 0.3, and 0.2 mM, respectively. If needed, add PbCl<sub>2</sub> to a final concentration of 1.0 mM (see Note 4).
2. Incubate the mixture at room temperature for 15 min (see Note 5).
3. Identify the crystallization conditions using Gryphon crystallization robot system and commercial crystallization kits. Every kit has 96 different conditions (buffer, pH, salt, and precipitant). Dispense 40  $\mu$ L of reservoir solution onto the 96-well crystallization plate. Mix 0.3  $\mu$ L of sample and 0.3  $\mu$ L of reservoir solution in each drop.
4. Seal the plate with transparent cover tape and gently place it in an 18 °C incubator.
5. Check the sample drops under a microscope.
6. Both DNAzyme and DNAzyme-Pb<sup>2+</sup> crystals appeared in the condition composed of 0.2 M di-ammonium hydrogen phosphate and 20% PEG 3350. Repeat and optimize the crystals using 24-well vapor diffusion crystallization plates at 16 °C. The well contains 500  $\mu$ L of crystallization buffer. The drop contains 1  $\mu$ L of sample and 1  $\mu$ L of crystallization solution. Crystals (Fig. 3a, b) appear within 2 days and reach their full size in about a week.



**Fig. 3** Crystals of 8–17 DNAzyme in the (A) presence and (B) absence of  $\text{Pb}^{2+}$ . (C) Typical diffraction image of the 8–17 DNAzyme- $\text{Pb}^{2+}$  complex crystals

### 3.5 Data Collection and Phase Determination

1. Cryo-protect all crystals by sequential soaking in their mother liquor supplemented with glycerol (from 5% to 20%). Flash-freeze the crystals in liquid nitrogen.
- 2.
3. Collect diffraction images (Fig. 3c) on the biological macromolecule crystallography beamline BL17U or BL19U of the Shanghai Synchrotron Radiation Facility (SSRF) at cryogenic temperature. Record diffraction patterns with image per  $1^\circ$  on a full rotation of  $360^\circ$ .
4. Process the diffraction images in a dataset with HKL2000 or HKL3000 program [16], determining and storing the unit cell parameters, space group, Miller indexes of reflections, intensities, and error estimates of reflections in a \*.sca file. Data collection statistics are stored in a \*.log file.
5. Solve the phase problem by molecular replacement method [14] as follows: Convert the \*.sca file to an \*.mtz file using the program SCALPACK2MTZ of the CCP4i Suite. Calculate the solvent content of the structure by the MATTHEWS\_COEF. Run with PHASER MR with *AsfvPolX* structure (PDB ID: 5HRB) as the search model. In our experiment, the PHASER program identified four *AsfvPolX* molecules.

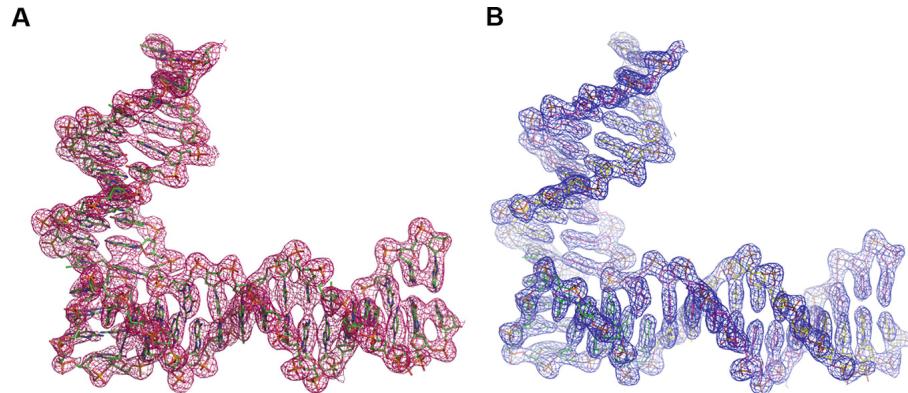
### 3.6 Model Building and Refinement

1. Refine the initial model using Refmac5 program in the CCP4i suite, which—in our experiment—generated one interpretable  $2F_o - F_c$  electron density map for the missing 8–17 DNAzyme and its substrate DNA analog.
2. Manually build DNAzyme and substrate sequence into the according positive electron density using COOT.
3. Refine the model using Refmac5. Use the resulting  $2F_o - F_c$  and  $F_o - F_c$  electron density maps as guide for further model adjustment and building of missing phosphate ions,  $\text{H}_2\text{O}$  molecules, and  $\text{Pb}^{2+}$  ions by COOT.
4. Refine the final model with phenix.refine program in the PHENIX suite.
5. Analyze the Ramachandran plot, geometry, rotamer, and other parameters of the structure by validation programs embedded in COOT.

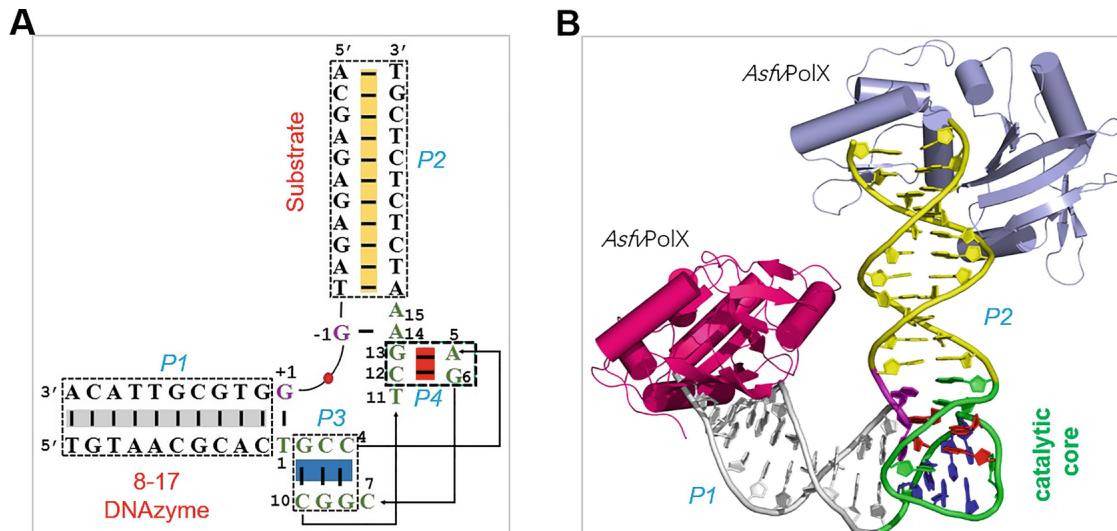
6.

Deposit the final model into the protein database bank.

The structure belongs to space group C222<sub>1</sub>, it contains four AsfvPolX molecules and two 8–17 DNAzyme/substrate DNA analog complexes per asymmetric unit. The overall folding of 8–17 DNAzyme/substrate DNA analog complex matches well with the omit map and the  $2F_o - F_c$  electron density map (Fig. 4). The structure reveals unexpected base pairing patterns in the 8–17 DNAzyme, including the A5:G13 and G6:C12 pairs at the P3 stem region (Fig. 5a). Each 8–17 DNAzyme/substrate DNA analog complex is flanked by two AsfvPolX molecules (Fig. 5b). Each AsfvPolX molecule recognizes six base pairs at the blunt end of 8–17 DNAzyme/substrate complex (see Notes 6 and 7).



**Fig. 4** The 8–17 DNAzyme/substrate DNA analog complex outlined with (A) the annealed  $F_o - F_c$  omit map (contoured at 3.0  $\sigma$  level) and (B) the refined  $2F_o - F_c$  map (contoured at 1.5  $\sigma$  level), respectively

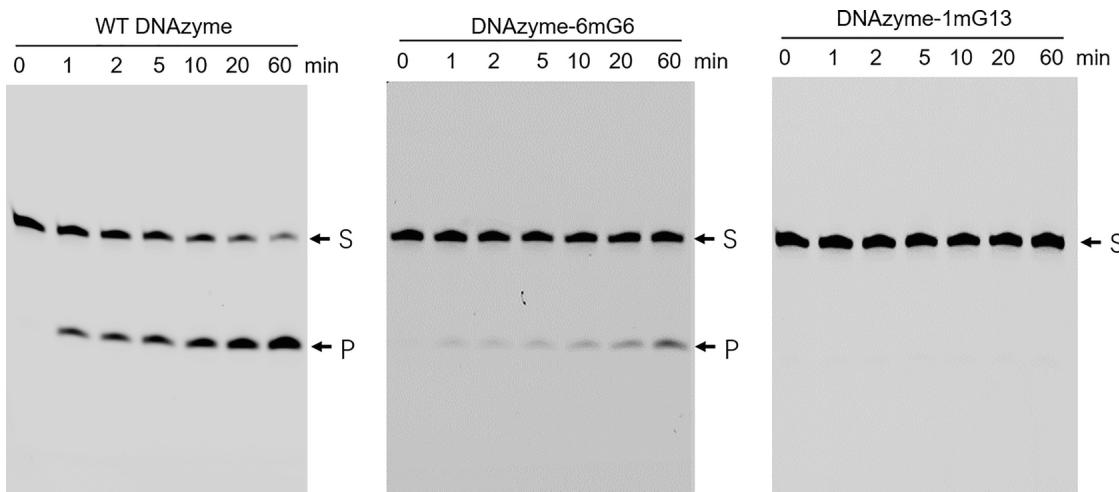


**Fig. 5** (A) Sequences and base pairing between 8–17 DNAzyme and the substrate DNA analog. (B) Interaction between AsfvPolX and the 8–17 DNAzyme/substrate DNA analog complex. AsfvPolX molecules are colored in purple and light blue

### 3.7 In Vitro Cleavage Assay

1. According to the structure, design a series of 8–17 DNAzymes with mutation, deletion, and modification.
2. Purify the purchased 5'-FAM labeled substrates by 18% PAGE gel with 8 M urea.
3. Set up 6  $\mu$ L of reaction system by mixing 3  $\mu$ L 2× reaction buffer, 1.5  $\mu$ L of DNAzyme (0.2  $\mu$ M), and 1.5  $\mu$ L of DNA substrate (0.4  $\mu$ M). Incubate the mixture at 26 °C for 1 h.

4. Quench the reaction by adding 6  $\mu$ L of Gel loading buffer at various time points.
5. Heat the samples at 95 °C for 3 min, centrifuge 14,500  $\times$  g for 5 min, and load onto an 18% PAGE gel with 8 M urea.
6. Run the gel at 10 W for 60 min in 0.5 $\times$  TBE buffer.
7. Scan the gel by Typhoon™ FLA 9000 Imaging Scanner (Fig. 6).
8. Quantify the substrate and product bands, analyze the data using the GraphPad Prism 5 software.



**Fig. 6** Comparison of substrate cleavage by wild-type (WT), 6mG6-modified, and 1mG13-modified 8–17 DNAzyme

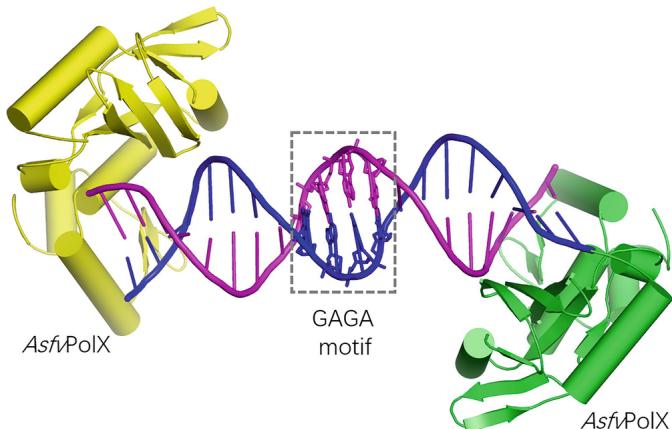
## 4 Notes

1. *AsfvPolX* functions in the base excision repair (BER) pathway of ASFV. Different from other DNA repair polymerases that are usually composed of multiple domains, *AsfvPolX* only contains two domains: the finger and the palm domain.
2. To maximize the crystallization success rate, systematically increase or decrease the pairing nucleotides at the P1 and P2 regions of 8–17 DNAzyme and the substrate analog.
3. To prevent occurrence of substrate cleavage, native or 2'-O-methyl modified DNAs are used as substrate analogs in the crystallization trials.
4. The catalytic activity of 8–17 DNAzymes can be enhanced by various cations, but it is most active in the presence of Pb<sup>2+</sup>. For convenience, Pb<sup>2+</sup> can be pre-included in the reaction buffer.
- 5.

When mixing *AsfvPolX* with certain DNA strands, precipitation can occur. In such cases, adjust the salt concentration to 500 mM by adding 5 M NaCl solution to increase the solubility of the complex. Remove the remaining precipitation by centrifugation at 4 °C for 2 min at 14,500  $\times$  g.

6. The interactions between *AsfvPolX* and the 8–17 DNAzyme/substrate complex are not sequence-specific. Instead of interacting with nucleobases, *AsfvPolX* mainly forms H-bond interactions with the phosphate backbone of the DNAzyme and substrate analog.
7. The sequence-unspecific interaction model and the strong binding affinity make *AsfvPolX* very useful in structural studies of nucleic acids. Following the protocols described in this chapter, we performed

crystallographic studies for various DNAs and solved one complex structure of a DNAzyme with self-catalyzed depurination activity [20]. The structure was refined at 2.85 Å and has been deposited in the protein database bank (PDB 7CPW). Due to G:A Hoogsteen interaction between the –GAGA– motifs, two DNAzymes with depurination activity formed an inactive duplex in the structure. Like the 8–17 DNAzyme, the depurination DNAzyme is flanked by two *Asfv*PolX molecules at the blunt ends (Fig. 7). The depurination DNAzyme study further demonstrated that *Asfv*PolX is useful in nucleic acid structural determination. In principle, *Asfv*PolX can also be utilized in the structural studies of many other DNAzymes, such as DNAzymes with DNA cleavage [21], DNA-ligation [22], and DNA phosphorylation [23] activities. The conserved *Asfv*PolX-DNA binding model can be considered as a general strategy and molecular tool for the experimental design of new DNA targets.



**Fig. 7** Crystal structure of *Asfv*PolX-complexed DNAzyme with depurination activity

## Acknowledgments

This work is supported by the National Natural Science Foundation of China (32171197 and 31870721), and US National Science Foundation (NSF CHE-1845486 and MCB-1715234).

## References

- Brown AK, Li J, Pavot CM, Lu Y (2003) A lead-dependent DNAzyme with a two-step mechanism. *Biochemistry* 42:7152–7161  
[Crossref]
- Santoro SW, Joyce GF (1997) A general purpose RNA-cleaving DNA enzyme. *Proc Natl Acad Sci U S A* 94:4262–4266  
[Crossref]
- Kim HK, Liu JW, Li J, Nagraj N, Li MX et al (2007) Metal-dependent global folding and activity of the 8-17 DNAzyme studied by fluorescence resonance energy transfer. *J Am Chem Soc* 129:6896–6902  
[Crossref]
- Mazumdar D, Nagraj N, Kim HK, Meng X, Brown AK et al (2009) Activity, folding and Z-DNA formation of the 8-17 DNAzyme in the presence of monovalent ions. *J Am Chem Soc* 131:5506–5515  
[Crossref]
- Luciano DJ, Vasilyev N, Richards J, Serganov A, Belasco JG (2017) A novel RNA phosphorylation state enables 5' end-dependent degradation in *Escherichia coli*. *Mol Cell* 67(44–54):e46
- Chakraborti S, Banerjea AC (2003) Identification of cleavage sites in the HIV-1 TAR RNA by 10-23 and 8-17 catalytic motif containing DNA enzymes. *Biomacromolecules* 4:568–571  
[Crossref]
- Ekesan S, York DM (2019) Dynamical ensemble of the active state and transition state mimic for the RNA-cleaving 8-17 DNAzyme in solution. *Nucleic Acids Res* 47:10282–10295  
[Crossref]
- Schlosser K, Li YF (2010) A versatile endoribonuclease mimic made of DNA: characteristics and applications of the 8-17 RNA-

- cleaving DNAzyme. *Chembiochem* 11:866–879  
[Crossref]
- 9. Pun SH, Tack F, Bellocq NC, Cheng J, Grubbs BH et al (2004) Targeted delivery of RNA-cleaving DNA enzyme (DNAzyme) to tumor tissue by transferrin-modified, cyclodextrin-based particles. *Cancer Biol Ther* 3:641–650  
[Crossref]
  - 10. Schlosser K, Gu J, Sule L, Li YF (2008) Sequence-function relationships provide new insight into the cleavage site selectivity of the 8-17 RNA-cleaving deoxyribozyme. *Nucleic Acids Res* 36:1472–1481  
[Crossref]
  - 11. Liu HH, Yu X, Chen YQ, Zhang J, Wu BX et al (2017) Crystal structure of an RNA-cleaving DNAzyme. *Nat Commun* 8:2006  
[Crossref]
  - 12. Chen YQ, Zhang J, Liu HH, Gao YQ, Li XH et al (2017) Unique 5'-P recognition and basis for dG:dGTP misincorporation of ASFV DNA polymerase X. *PLoS Biol* 15:e1002599  
[Crossref]
  - 13. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN et al (2000) The Protein Data Bank. *Nucleic Acids Res* 28:235–242  
[Crossref]
  - 14. Read RJ, Mccoy AJ, Storoni LC (2007) Likelihood-based molecular replacement in Phaser. *Nato Sci Ser II* 245:91–100
  - 15. Wang XH, Liu HF, Liu YW, Li YT, Yan L et al (2016) A novel strategy for the preparation of codon-optimized truncated Ulp1 and its simplified application to cleavage the SUMO fusion protein. *Protein J* 35(2):115–123  
[Crossref]
  - 16. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M (2006) HKL-3000: the integration of data reduction and structure solution-from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr* 62:859–866  
[Crossref]
  - 17. Potterton E, Briggs P, Turkenburg M, Dodson E (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr* 59:1131–1137  
[Crossref]
  - 18. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ et al (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* 58:1948–1954  
[Crossref]
  - 19. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132  
[Crossref]
  - 20. Fresco JR, Amosova O (2017) Site-specific self-catalyzed DNA depurination: a biological mechanism that leads to mutations and creates sequence diversity. *Annu Rev Biochem* 86:461–484  
[Crossref]
  - 21. Chandra M, Sachdeva A, Silverman SK (2009) DNA-catalyzed sequence-specific hydrolysis of DNA. *Nat Chem Biol* 5:718–720  
[Crossref]
  - 22. Cuenoud B, Szostak JW (1995) A DNA metalloenzyme with DNA ligase activity. *Nature* 375:611–614  
[Crossref]
  - 23. Wang W, Billen LP, Li YF (2002) Sequence diversity, metal specificity, and catalytic proficiency of metal-dependent phosphorylating DNA enzymes. *Chem Biol* 9:507–517  
[Crossref]

## 10. Solution NMR Spectroscopy as a Tool to Study DNAzyme Structure and Function

Jan Borggräfe<sup>1, 2</sup>✉ and Manuel Etzkorn<sup>1, 2, 3</sup>✉

- (1) Institute of Physical Biology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
- (2) Institute of Biological Information Processing, IBI-7: Structural Biochemistry, Forschungszentrum Jülich, Jülich, Germany
- (3) Jülich Center for Structural Biology (JuStruct), Forschungszentrum Jülich, Jülich, Germany

✉ **Jan Borggräfe**

Email: [jan.borggraefe@hhu.de](mailto:jan.borggraefe@hhu.de)

✉ **Manuel Etzkorn (Corresponding author)**

Email: [manuel.etzkorn@hhu.de](mailto:manuel.etzkorn@hhu.de)

### Abstract

Catalytically active DNA oligomers (or DNAzymes) offer a broad spectrum of functions as well as applications. Although known for over two decades, the DNAzyme's mode-of-actions are still poorly understood, mainly due to lack of high-resolution structural insights. Due to their molecular size, structural flexibility, and dynamic interactions with metal-ion cofactors, solution nuclear magnetic resonance spectroscopy (NMR) can serve as optimal tool to obtain mechanistic insights of DNAzymes. In this respect, nearly all states of the DNAzyme and its substrate during the catalytic cycle are accessible. The instructions and protocols provided in the following

may assist the initial steps of an NMR-based characterization of DNAzymes. To reduce the initial setup requirements and foster exciting new research projects, the discussed approaches focus on experiments that do not require cost-intensive isotope labeling strategies.

**Key words** DNAzymes – Nucleic acids – NMR spectroscopy – Homonuclear NMR – Metal-ion cofactors

---

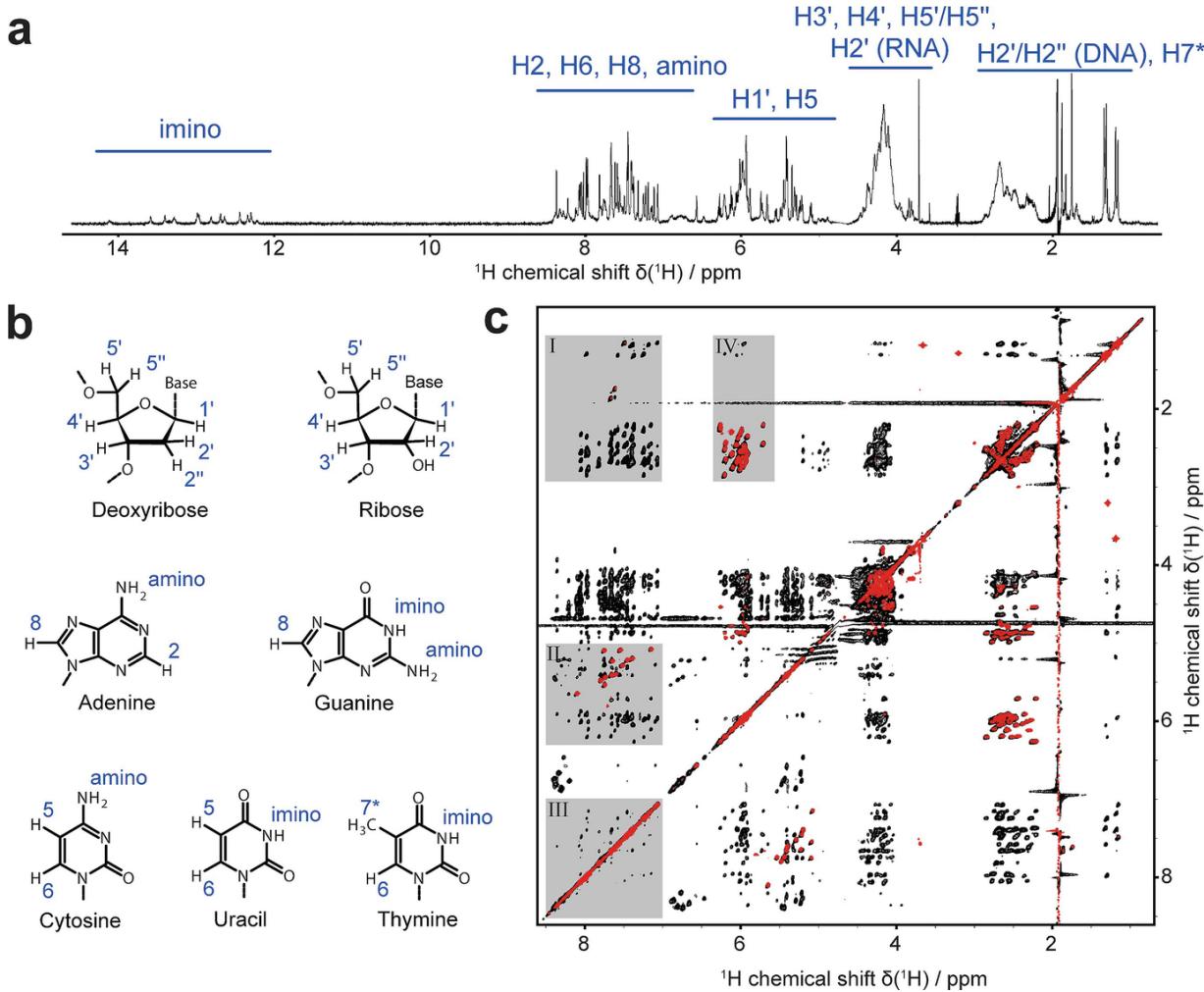
## 1 Introduction

DNAzymes (Dz) are artificially designed, catalytically active deoxyribonucleic acid oligomers with a length of up to a few dozens of nucleotides [1, 2]. They can facilitate a broad range of reactions, such as RNA/DNA cleavage [3, 4] or ligation [5, 6], peptide modifications [7–11], or thymine dimerization [12]. Due to their various functions, Dz became increasingly interesting for applications in bio-sensing, diagnostics [13–17], and therapy [18–23]. However, many aspects of their mode-of-action are unknown, mainly due to the lack of high-resolution structural information of the various Dz types and their respective individual states during the catalytic cycle.

So far, only three high-resolution snapshots of single states of Dz systems could be determined using X-ray crystallography [24–26] and none via other techniques such as nuclear magnetic resonance spectroscopy (NMR).<sup>1</sup> In general, NMR has a great potential to investigate structure, dynamics, and interactions of Dz in solution. Unlike to the previous crystallographic studies, NMR should also allow the investigation of dynamic equilibria as well as of transient and low populated states that may only occur in solution under physiologically relevant buffer and temperature conditions, as well as in the absence of artificial factors and modifications. NMR is capable of detecting even small changes of the environment of suitable nuclei. These sensors can thus directly report on structural changes and/or interactions with cofactors. With increasing knowledge of the accessible sensors (i.e., via resonance assignment of suitable nuclei), an increasing level of detail can be obtained. The following protocols describe the key steps of sample preparation, data acquisition, and data interpretation to initiate a new NMR-based characterization of a Dz system.

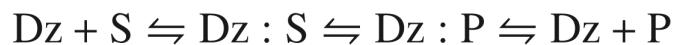
The presented methods focus on homonuclear  $^1\text{H}$ - $^1\text{H}$  NMR experiments, which can be performed without expensive isotope-enrichment and by using standard pulse programs. This should allow researchers without prior experience in RNA/DNA NMR to easily obtain initial results for their Dz projects. The discussed protocols assume a Dz system targeting other nucleic acids such as a DNA- or RNA-ligating/cleaving Dz. While many aspects should be independent of the specific Dz function, the high variety of Dz with different catalytic properties may enforce, in parts, an adaptation of the experimental setups.

Figure 1a shows a typical  $^1\text{H}$  NMR spectrum of a Dz system. In general, the proton chemical shifts of the ribose moiety in nucleotides are poorly dispersed due to their geminal hydroxyl group and cluster mainly in the region between 3.5 and 5.0 ppm. Fortunately, the absence of the 2' oxygen in DNA leads to a considerable up-field shift for the protons at this position, which clearly separates them from the other signals (Fig. 1a, see Fig. 1b for  $^1\text{H}$  nomenclature). In addition, cross peaks between H6 and H5 (or methyl H7\*) base protons of pyrimidines (cytosine, uracil, thymine) are clearly distinguishable from other signals in [ $^1\text{H}$ ,  $^1\text{H}$ ]-TOCSY spectra, which makes them ideal sensors to observe Dz states (Fig. 1c, area I+II, red spectrum, see also Note 1). This allows observing site-specific effects by standard solution NMR techniques even for complexes beyond 50 nucleotides.

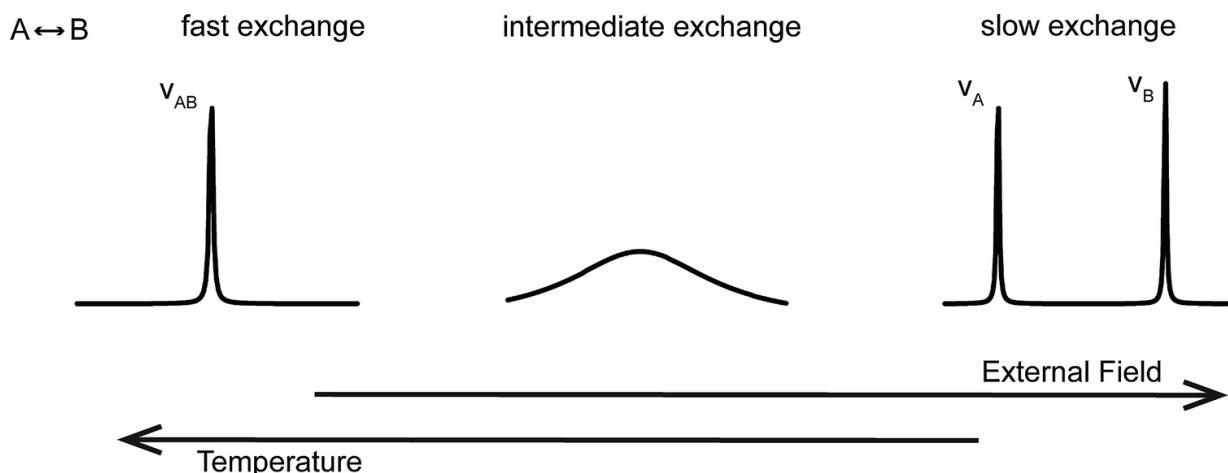


**Fig. 1** Proton chemical shifts in Dz systems. **(a)** One-dimensional  $^1\text{H}$  spectrum of a hybridized DNA:RNA complex with chemical shift ranges of non-exchangeable as well as amino and imino protons. Notably, imino proton signals are often only visible for base-paired nucleotides. **(b)** Lewis structures and proton nomenclature of (deoxy)ribose and nucleobases. **(c)** Two-dimensional  $[^1\text{H}, ^1\text{H}]$ -NOESY (black) and  $[^1\text{H}, ^1\text{H}]$ -TOCSY (red) spectra. The indicated areas I, II, III, and IV are used for assignment in Fig. 3a

Generally, Dz occur in at least three states during their catalytic cycle: (1) their single-stranded unbound form, (2) in complex with their substrate(s) (Dz:S), and (3) in complex with their product(s) (Dz:P), while the second and third state might consist of several (non-stable) intermediates.



Exchange processes between the different states (Fig. 2) as well as sample instability (see Note 2) might lead to inhomogeneity or peak broadening. Stabilization of different Dz states and improvements of spectral quality can be achieved by changing buffer conditions (see Note 3), sample concentration (see Note 4), or acquisition temperature (see Note 5). In general, pyrimidine signals serve as ideal sensors to study the Dz's behavior at different conditions [27]. Initially, perturbations of the chemical shift between Dz, Dz:S, as well as Dz:P can be observed, which can, e.g., give insights into complex formation and substrate cleavage. Detected chemical shift perturbations (CSPs) report on changes in the chemical environment around the observed protons and therefore can be induced by structural changes as well as cofactor binding. Analysis of the CSPs allows to locate regions with largest structural changes and/or binding sites. Following the CSPs via titration of cofactors (e.g., metal ions) can also enable determination of nucleotide-specific binding constants ( $K_D$ , Fig. 4b, left and middle, see Subheading 3.6). For cases in which cofactor interactions occur in the intermediate exchange regime, the observed line broadening effects (Fig. 4b, right) can still give valuable insights into binding dynamics (compare Fig. 2).



**Fig. 2** Chemical exchange processes between two Dz conformations A and B. Fast exchange rates lead to a singular signal and a homogenous spectrum (left), while in the intermediate exchange regime signals are severely broadened (middle) and in the slow exchange regime signals for multiple states are appearing (right). Variation of temperature and/or spectrometer field strengths may tip the scales towards either a homogenous, intense spectrum or two well-separated states. Both options are in general favorable over intermediate exchange effects

Divalent metal ions are often an essential cofactor for Dz-mediated catalysis. Due to its physiological abundance,  $Mg^{2+}$  is a frequent first choice for the characterization of the effects of divalent metal ions. However, the resulting CSP data derived from titration experiments cannot distinguish between direct binding or binding-induced changes in different regions of the Dz. The paramagnetic properties of  $Mn^{2+}$  provide an attractive alternative/addition (*see* Subheading 3.7 and Note 6). The proximity of  $Mn^{2+}$  induces paramagnetic relaxation enhancement (PRE) effects around the interacting atoms, leading to selective line broadening of proton signals close to the  $Mn^{2+}$ -binding sites (Fig. 5a). Another useful mimetic for hexahydrated  $Mg^{2+}$  is  $Co(NH_3)_6$ , which can also be utilized to locate divalent ion binding sites by observing direct Nuclear Overhauser Effect (NOE) contacts between the amine protons of cobalt hexamine and Dz protons [28] (*see* Subheading 3.8).

Overall, it can be summarized that NMR-resonance assignments of Dz systems in the range of up to 50–60 nucleotides are well feasible even in the absence of expensive isotope labeling strategies (*see* Subheading 3.5 and Note 7). Usage of homonuclear spectra can provide valuable information about occurring states, sample integrity, homogeneity, and presence of dynamic exchange processes. Several hundreds of atom-specific sensors are accessible and can further provide unique insights into the interaction with different cofactors or effects induced before, during, or after the catalysis reaction. While a detailed structure determination and/or characterization of nucleotide dynamics will require substantial work and financial means (*see* Note 8), initial structural parameters, such as base-pairing, secondary structure, and overall arrangement can already be generated with moderate amount of work and unlabeled samples.

---

## 2 Materials

1. Purchased oligonucleotides solubilized in deionized water to final concentrations of 1–3 mM.
2. Buffer stock solution with 2- to 10-fold concentration, e.g., 2-fold buffer 100 mM  $NaH_2PO_4/Na_2HPO_4$  pH 7, 200 mM NaCl and 10–20% (v/v)  $D_2O$  for deuterium lock.

3. 3 mm NMR tubes.
  4. Appropriate stock solutions of MgCl<sub>2</sub>, MnCl<sub>2</sub> (freshly prepared), and Co(NH<sub>3</sub>)<sub>6</sub>.
  5. Heating block (37–73 °C).
  6. High-field NMR spectrometer equipped with cryoprobe.
- 

## 3 Methods

### 3.1 Sample Preparation

1. Solubilize purchased oligonucleotides in deionized water to final concentrations of 1–3 mM.
2. Prepare buffer stock solution with 2- to 10-fold concentration, e.g., 2-fold buffer 100 mM NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub> pH 7, 200 mM NaCl and 10–20% (v/v) D<sub>2</sub>O for deuterium lock.
3. Prepare a 200 µl sample with the desired ratio of Dz and substrate, e.g., for a desired 200 µM Dz:RNA (1:1) complex, and using 1 mM oligomer stock solutions, and 2-fold buffer stock:
  - (a) 40 µl Dz stock solution.
  - (b) 40 µl substrate stock solution.
  - (c) 100 µl two-fold buffer stock solution.
  - (d) 20 µl ddH<sub>2</sub>O.
  - (e) Mix and spin down the sample solution.

4. Incubate the sample at 73 °C for 5 min in a heating block and let it cool down to room temperature afterwards. This procedure should destroy possible secondary structures of the single-stranded oligomers and allow reproducible complex formation.
5. Fill the sample into a clean 3 mm NMR tube and seal the tube.
6. Samples are usually stable over weeks at room temperature or 4 °C. However, note that decreased temperatures might induce a basic pH shift for several buffer systems, which can harm RNA substrates.
7. If higher D<sub>2</sub>O content is desired, lyophilize the sample overnight and re-suspend with the correct D<sub>2</sub>O/H<sub>2</sub>O ratio. Repeat **step 4** afterwards.

## 3.2 Data Acquisition

1. Wipe outside of sample tube clean, place it into the NMR magnet, and set temperature.
2. After temperature equilibration, lock, tune, and shim.
3. Calibrate <sup>1</sup>H 90° hard pulse.
4. Create new TOCSY and NOESY experiment with desired pulse program (e.g., *dipsi2esgpph* and *noesyegpph* on Bruker instruments) and calculate/calibrate essential pulses.
5. Set acquisition parameters as suggested in Table 1 and start experiment(s).
6. Process data with appropriate zero-filling, window function, and linear prediction.

**Table 1** Suggested NMR acquisition parameters for sample with Dz concentration in the range of 200–1000 μM, and using a cryoprobe setup at 600–900 MHz

Experiment	Spectral width/ppm	Data points	Number of scans	Mixing time/ms
------------	--------------------	-------------	-----------------	----------------

Experiment	Spectra F1		Data points	F1	Number of scans	Mixing time/ms
	F2	F1				
<i>Initial experiments</i>						
[ <sup>1</sup> H, <sup>1</sup> H]-TOCSY	20	10	2048	256	8–16	80
<i>High-resolution spectra</i>						
[ <sup>1</sup> H, <sup>1</sup> H]-TOCSY	20	10	2048	512	≥16	80
[ <sup>1</sup> H, <sup>1</sup> H]-NOESY	20	10–20	2048	1024	≥32	300

### 3.3 Initial Experiments

First, it should be determined how good the Dz states are accessible for NMR analysis. Therefore, [<sup>1</sup>H,<sup>1</sup>H]-TOCSY spectra of all stable Dz states should be recorded. This includes free Dz, stabilized Dz:S and Dz:P, as well as of free substrate. To generate Dz:S, a stabilized variant of the substrate or the absence of (metal ion) cofactors may be required (*see Note 2*). A stable Dz:P complex may require high sample concentrations and/or low temperatures. However, alternatively also non-hybridized Dz + P can offer interesting insights. All spectra should be recorded in the same buffer and the same acquisition setting. Pyrimidine cross peaks in each spectrum, which are appearing around 7.0–8.5 ppm for H6 and 5.0–6.5 ppm for H5 (if cytosine or uracil) or 0.5–2.0 ppm for H7\* (if thymine) (compare Fig. 1c, I+II, red spectrum), should be counted, since they give direct insights into sample integrity and exchange processes (*see Notes 9 and 10* for detailed discussion). Sample and/or acquisition parameters might need to be altered to yield a homogenous spectrum (*see Notes 3–5*).

### 3.4 High-Resolution Experiments

Once conditions leading to a homogenous spectrum with maximal signal dispersion could be established, high-resolution [<sup>1</sup>H,<sup>1</sup>H]-TOCSY and [<sup>1</sup>H,<sup>1</sup>H]-NOESY spectra should be acquired for the Dz complex at higher concentrations as well as for the free Dz and substrate. To identify

exchangeable protons and, if required, to reduce solvent signal, high-resolution spectra can additionally be recorded in 100% D<sub>2</sub>O. Normally, the same sample can be recovered after lyophilization and re-solubilization in D<sub>2</sub>O. To facilitate analysis and increase level of details, spectra should be recorded with highest possible spectral resolution and signal-to-noise ratio, so several days of measurement time on a high-field spectrometer should be planned for. The resulting spectra should be used to obtain a comprehensive resonance assignment.

### 3.5 Sequential Assignment

In the following, a protocol for sequential resonance assignment is described, which uses several NOE contacts between adjacent nucleotides to resolve ambiguities. This protocol works robustly for stacked nucleotides such as present in A- and B-form double helices, and should also be well transferable to unknown secondary structures. While some NOE signals might be absent, weak, or overlapping, the usage of different spectral areas may still enable full sequential resonance assignment. To start with the sequential assignment, one should initially focus on the spectral regions of 7.0–8.5 ppm in the direct dimension (H<sub>6</sub> and H<sub>8</sub> base protons) and 5.0–6.5 ppm (H<sub>1'</sub> ribose protons and H<sub>5</sub> base protons) and 0.5–3.0 ppm in the indirect dimension (H<sub>2'</sub> and H<sub>2''</sub> deoxyribose protons, H<sub>7\*</sub> thymine methyls, *see* Figs. 1c and 3a). An extensive description of nucleotide assignment is published [29] and additional information can be found in Note 7.

1. Load the recorded high-resolution [<sup>1</sup>H,<sup>1</sup>H]-TOCSY and [<sup>1</sup>H,<sup>1</sup>H]-NOESY spectra in an appropriate NMR analysis software.
2. Open the NOESY spectrum and overlay the TOCSY spectrum.
3. For sequential assignment, concentrate on the highlighted spectral areas shown in Fig. 1c.
4. Draw vertical lines through areas I and II connecting the most prominent cross peaks and define them as different spin systems (*see* Fig. 3a).  
Some strong NOESY cross peaks are overlapping with TOCSY cross

5. peaks, which identifies pyrimidine spin systems. A thymidine will show a TOCSY cross peak in area I, and a cytidine in area II (marked with arrows in Fig. 3a).
6. In the best case, four additional peaks on the respective vertical line can be found in area I, which represent cross peaks between base proton H6 or H8 and deoxyribose protons H2'/H2'' of the own and the 5' neighboring sugar moiety. In addition, usually two signals on the vertical line are visible in area II belonging to cross peaks between H6 or H8 and H1' of the own and the 5' neighboring sugar moiety.
7. Normally, cross peaks between H6/H8 and sugar protons within the same nucleotide (same spin system) appear stronger than between neighboring nucleotides. However, to identify them unambiguously, draw vertical lines from the respective NOESY cross peaks in area I towards area IV. If two cross peaks, which are vertically matching in area I, are also vertically matching in area IV, they belong to the same nucleotide (and spin system). If a line is drawn via the matching diagonal position, also the correct H1' frequency can be identified. This is shown in Fig. 3a for spin systems #1, #2, and #4.
8. Next, assign spin systems sequentially. From the cross peaks on the vertical line of a spin system that have been identified as not belonging to this spin system, draw horizontal lines to find other matching peaks, which are assigned to another spin system. These cross peaks normally belong to a spin system of the 5' neighboring nucleotide. Additionally, stacking contacts between base protons H6 or H8 of neighboring nucleotides can be utilized, as seen in Fig. 3a, area III. Also, protons H5 or methyl H7\* of pyrimidines form contacts with H6 or H8 base protons of their 3' neighbor.
9. Using these connections, a sequential walk can be realized. The principle is shown in Fig. 3a for spin systems #1 to #10, which are connected to each other in 3' to 5' direction. In the shown example, spin systems #1, #3, and #7 could be identified as thymidines, and #5 and #10 as cytidines. The assigned spin system sequence can then be written as

Sequence:	5'- #10 #9 #8 #7 #6 #5 #4 #3 #2 #1 -3'
Nucleotide:	C X X T X C X T X...T

with X being a purine. The used DNA:RNA complex in this example is shown below, with the sequence matching to the assigned spin systems underlined.

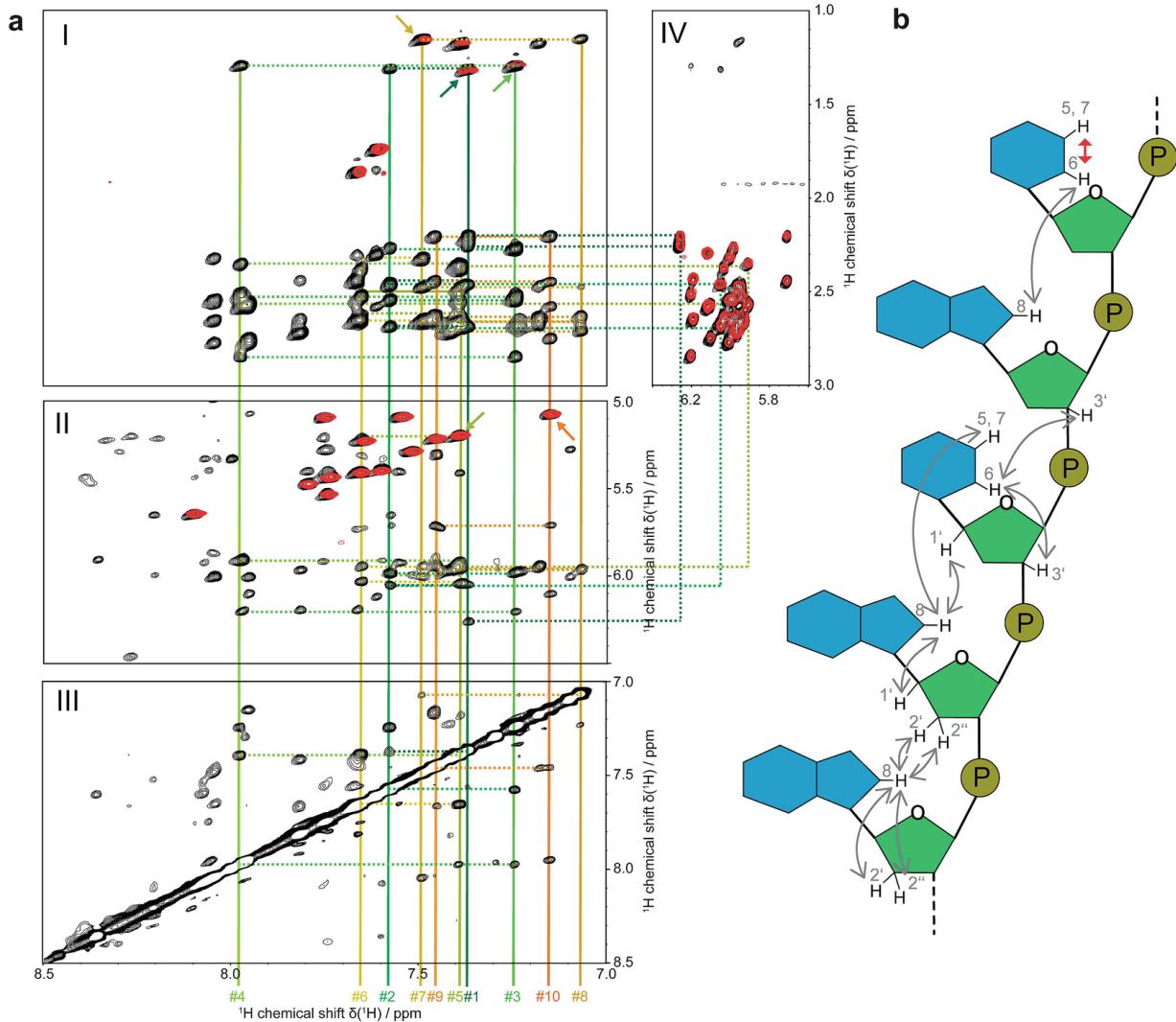
DNA:	5' -TTGGGG <u>TAA<u>CGGTGCATGT</u>-3'</u>
RNA:	3' -AAC <u>CCCCAUUGCCACGUACA</u> -5'

10.

In case the Dz targets an RNA substrate, perform RNA assignment after full Dz assignment. RNA is lacking the characteristic dispersion of H2'/H2'' chemical shifts and assignment relies mainly on H1' cross peaks in area II.

11.

H3' assignment is also often possible using the same strategy; however, assignment of H4', H5'/H5'' and ribose H2' can be very challenging using homonuclear spectra.



**Fig. 3** Sequential resonance assignment strategy for DNA. **(a)** Spectral regions indicated in Fig. 1c showing NOE contacts to one or two adjacent nucleotides (black spectrum). The shown patterns involve base protons H6 or H8 and deoxyribose protons H2'/H2'' as well as thymine methyl protons H7\* (area I), ribose H1' and cytosine/uracil protons H5 (area II), and inter-nucleotide contacts between H6 or H8 of two different bases (area III). The vertical lines are crossing the diagonal at chemical shift position of H6 or H8 (area III) and the corresponding spin system (#1 to #10) is labeled. The NOESY cross peaks along these lines either correspond to the same spin system (intra-nucleotide NOEs) or represent through space inter-nucleotides contacts (usually involving the direct neighbor). Matching peaks on different vertical lines indicate sequential connections (horizontal dashed lines in areas I–III). For pyrimidines, the vertical line is also crossing a TOCSY cross peak in area I or II (red spectrum, indicated by arrows). The TOCSY peak pattern in area IV (red spectrum) can be used to determine if H1' and H2'/H2'' belong to the same spin system. **(b)** The scheme displays expected NOE contacts between protons of two adjacent nucleotides

(gray arrows). The red arrow indicates the TOCSY transfer between H6 and H5 (or H7\*) in pyrimidines

### 3.6 NMR-Titration Experiments

Dz often require mono- or divalent cations as cofactors, either to stabilize a catalytically relevant structure or as a reaction partner itself (see [30] for a review). Titration with increasing concentrations of metal ions can give nucleotide- or even atom-specific information about binding events (Fig. 4).

To perform titration experiments with metal ions (e.g.,  $Mg^{2+}$ ), a single suitable sample (e.g., 200  $\mu l$  of 200  $\mu M$  Dz or Dz:S complex) and a series of different metal-ion stock solutions are normally sufficient. We, e.g., obtained decent titration results by iteratively adding 2  $\mu l$  of each stock solution on the upper wall inside the sample tube, closing the tube tightly, and mixing the drop by shaking the solution up and down the tube several times. This method gave highly reproducible results for metal-ion cofactors with negligible/calculable dilution effects.

1. Prepare a 200  $\mu l$  sample of 200  $\mu M$  Dz or Dz:S complex in suitable buffer conditions in absence of the titrant.
2. Prepare a series of stock solutions of titrant in a way that stepwise titration with a fixed volume of each stock leads to the correct end concentration, e.g., for stepwise addition of 2  $\mu l$  in 200  $\mu l$  starting volume with desired end concentrations of 0.25, 0.5, 1, 2, 5, 10, 20, and 50 mM, the prepared stock concentrations should be 25.25, 25.5, 51.5, 104, 315, 530, 1070, and 3240 mM.
3. Fill the sample into a 3 mm NMR tube, place the tube in the spectrometer, and calibrate relevant parameters for a [ $^1H, ^1H$ ]-TOCSY experiment.
4. Set number of scans to allow for approx. 5–12 2D experiments with sufficient signal-to-noise and good resolution in the indirect dimension. For example, with our setup, 2D spectra with 16 scans and 256 indirect increments resulted in decent spectra for a 20 kDa Dz:S complex within 2 h at 37 °C and 600 MHz. However, keep the experimental parameters constant over the whole titration series.  
After recording the initial reference experiment in the absence of

5. After recording the initial reference experiment in the absence of titrant, place a 2  $\mu$ l drop of the titrant stock with lowest concentration on the inner wall at the upper part of the NMR tube and close the tube tightly with the lid. Shake or spin the sample solution 5–10 times up and down the tube and start the next NMR acquisition. Repeat the step iteratively with increasing stock concentrations until the titration is done.
6. Longer acquisition periods (overnight, over weekend) can be used to acquire [ $^1\text{H}$ ,  $^1\text{H}$ ]-NOESY spectra to ease assignments and to record longer [ $^1\text{H}$ ,  $^1\text{H}$ ]-TOCSY spectra, which might be necessary if signals “broaden away” during titration.
7. Process all recorded spectra with appropriate zero-filling, window function, and linear prediction. However, keep processing parameters consistent for all experiments.
8. For assignment, overlay the peak list of your high-resolution spectra with the titration spectra recorded under conditions closest to the one used for assignment. If titration steps are small enough, peak shifts can easily be followed during the titration. Additionally, recorded NOESY spectra can resolve ambiguous assignments.
9. For data analysis, first extract relative peak positions and peak intensities of pyrimidine H5-H6 cross peaks (compare Fig. 4a). Calculate relative peak positions using the combined chemical shift (CCS) between proton shifts of the initial spectrum ( $\delta(\text{H5})_{\text{init}}$ ,  $\delta(\text{H6})_{\text{init}}$ ) and of each titration spectrum ( $\delta(\text{H5})$ ,  $\delta(\text{H6})$ ):

$$\text{CCS} = \sqrt{(\delta(\text{H5})_{\text{init}} - \delta(\text{H5}))^2 + (\delta(\text{H6})_{\text{init}} - \delta(\text{H6}))^2} \quad (1)$$

10. Plotting the CCS of each peak against the titrant concentration and fitting with Eq. (2) yields nucleotide-specific  $K_D$  values (see Fig. 4b, I+II):

$$\text{CCS} = \frac{\text{CCS}_{\text{max}} \times [\text{titrant}]}{K_D + [\text{titrant}]} \quad (2)$$

with CCS being the CCS at saturation

with  $I_{\text{max}}$  being the intensity at saturation.

11.

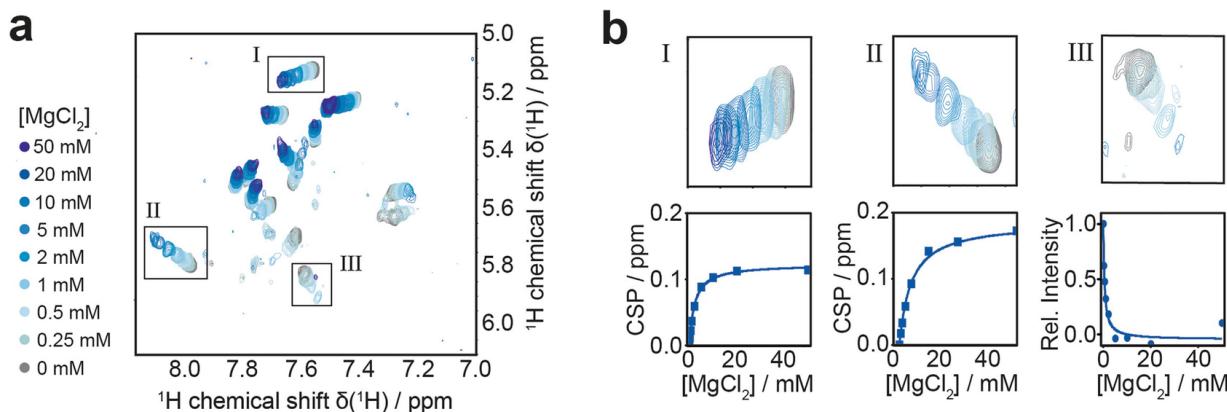
Some titrants lead to line broadening of peaks due to intermediate exchange processes of binding events. By increasing the acquisition time at higher titrant concentrations, peaks might still become visible. While  $K_D$  values can be roughly estimated by plotting the peak intensity against the titrant concentration and fitting with Eq. (3) (see Fig. 4b, III):

$$I = I_{\text{max}} - \frac{I_{\text{max}} \times [\text{titrant}]}{K_D + [\text{titrant}]} \quad (3)$$

with  $I_{\text{max}}$  being the intensity of the initial experiment, the exchange effects obscure a reliable  $K_D$  calculation with this method.

12.

While H5-H6 correlations are usually best resolved and easiest to follow in titration experiments, if required, CSP information can also be extracted for all other individually resolved cross correlation to get atom-specific binding information.



**Fig. 4** NMR-based  $\text{Mg}^{2+}$  titration experiment. (a) Overlay of  $[^1\text{H}, ^1\text{H}]$ -TOCSY spectra (cytosine/uracil signals) recorded for a 10–23 DNAzyme in complex with stabilized RNA substrate at different  $\text{MgCl}_2$  concentrations. Clear chemical shift perturbation (CSP) and signal broadening are visible. (b) Enlarged view of the peaks indicated in a) (top). Plotting of CSPs (bottom left and middle) or relative intensities (bottom right) against  $\text{MgCl}_2$  concentration and fitting with Eqs. (2) and (3) gives estimates for nucleotide-specific  $K_D$  values. (Data refer to experimental results reported in [27])

### 3.7 Titration Experiments with Mn<sup>2+</sup>

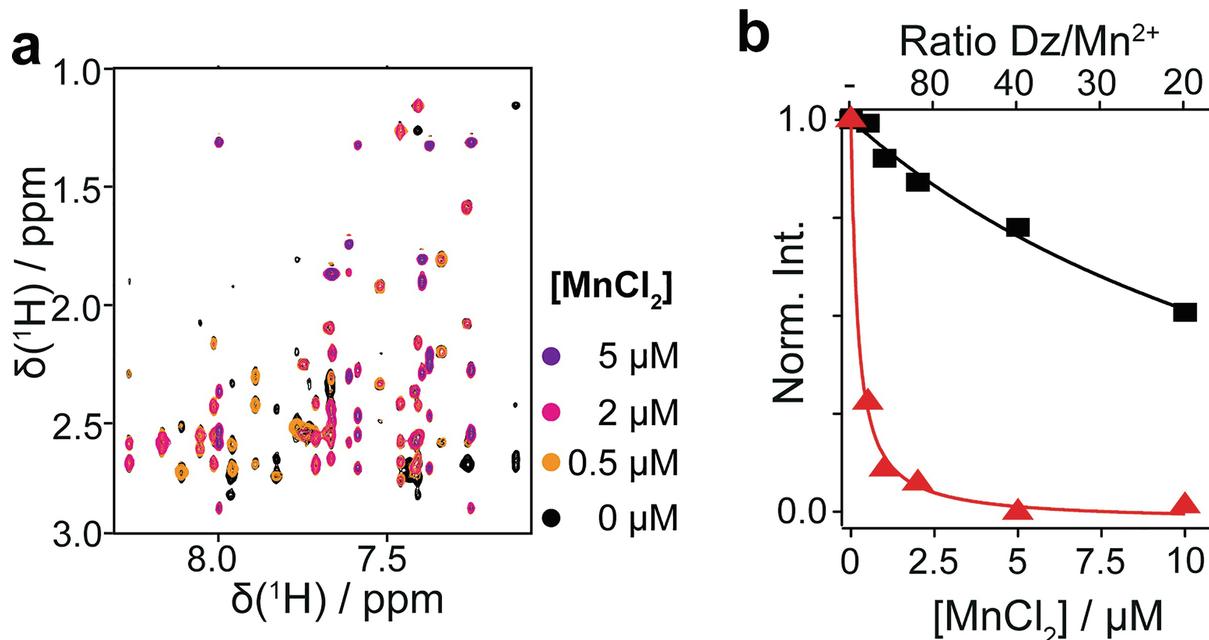
The protocol below describes the experimental setup of a Mn<sup>2+</sup> titration with prior Mg<sup>2+</sup> equilibration to map divalent metal-ion binding sites with atomic resolution (see also **Note 6**).

1. Prepare 200 µl Dz or Dz:S sample in the appropriate buffer preferably in presence of MgCl<sub>2</sub>. The Mg<sup>2+</sup> concentration should be considerably higher (e.g., 1 mM in this example) than the maximal MnCl<sub>2</sub> concentration. Perform resonance assignments of a spectrum under the chosen condition in absence of MnCl<sub>2</sub> beforehand.
2. Prepare a series of stock solutions of MnCl<sub>2</sub> in a way that stepwise titration with a fixed volume of each stock leads to the correct end concentration, e.g., for stepwise addition of 2 µl in 200 µl starting volume with desired end concentrations of 0.5, 1, 2, 5, and 10 µM, the prepared stock concentrations should be 50.5, 51, 103, 312, and 525 µM. (Note that Mn<sup>2+</sup> is often not stable under the applied buffer conditions and hence should be always prepared freshly and added directly before starting the measurements.)
3. Fill the sample into a 3 mm NMR tube, place the tube in the spectrometer, and calibrate relevant parameters for a [<sup>1</sup>H,<sup>1</sup>H]-NOESY and/or [<sup>1</sup>H,<sup>1</sup>H]-TOCSY experiment.
4. Set number of scans to allow for approx. 5–12 2D experiments with sufficient signal-to-noise and good resolution in the indirect dimension. For example, with our setup, 2D spectra with 16 scans and 256 indirect increments resulted in decent spectra for a 20 kDa Dz:S complex within 2 h at 37 °C and 600 MHz.
5. After recording the initial experiment, place a 2 µl drop of the MnCl<sub>2</sub> stock with lowest concentration on the inner wall at the upper part of the NMR tube and close the tube tightly with the lid. Shake or spin the sample solution 5–10 times up and down the tube and start the next NMR acquisition. Repeat the step iteratively with increasing stock concentrations until the titration is done.

6. Process all recorded spectra with appropriate zero-filling, window function, and linear prediction. Keep processing parameters constant for all experiments.
7. Open the initial spectrum with an appropriate NMR analysis software and import the assigned peak list. Remove all overlapping peaks as well as diagonal peaks. Open the remaining spectra of the titration experiment, import the newly created peak list, and integrate all peaks in all spectra.
8. Plotting peak intensities against  $\text{Mn}^{2+}$  concentration and fitting with Eq. (4) yields peak-specific PRE rates (see Fig. 5b).

$$I = e^{-k \times [\text{Mn}^{2+}]} \quad (4)$$

9. To extract proton-specific PRE rates, average PRE rates of all cross peaks involving the same proton.



**Fig. 5**  $\text{Mn}^{2+}$ -induced paramagnetic relaxation enhancement. (a) Section of  ${}^1\text{H}, {}^1\text{H}$ -NOESY spectra recorded for a 10–23 DNAzyme in complex with stabilized RNA substrate in presence of 1 mM  $\text{MgCl}_2$  and indicated increasing concentrations of  $\text{MnCl}_2$ . (b) Normalized cross-peak intensities of a peak with fast (red) and a slow (black) relaxation

rate. Faster relaxation rates indicate closer proximity to the Mn<sup>2+</sup> interaction site. Data refer to experimental results reported in Borggrøfe et al. [29]

### 3.8 Titration Experiments with Co(NH<sub>3</sub>)<sub>6</sub>

Amines of Co(NH<sub>3</sub>)<sub>6</sub> can lead to a strong background signal, which can be reduced to a tolerable level by using an increased deuteration level of around 80%. Acquisition temperature might need to be lowered to slow down exchange of the amine protons of Co(NH<sub>3</sub>)<sub>6</sub> with the solvent.

Experiments with Co(NH<sub>3</sub>)<sub>6</sub> should also be performed as a titration with 5–10 fold excess of Co(NH<sub>3</sub>)<sub>6</sub> at the final step. The appearing Co(NH<sub>3</sub>)<sub>6</sub>-derived NOESY cross peaks are then better distinguishable from noise, baseline distortions, and impurities.

1. Prepare 200 µl Dz or Dz:S sample in the appropriate buffer preferably in 80% (v/v) D<sub>2</sub>O/H<sub>2</sub>O (see Subheading 3.1).
2. Prepare a series of stock solutions of Co(NH<sub>3</sub>)<sub>6</sub> in a way that stepwise titration with a fixed volume of each stock leads to the correct end concentration, e.g., for stepwise addition of 2 µl in 200 µl starting volume with desired end concentrations of 200, 500, and 1000 mM, the prepared stock concentrations should be 20.2, 30.6, and 51.5 mM.
3. Fill the sample into a 3 mm NMR tube, place the tube in the spectrometer, and calibrate relevant parameters for a [<sup>1</sup>H,<sup>1</sup>H]-NOESY experiment.
4. Set number of scans to record experiments with sufficient signal-to-noise and resolution in the indirect dimension (Note that amine signals of Co(NH<sub>3</sub>)<sub>6</sub> are usually well-separated from other signals, therefore resolution in indirect dimension is not crucial). We recorded 2D spectra with 16 scans and 256 indirect increments for a 20 kDa Dz:S complex within 2 h at 10 °C and 600 MHz.
5. Process all recorded spectra with appropriate zero-filling, window function, and linear prediction. Keep processing parameters constant for all experiments.

6. Open and overlay all recorded spectra with an appropriate NMR analysis software. Amines of  $\text{Co}(\text{NH}_3)_6$  show a strong signal around 3.38 ppm and potential NOE cross peaks with the Dz's protons should appear at this frequency. Cross peaks are best identified in less crowded spectral areas, e.g., of the H6/H8 base protons around 7.0–8.5 ppm.
  7. The use of several spectra with different  $\text{Co}(\text{NH}_3)_6$  helps to distinguish appearing cross peaks between baseline distortions. Alternatively, one can displace  $\text{Co}(\text{NH}_3)_6$  with an excess of  $\text{Mg}^{2+}$  leading to a disappearance of  $\text{Co}(\text{NH}_3)_6$ -derived cross peaks.
- 

## 4 Notes

1. Some Dz sequences have variable parts which (presumably) do not have an impact on its catalytic mode-of-action. For instance, the substrate-specific binding arms of DNA- or RNA-ligating or cleaving Dz can be chosen rather freely to target virtually any substrate sequence [3–6]. This may offer the opportunity to consider NMR properties before/during the selection process of the desired Dz system. Introduction of well-distributed pyrimidine nucleotides in these parts can, e.g., ease assignment and serve as useful sensors for later titration experiments. Repeating strings of the same nucleotides should be avoided since this might lead to very similar chemical shifts (and less signal dispersion) in the resulting spectra.
2. Inhomogeneous spectra and peak broadening might be caused by sample instability. Single-stranded DNA oligomers are usually stable within a wide range of pH and temperatures. However, RNA substrates are very sensitive to a basic environment and RNase contamination, which leads to fragmented short nucleotide oligomers. To access Dz:S via NMR often requires experiments of several hours or days, and one has to stabilize the substrate-bound state, e.g., in absence of essential cofactors (such as divalent ions for RNA-cleaving DNzyme [4]) or using non-reactive substrate mimetics. If chosen wisely, the latter may enable studies of “activated” Dz states. However, also non-native artifacts may be introduced. Previous

examples include the crystallographic study of the 8–17 DNAzyme in complex with a biological inactive DNA target [25] or NMR studies of the 10–23 DNAzyme with an RNA target stabilized by a single fluorine substitution at the RNA's cleavage site [27]. Sample integrity should be checked at least once before and after NMR acquisition by another independent method such as polyacrylamide gel electrophoresis (PAGE) [27]. Once a spectrum can be linked to PAGE results, the NMR data will normally be a very reliable readout of the different states of the sample. In this respect also long-term stability and functionality of substrate ligating or cleaving Dz can and should be checked.

3. Buffer conditions rely strongly on the specific needs of the respective Dz system, but acidic to neutral pH and low to medium ionic strength is preferable in NMR spectroscopy. However, very low ionic strength should be avoided since DNA/RNA complexes usually need an elevated number of cationic counterions for full hybridization and stability, and suitable conditions should be identified in initial experiments. Obviously, buffer systems with overlapping signals in the  $^1\text{H}$ -spectra should be avoided. However, fully deuterated alternatives such as Tris-d11 are available at low costs. Ionic cofactors are often essential for the Dz's function and, if required, their optimal concentrations should be carefully screened in initial experiments. A suitable buffer for initial experiments might be 50 mM  $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$ , 100 mM NaCl, 10% (v/v)  $\text{D}_2\text{O}$ , pH 7.
4. A high sample concentration is crucial for acquisition of decent NMR signals in multidimensional experiments. Traditionally most reported NMR experiments on nucleotide samples have been performed using DNA/RNA concentrations of 1 mM or higher. However, structural and functional studies by other methods (such as fluorescence techniques or electrophoresis) apply far lower concentrations. Therefore, properties derived from these methods cannot always directly be transferred to the NMR setup. Note that the high concentrations will change the effective dissociation temperatures of the hybridized nucleotides, and may also induce artificial (super-)complexes via palindromic regions in the Dz sequence [26]. Therefore, in addition to reducing the costs of the samples, one should generally consider to work with not too high Dz concentrations. In our

~~generally consider to work with not too high Dz concentrations. In our experience we could, e.g., acquire decent [<sup>1</sup>H,<sup>1</sup>H]-TOCSY spectra within 2 h acquisition time using a Dz sample concentration in the range of 200 μM (at 600–800 MHz spectrometer equipped with a cryoprobe).~~

5. In the absence of exchange processes, higher acquisition temperatures normally directly translate to reduced linewidths and increased signal intensities in the resulting spectra. However, for hybridized complexes the maximum temperature is usually limited by the stability of the nucleotide complex. The melting temperature of nucleic acid-cleaving or -ligating Dz complexes can be altered by length and basepair-composition of their substrate-binding arms [31, 32]. For therapeutic relevant RNA-cleaving Dz the substrate-specific sequence is usually chosen to allow hybridization with its substrate and dissociation with the cleaved products below 37 °C using Dz concentrations in the nM range [33]. It should be taken into account that the melting temperatures are increased at concentrations usually used in NMR experiments. Higher temperatures also accelerate the exchange of imino/amino protons with the solvent, which might render their signals invisible. Acquisition temperature of 10 °C or below highly improves their visibility, but severely broadens the overall linewidth. It is in general advisable to record spectra at different temperatures to optimize spectral properties for the desired readout.
6. Often only Mn<sup>2+</sup> concentrations in the low μM range can be used for Mn<sup>2+</sup> titration experiments since higher concentrations lead to too severe signal quenching of the whole sample. However, Mn<sup>2+</sup> titration can be carried out in the presence of an excess of a constant concentration of, e.g., Mg<sup>2+</sup>. The Mn<sup>2+</sup> ions are then competing with Mg<sup>2+</sup> and can, due to fast exchange processes, induce PRE effects without inducing structural changes (CSPs) (Fig. 5a). This enables a much more direct determination of binding sites. To gain maximal insights, recording [<sup>1</sup>H,<sup>1</sup>H]-NOESY spectra should be preferred over [<sup>1</sup>H,<sup>1</sup>H]-TOCSY spectra. To evaluate the PRE data, the peak intensity of each peak in the different spectra is fitted with a simple exponential decay (Fig. 5b). If an atom-specific analysis is desired, the average decay rate of all cross peaks involving the same proton should be calculated

calculated.

7.

Resonance assignments using homonuclear [ $^1\text{H}$ ,  $^1\text{H}$ ] spectra mainly rely on NOE contacts between neighboring nucleotides (Fig. 3). Since NOEs are depending on proton-proton distances, this approach is highly depending on the DNA/RNA secondary structure. To avoid spectral overlap and distortions, the sample purity is important and can simplify analysis and maximize readout. Sample purity may considerably differ from vendor to vendor. Therefore, the ordered samples should always be checked for impurities such as remaining organic solvents. Changing the vendor or requesting a more thorough purification strategy may largely simplify future work on the system. In case Dz:S complexes should be assigned, spectra of free Dz and substrate should be recorded beforehand to identify signals originating from incomplete complexation or small pipetting errors.

Exchangeable protons, such as imino and amino base protons, are usually not interfering with the assignment procedure, but one can record additional spectra in 100%  $\text{D}_2\text{O}$  to eliminate these peaks.

Assignment of imino signals can provide additional information [34]; however, this requires low temperatures, which may complicate subsequent interaction/functional studies.

8. Most experiments to more thoroughly study structure and dynamics require cost-intensive isotope labeling strategies and feasibility should be evaluated beforehand. With sufficient sample concentration and acquisition time, initial two-dimensional heteronuclear experiments can also be recorded using the natural abundance of  $^{13}\text{C}$ , but this is not feasible for higher-dimensional experiments. Looking on a [ $^1\text{H}$ ,  $^{13}\text{C}$ ]-HSQC spectrum of nucleic acids (Fig. 6), most signals strongly overlap as also seen for homonuclear spectra. However, particularly the C1'-H1' cross correlations often show significantly high dispersion suitable for first resonance assignments. These peaks can subsequently be used as sensors for acquisition of structural parameters such as residual dipolar couplings (RDC) or dynamic experiments such as T1/T2 relaxation or HetNOE experiments (see [35] for an overview of experiments on nucleic acid dynamics).

However, in these cases  $^{13}\text{C}$ -enrichment is normally necessary to obtain reliable data within realistic experimental times.

The dilute proton network paired with the limited dispersion, complicates NMR-based structure determination purely based on semi-quantitative NOE contacts. An alternative to generate long-distant structural information, which is also applicable for non-isotope enriched samples, is the introduction of paramagnetic spin labels. The PRE effect can generate a larger number of long-range distance restraints that can provide valuable information about the overall structural arrangements of the Dz at lower resolution [36]. Dihedral angular information of nucleotides is also often extracted from scalar coupling constants, e.g., by usage of the E.COSY experiments [37]. However, this approach can be obstructed by severe signal overlap in the recorded spectra when applied to larger nucleotide systems. We also observed strong limitations of conventional  $^{31}\text{P}$ -based experiments that, unlike to high concentrated and smaller test systems, did not provide useful results for the relevant 20 kDa Dz system at the used conditions. Alternatively, intra-nucleotide NOEs and NOE contacts between adjacent nucleotides can be used to determine secondary structure elements such as base-sugar configurations and sugar puckering. Since accurate quantification of these “short-range” NOEs is not feasible using data of only one NOESY spectrum, one usually records NOE peak build-up curves from spectra with different mixing times, e.g., between 20 and 400 ms. To accommodate relaxation effects from surrounding protons, suitable software like eNORA [38] is able to extract NOE build-up rate constants and convert them into more accurate distance restraints also known as exact NOEs (eNOEs) [39]. However, this approach requires a good estimate of the structure as a starting condition.

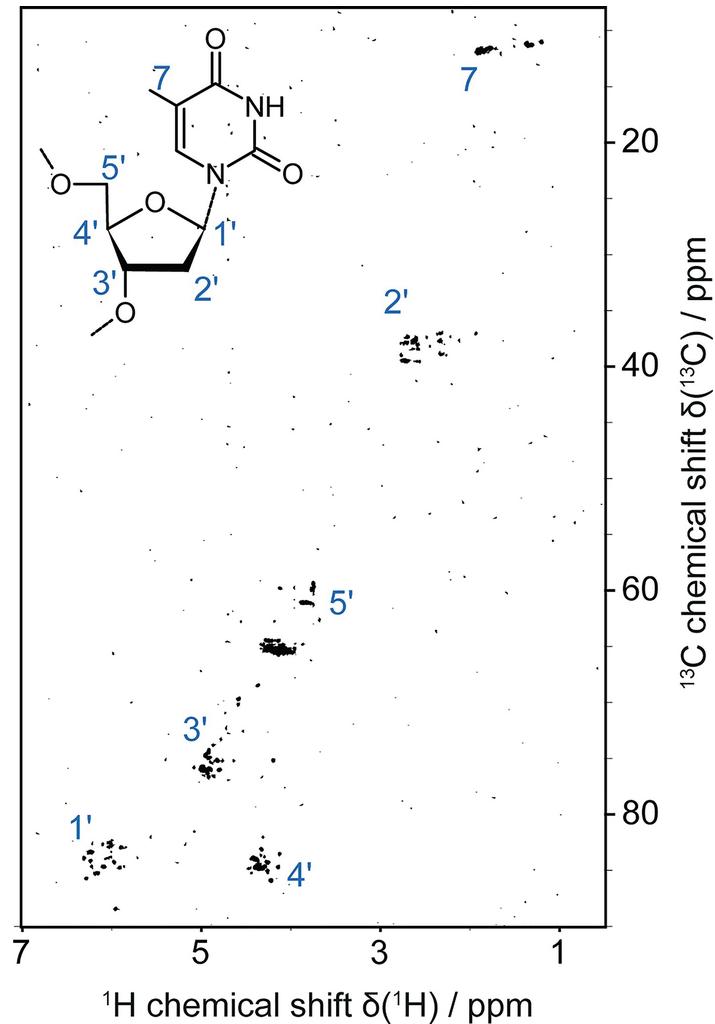
9. If more pyrimidine cross peaks are appearing in the spectrum than there are pyrimidines in the sequence(s), the sample might be in a condition containing several slowly interchanging states (Fig. 2, right). In case of complexed Dz:S or Dz:P, these peaks might also arise from free, single-stranded oligomers, which can be verified by comparison to the spectra recorded of the respective free Dz or substrate. On the other hand, free or complexed Dz itself might exist in two or more stable conformations, which lead to a heterogeneous

peak pattern for the pyrimidine cross peaks. These conformations might already provide relevant details of the Dz's mode-of-action or could also be artificially induced by non-suitable sample conditions, e.g., high sample concentration or too low ionic strength. Alteration in

ionic strength, but also pH, temperature or the addition of cofactors might lead to sample homogenization.

10.

If less pyrimidine cross peaks are counted in the spectrum than there are pyrimidines in the sequence(s), one should carefully check peak overlap in both dimensions. To consider also weaker peak intensities and maximize spectral resolution, it may be useful to repeat the experiment with increased number of scans and elongated indirect evolution times. The latter can also be realized via non-uniform sampling approaches. In addition, either conservative or progressive spectral processing may also help to identify weak peaks or spectral overlap, respectively. If peaks are still missing or appearing with low intensity and broad linewidth, the sample might be affected by intermediate exchange processes (Fig. 2, middle). Exchange processes can be altered into a fast or slow regime by changing acquisition temperature or external field strength. However, also binding of (divalent metal ion) cofactors often occur in the intermediate exchange regime (see Fig. 4). The aim of the initial experiments is to find most suitable conditions, which lead to a homogenous spectrum with optimal signal dispersion.



**Fig. 6**  $[{}^1\text{H}, {}^{13}\text{C}]$ -HSQC spectrum recorded with 500  $\mu\text{M}$  10–23 DNAzyme in complex with its stabilized RNA target with 128 scans at 37 °C with natural  ${}^{13}\text{C}$  abundance. The spectral width is covering (deoxy)ribose methylene and thymine methyl signals (indicated by blue labels). Usually, C1'-H1' cross peaks show the best dispersion

## References

1. Breaker RR, Joyce GF (1994) A DNA enzyme that cleaves RNA. *Chem Biol* 1:223–229. [https://doi.org/10.1016/1074-5521\(94\)90014-0](https://doi.org/10.1016/1074-5521(94)90014-0) [Crossref][PubMed]
2. Santoro SW, Joyce GF (1998) Mechanism and utility of an RNA-cleaving DNA enzyme. *Biochemistry* 37:13330–13342. <https://doi.org/10.1021/bi9812221> [Crossref][PubMed]
3. Chandra M, Sachdeva A, Silverman SK (2009) DNA-catalyzed sequence-specific

- hydrolysis of DNA. *Nat Chem Biol* 5:718–720. <https://doi.org/10.1038/nchembio.201> [Crossref][PubMed][PubMedCentral]
4. Santoro SW, Joyce GF (1997) A general purpose RNA-cleaving DNA enzyme. *Proc Natl Acad Sci U S A* 94:4262–4266. <https://doi.org/10.1073/pnas.94.9.4262> [Crossref][PubMed][PubMedCentral]
  5. Cuenoud B, Szostak JW (1995) A DNA metalloenzyme with DNA ligase activity. *Nature* 375:611–614. <https://doi.org/10.1038/375611a0> [Crossref][PubMed]
  6. Flynn-Charlebois A, Wang Y, Prior TK et al (2003) Deoxyribozymes with 2'–5' RNA ligase activity. *J Am Chem Soc* 125:2444–2454. <https://doi.org/10.1021/ja028774y> [Crossref][PubMed]
  7. Chandrasekar J, Silverman SK (2013) Catalytic DNA with phosphatase activity. *Proc Natl Acad Sci U S A* 110:5315–5320. <https://doi.org/10.1073/pnas.1221946110> [Crossref][PubMed][PubMedCentral]
  8. Chu C-C, Wong OY, Silverman SK (2014) A generalizable DNA-catalyzed approach to peptide-nucleic acid conjugation. *Chembiochem* 15:1905–1910. <https://doi.org/10.1002/cbic.201402255> [Crossref][PubMed][PubMedCentral]
  9. Pradeepkumar PI, Höbartner C, Baum DA, Silverman SK (2008) DNA-catalyzed formation of Nucleopeptide linkages. *Angew Chem Int Ed* 47:1753–1757. <https://doi.org/10.1002/anie.200703676> [Crossref]
  10. Silverman SK (2015) Pursuing DNA catalysts for protein modification. *Acc Chem Res* 48:1369–1379. <https://doi.org/10.1021/acs.accounts.5b00090> [Crossref][PubMed][PubMedCentral]
  11. Wong OY, Pradeepkumar PI, Silverman SK (2011) DNA-catalyzed covalent modification of amino acid side chains in tethered and free peptide substrates. *Biochemistry* 50:4741–4749. <https://doi.org/10.1021/bi200585n> [Crossref][PubMed]
  12. Chinnapen DJ-F, Sen D (2004) A deoxyribozyme that harnesses light to repair thymine dimers in DNA. *Proc Natl Acad Sci U S A* 101:65–69. <https://doi.org/10.1073/pnas.0305943101> [Crossref][PubMed]
  13. Baum DA, Silverman SK (2007) Deoxyribozyme-catalyzed labeling of RNA. *Angew Chem Int Ed* 46:3502–3504. <https://doi.org/10.1002/anie.200700357>

[Crossref]

14. Shen Z, Wu Z, Chang D et al (2016) A catalytic DNA activated by a specific strain of bacterial pathogen. *Angew Chem Int Ed* 55:2431–2434. <https://doi.org/10.1002/anie.201510125>  
[Crossref]
15. Torabi S-F, Wu P, McGhee CE et al (2015) In vitro selection of a sodium-specific DNAzyme and its application in intracellular sensing. *Proc Natl Acad Sci U S A* 112:5903–5908. <https://doi.org/10.1073/pnas.1420361112>  
[Crossref][PubMed][PubMedCentral]
16. Wu P, Hwang K, Lan T, Lu Y (2013) A DNAzyme-gold nanoparticle probe for uranyl ion in living cells. *J Am Chem Soc* 135:5254–5257. <https://doi.org/10.1021/ja400150v>  
[Crossref][PubMed][PubMedCentral]
17. Xiao Y, Pavlov V, Niazov T et al (2004) Catalytic beacons for the detection of DNA and telomerase activity. *J Am Chem Soc* 126:7430–7431. <https://doi.org/10.1021/ja031875r>  
[Crossref][PubMed]
18. Basu S, Sriram B, Goila R, Banerjea AC (2000) Targeted cleavage of HIV-1 coreceptor-CXCR-4 by RNA-cleaving DNA-enzyme: inhibition of coreceptor function. *Antivir Res* 46:125–134. [https://doi.org/10.1016/S0166-3542\(00\)00075-9](https://doi.org/10.1016/S0166-3542(00)00075-9)  
[Crossref][PubMed]
19. Cho E-A, Moloney FJ, Cai H et al (2013) Safety and tolerability of an intratumorally injected DNAzyme, Dz13, in patients with nodular basal-cell carcinoma: a phase 1 first-in-human trial (DISCOVER). *Lancet* 381:1835–1843. [https://doi.org/10.1016/S0140-6736\(12\)62166-7](https://doi.org/10.1016/S0140-6736(12)62166-7)  
[Crossref][PubMed][PubMedCentral]
20. Goila R, Banerjea AC (1998) Sequence specific cleavage of the HIV-1 coreceptor CCR5 gene by a hammer-head ribozyme and a DNA-enzyme: inhibition of the coreceptor function by DNA-enzyme. *FEBS Lett* 436:233–238. [https://doi.org/10.1016/S0014-5793\(98\)01137-5](https://doi.org/10.1016/S0014-5793(98)01137-5)  
[Crossref][PubMed]
21. Jakobsen MR, Haasnoot J, Wengel J et al (2007) Efficient inhibition of HIV-1 expression by LNA modified antisense oligonucleotides and DNAzymes targeted to functionally selected binding sites. *Retrovirology* 4:29. <https://doi.org/10.1186/1742-4690-4-29>  
[Crossref][PubMed][PubMedCentral]
- 22.

- Krug N, Hohlfeld JM, Kirsten A-M et al (2015) Allergen-induced asthmatic responses modified by a GATA3-specific DNAzyme. *N Engl J Med* 372:1987–1995. <https://doi.org/10.1056/NEJMoa1411776>  
[Crossref][PubMed]
23. Zhang X, Xu Y, Ling H, Hattori T (1999) Inhibition of infection of incoming HIV-1 virus by RNA-cleaving DNA enzyme. *FEBS Lett* 458:151–156. [https://doi.org/10.1016/S0014-5793\(99\)01149-7](https://doi.org/10.1016/S0014-5793(99)01149-7)  
[Crossref][PubMed]
  24. Aranda J, Terrazas M, Gómez H et al (2019) An artificial DNAzyme RNA ligase shows a reaction mechanism resembling that of cellular polymerases. *Nat Catal* 2:544–552. <https://doi.org/10.1038/s41929-019-0290-y>  
[Crossref]
  25. Liu H, Yu X, Chen Y et al (2017) Crystal structure of an RNA-cleaving DNAzyme. *Nat Commun* 8:2006. <https://doi.org/10.1038/s41467-017-02203-x>  
[Crossref][PubMed][PubMedCentral]
  26. Nowakowski J, Shim PJ, Prasad GS et al (1999) Crystal structure of an 82-nucleotide RNA-DNA complex formed by the 10-23 DNA enzyme. *Nat Struct Biol* 6:151–156. <https://doi.org/10.1038/5839>  
[Crossref][PubMed]
  27. Rosenbach H, Borggräfe J, Victor J et al (2020) Influence of monovalent metal ions on metal binding and catalytic activity of the 10–23 DNAzyme. *Biol Chem* 402:99–111. <https://doi.org/10.1515/hsz-2020-0207>  
[Crossref][PubMed]
  28. Davis JH, Foster TR, Tonelli M, Butcher SE (2006) Role of metal ions in the tetraloop-receptor complex as analyzed by NMR. *RNA* 13:76–86. <https://doi.org/10.1261/rna.268307>  
[Crossref][PubMed]
  29. Borggräfe J, Victor J, Rosenbach H, Viegas A, Gertzen C, Wuebben C, Kovacs H, Gopalswamy M, Riesner D, Steger G, Schiemann O, Gohlke H, Span I, Etzkorn M: Time-resolved structural analysis of an RNA-cleaving DNA catalyst. *Nature* (accepted)
  30. Wijmenga SS, van Buuren BNM (1998) The use of NMR methods for conformational studies of nucleic acids. *Prog Nucl Magn Reson Spectrosc* 32:287–387. [https://doi.org/10.1016/S0079-6565\(97\)00023-X](https://doi.org/10.1016/S0079-6565(97)00023-X)  
[Crossref]
  - 31.

- Rosenbach H, Victor J, Etzkorn M et al (2020) Molecular features and metal ions that influence 10-23 DNAzyme activity. *Molecules* 25:3100. <https://doi.org/10.3390/molecules25133100>  
[Crossref][PubMedCentral]
32. Steger G (1994) Thermal denaturation of double-stranded nucleic acids: prediction of temperatures critical for gradient gel electrophoresis and polymerase chain reaction. *Nucleic Acids Res* 22:2760–2768. <https://doi.org/10.1093/nar/22.14.2760>  
[Crossref][PubMed][PubMedCentral]
  33. Sugimoto N, Okumoto Y, Ohmichi T (1999) Effect of metal ions and sequence of deoxyribozymes on their RNA cleavage activity. *J Chem Soc Perkin Trans 2*:1381–1386. <https://doi.org/10.1039/a901461f>  
[Crossref]
  34. Victor J, Steger G, Riesner D (2018) Inability of DNAzymes to cleave RNA in vivo is due to limited Mg<sup>2+</sup> concentration in cells. *Eur Biophys J* 47:333–343. <https://doi.org/10.1007/s00249-017-1270-2>  
[Crossref][PubMed]
  35. Dingley AJ, Steger G, Esters B et al (2003) Structural characterization of the 69 nucleotide potato spindle tuber viroid left-terminal domain by NMR and thermodynamic analysis. *J Mol Biol* 334:751–767. <https://doi.org/10.1016/j.jmb.2003.10.015>  
[Crossref][PubMed]
  36. Marušič M, Schlagnitweit J, Petzold K (2019) RNA dynamics by NMR spectroscopy. *Chembiochem* 20:2685–2710. <https://doi.org/10.1002/cbic.201900072>  
[Crossref][PubMed][PubMedCentral]
  37. Cetiner EC, Jonker HRA, Helmling C et al (2019) Paramagnetic-iterative relaxation matrix approach: extracting PRE-restraints from NOESY spectra for 3D structure elucidation of biomolecules. *J Biomol NMR* 73:699–712. <https://doi.org/10.1007/s10858-019-00282-0>  
[Crossref][PubMed]
  38. Griesinger C, Ernst R (1987) Frequency offset effects and their elimination in NMR rotating-frame cross-relaxation spectroscopy. *J Magn Reson* 75:261–271. [https://doi.org/10.1016/0022-2364\(87\)90035-7](https://doi.org/10.1016/0022-2364(87)90035-7)  
[Crossref]
  39. Strotz D, Orts J, Chi CN et al (2017) eNORA2 exact NOE analysis program. *J Chem Theory Comput* 13:4336–4346. <https://doi.org/10.1021/acs.jctc.7b00436>  
[Crossref][PubMed]

40. Vögeli B, Olsson S, Güntert P, Riek R (2016) The exact NOE as an alternative in ensemble structure determination. *Biophys J* 110:113–126. <https://doi.org/10.1016/j.bpj.2015.11.031>  
[Crossref][PubMed][PubMedCentral]
  41. Borggrøfe J, Victor J, Rosenbach H, Viegas A, Gertzen CGW, Wuebben C, Kovacs H, Gopalswamy M, Riesner D, Steger G, Schiemann O, Gohlke H, Span I, Etzkorn M (2022) Time-resolved structural analysis of an RNA-cleaving DNA catalyst. *Nature* 601 (7891):144–149. <https://doi.org/10.1038/s41586-021-04225-4>
- 

## Footnotes

<sup>1</sup> During typesetting of the book, Borggrøfe et al. [30] elucidated structure and detailed mechanism of the 10-23 DNAzyme.

## 11. Molecular Modeling and Simulations of DNA and RNA: DNAzyme as a Model System

Christoph G. W. Gertzen<sup>1, 2</sup> and Holger Gohlke<sup>1, 3✉</sup>

- (1) Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
- (2) Center for Structural Studies (CSS), Heinrich Heine University Düsseldorf, Düsseldorf, Germany
- (3) John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), Institute of Biological Information Processing (IBI-7: Structural Biochemistry) and Institute of Bio- and Geosciences (IBG-4: Bioinformatics), Forschungszentrum Jülich GmbH, Jülich, Germany

✉ **Holger Gohlke**  
Email: [gohlke@uni-duesseldorf.de](mailto:gohlke@uni-duesseldorf.de)

### Abstract

Nowadays, the structural dynamics of DNA and RNA is accessible on an atomistic level on a micro- to millisecond time scale via molecular dynamics simulations. However, as DNA or RNA are highly charged molecules, performing such simulations is challenging as to the representation of intramolecular electrostatic interactions and those to solvent molecules and ions. This is particularly true for DNAzymes, where DNA and RNA backbones can come as close as 2.4 Å with their charged phosphate groups during the catalytic cycle. Here, we present tools to simulate the structural dynamics of a DNAzyme, with a focus on detailed

instructions for the Amber suite of programs. Furthermore, we will show how to analyze metal ion binding within the DNAzyme.

**Key words** Structural dynamics – Amber suite of programs – Nucleic acid – Metal ion – Catalysis

---

## 1 Introduction

DNAzymes are catalytically active, non-coding DNAs that, among others, can cleave DNA or RNA targets [1]. The ability to cleave RNA has enormous therapeutic potential, which ranges from targeting RNA-based viruses, like SARS-CoV-2, to applications for personalized therapy.

However, DNAzymes show low catalytic activity in cells, which currently limits their *in vivo* applications. A detailed understanding of their structural dynamics in different environments and its relation to their function is expected to overcome these limitations. Molecular dynamics (MD) simulations allow analyzing DNAzymes in different environments. This knowledge will help design DNAzymes that specifically cleave their targets in the desired environment. To do so, several factors must be considered.

DNAzymes from the 10–23 family form a complex with their target RNA and catalytically cleave the RNA. 10–23 DNAzymes do so in a divalent cation-dependent manner, others are divalent cation-independent [2, 3]. This difference influences the setup of respective MD simulations. Here, we will focus on a setup with monovalent cations. For setting up nucleic acid simulations with divalent cations, see Ref. 4 and recent reviews [5, 6].

MD simulations of DNAzymes can be carried out using different software suites. We will describe the setup using the Amber suite of programs [7, 8] with the OL15 parameters for DNA and the OL3 parameters for RNA for computations using the highly efficient GPU implementation of the *pmemd* program [9]. The OL15 parameters consist of the Amber force field ff99 [10] with the parmbsc0  $\alpha/\gamma$  [11],  $\chi/\zeta$ OL1 [12], xOL4 [13], and  $\beta$ OL1 [14] modifications; the OL3 parameters consist of ff99 [10] with the “OL3”  $\chi$  distribution from ff14SB [15, 16] and the parmbsc0  $\alpha/\gamma$  [11] modifications. The modifications of the force fields lead to backbone torsion angle distributions closer to the experiment. These two parameter sets are the currently recommended parameterizations for

performing DNA and RNA MD simulations with Amber force fields. Other simulation software include CHARMM [17], Desmond [18], GROMACS [19, 20], NAMD [21], and TINKER [22], used with either their own or the Amber forcefields.

As a starting structure, an RNA-bound DNAzyme will be used. As no RNA-bound DNAzyme structure is currently available, we will demonstrate how to create such a complex from a DNA-bound X-ray crystal structure of the 8–17 DNAzyme (PDB ID: 5XM9)<sup>1</sup>. Hence, the described techniques can be used for RNA- and DNA-bound structures as DNA-bound DNAzymes are somewhat easier to simulate because the same forcefield can be used for the DNAzyme and the substrate, and the residue naming is consistent. Structures from other sources, including NMR or structure prediction, can also be used, although the accuracy of the starting structure may impact the simulation results.

Depending on the structure’s origin, great care must be taken for placing monovalent cations around it when setting up the system. If monovalent ions are available from the experimental structure, they can be included in the simulation system. Usually, ions detected by X-ray crystallography or NMR spectroscopy are structurally important. By contrast, in the case of structures determined in solution, close contacts between the backbone phosphates are possible because cations could have diffused into the structure and locally stabilized the charges. However, the ion localization may not have been resolved. When starting the simulations without proper cation placement, the charges of the phosphate backbone can repel each other, which distorts the structure. Therefore, the initial relaxation of the DNA/cation simulation system requires more time than for protein systems to allow the cations to diffuse around and into the structure to achieve local charge compensation. Using a “standard” relaxation protocol [24], it took up to 50 ns of unrestrained MD simulations for the cations to locate around a distorted structure. As the cations may have easier access to the backbone in the distorted structure, we recommend using 300 ns of restrained MD simulations to ensure the integrity of the DNAzyme structure while the cations localize around it. With this amount of restrained relaxation, we did not detect structural distortions among 160 test runs. Nonetheless, the integrity of the DNAzyme structure at the end of the relaxation phase should be checked.

Furthermore, we will describe how to analyze and visualize the structural dynamics of the simulated DNAzyme. For this, we will show how to perform fundamental structural analysis to determine the root mean square deviations and fluctuations (RSMD/RMSF). We will also show how to evaluate metal ion binding to the DNAzyme.

In the following, we will introduce the software and hardware needed to perform MD simulations of a DNAzyme (Subheading 2), explain the steps necessary to simulate a DNAzyme with an RNA substrate (Subheading 3), and note further important points (Subheading 4).

---

## 2 Materials

As mentioned earlier, several MD simulation software can be used to compute the structural dynamics of a DNAzyme. The steps, commands, and scripts needed to perform the MD simulations in these software can differ from the workflow described here for the Amber suite of programs. Nonetheless, the general principles described here could be adapted to the other MD simulation software. An overview of tools and resources is summarized in Table 1.

**Table 1** Tools and resources for MD simulations (of DNAzymes) ordered by the appearance in the text

Program/service	Purpose	Link
CUDA9.2	GPU calculation	<a href="https://www.nvidia.com">https://www.nvidia.com</a>
SLURM workload manager	Cluster management	<a href="https://slurm.schedmd.com/overview.html">https://slurm.schedmd.com/overview.html</a>
Amber GPU information	Hardware setup	<a href="https://ambermd.org/GPUHardware.php">https://ambermd.org/GPUHardware.php</a>
Xmgrace	Data visualization	<a href="https://plasma-gate.weizmann.ac.il/Grace/">https://plasma-gate.weizmann.ac.il/Grace/</a>
R	Data visualization	<a href="https://www.r-project.org/">https://www.r-project.org/</a>
Microsoft excel	Data visualization	<a href="https://www.microsoft.com/de-de/microsoft-365/excel">https://www.microsoft.com/de-de/microsoft-365/excel</a>

Program/service	Purpose	Link
VMD	MD visualization	<a href="http://www.ks.uiuc.edu/Research/vmd/">http://www.ks.uiuc.edu/Research/vmd/</a>
PyMOL	MD visualization	<a href="http://www.pymol.org">http://www.pymol.org</a>
Protein data bank (PDB)	Structure database	<a href="http://www.rcsb.org">http://www.rcsb.org</a>
Nucleic acid database	Structure database	<a href="http://ndbserver.rutgers.edu">http://ndbserver.rutgers.edu</a>
PDB format documentation	Documentation	<a href="http://www.wwpdb.org/documentation/format33/v3.3.html">http://www.wwpdb.org/documentation/format33/v3.3.html</a>
Amber manual	Documentation	<a href="http://ambermd.org/doc12/Amber20.pdf">http://ambermd.org/doc12/Amber20.pdf</a>
Amber tutorials	Tutorial	<a href="http://ambermd.org/tutorials/">http://ambermd.org/tutorials/</a>
SI scripts	Scripts	<a href="https://doi.org/10.25838/d5p-18">https://doi.org/10.25838/d5p-18</a>
Schrödinger maestro	Structure modeling	<a href="https://www.schrodinger.com/products/maestro">https://www.schrodinger.com/products/maestro</a>

For performing the simulations as described, versions 19 or 20 of Amber (including the respective AmberTools) [7] are required. AmberTools is free of charge, and most of its components are released under the GNU General Public License (GPL); Amber facilitates faster simulations on parallel CPU or GPU hardware and is distributed with a separate license structure. Although we recommend a Linux-based high-performance computing cluster with at least NVIDIA RTX 2080Ti GPUs, CUDA 9.2 (Table 1), and a SLURM workload manager [25] (Table 1) to perform the MD simulations efficiently, it is also possible to use a personal computer with high-end hardware. Further details can be found in the Amber GPU information (Table 1), where also an exemplary hardware setup for a high-end personal computer capable of performing GPU-based MD simulations is provided.

Furthermore, one needs a text editor to change input and PDB files, a program to visualize data such as *Xmgrace* (Table 1), *R* (Table 1), *Gnuplot* (Table 1), or *Microsoft Excel* (Table 1), and a program to visualize MD trajectories such as *VMD* [26] (Table 1) or *PyMOL* (Table 1).

### 3 Methods

In the following, we will describe how to prepare a DNAzyme structure for MD simulations and the steps necessary to run the simulations and analyze them. Figure 1 shows an overview of the workflow. As no RNA-bound DNAzyme structure is available, we will show how to convert the DNA “target” of a DNAzyme into RNA in **step 3**. Here, the DNA target is the analog of the RNA target without cleavage reaction and was used for crystallization purposes.

1. Retrieve the structure. The DNAzyme can be an experimentally determined structure, some of which are available at structure databases such as the Protein Data Bank (PDB; Table 1) or the Nucleic Acid Database (NDB; Table 1) or a predicted model. For the following, the structure should be in the PDB format (Table 1). Here, the structure with the PDB ID 5XM9 downloaded from the PDB will be used as an example.
2. Open the PDB file with a text editor and remove all unwanted parts. Depending on what one wants to simulate, the unwanted parts can include proteins, ions, water molecules, or co-solvents. In our example, we delete everything from the first line of the file, starting with the HEADER entry, until the first line starting with the ATOM entry. The file will now start with the line:

```
ATOM 1 N MET A 1 8.630 4.224 16.278 1.00 79.75 N
```

The “A” between MET and 1 indicates that this nitrogen atom belongs to chain A. For our example, delete the proteins and the second copy of the DNAzyme present in the PDB file. Hence, delete every line with atoms belonging to chains A, B, C, D, G, and H, and the second to last line starting with MASTER, and save the file under a new name, e.g., 5xm9\_cut.pdb. This file now contains only one DNAzyme (chain E) and its target (chain F).

3. As no RNA-bound structure of a DNAzyme is currently available, change the DNA target of 5XM9 into RNA. This also demonstrates how small changes to a structure such as a single mutation can be made during the MD simulations setup. For this, open

`5xm9_cut.pdb` with a text editor and replace the D in the residue name of chain F with a space. For example, change.

```
ATOM 3600 05' DA F 1 37.241 39.314 -9.026 1.00 80.16 0
```

to

```
ATOM 3600 05' A F 1 37.241 39.314 -9.026 1.00 80.16 0
```

Note the fixed-column width format and pay attention to conserve it during the changes.

When a deoxythymidine phosphate is present (“DT”), replace the name of the residue by “U” to convert it to a uridine phosphate and delete the line containing the atom C7. For example, delete line:

```
ATOM 3996 C7 DT F 19 -7.040 12.247 -4.395 1.00 82.57 C
```

This is necessary as uracil does not possess this methyl group; its presence would raise an error during parameterization of the DNAzyme/RNA complex. The missing 2'-OH groups at the ribose, on the other hand, will automatically be added during system setup.

Save the edited file under a new name, e.g., `5xm9_RNA.pdb`.

4. Now use *packmol-memgen* [27], available in AmberTools, to set up the system according to the environment of your interest. For our example, add 150 mM of NaCl to mimic an experimental setup. (Experimental conditions (see **Note 1**) or other ions, particularly Mg<sup>2+</sup> ions, (see **Notes 2** and **3**) can also be accounted for.) For this, we will use the following command:

```
packmol-memgen --solute --cubic --pdb 5xm9_RNA.pdb \
--output 5xm9_RNA_packmol.pdb --salt --salt_c Na+ --dist
15
```

The parameters used in this command have the following meaning:

solute	Do not add lipids while solvating the system
--------	--

cubic	Use a cubic box shape
pdb 5xm9_RNA.pdb	Use 5xm9_RNA.pdb as the structure to be set up
output 5xm9_RNA_packmol.pdb	The output file name is 5xm9_RNA_packmol.pdb
salt	In addition to neutralizing the system, add salt (default 150 mM)
salt_c Na <sup>+</sup>	Use NaCl as a salt
dist 15	The minimum distance between the solute and the water boundary will be 15 Å

A comprehensive list of possible options can be displayed with:

```
packmol-memgen --help
```

5. Use *pdb4amber* to analyze and clean up the PDB file for further usage:

```
pdb4amber -i 5xm9_RNA_packmol.pdb -o 5xm9_RNA_p4a.pdb
```

-i 5xm9_RNA_packmol.pdb	Specifies the input file
-o 5xm9_RNA_p4a.pdb	Specifies the output file

The resulting system should look similar to Fig. 2, when opened in a molecular viewer, e.g., via `pymol 5xm9_RNA_p4a.pdb`.

6. *Packmol-memgen* added the H2" hydrogen atom to the RNA, which should not be there. Open the PDB file `5xm9_RNA_p4a.pdb` with a text editor and delete the lines of the RNA containing the H2" atom and name the resulting file `5xm9_RNA_p4a_cut.pdb`.
7. Specify which force field parameters to use for the DNA, RNA, ions, and water molecules. For this, we will use *t leap* and a generated input file named `t leap_DNAzyme.in` with the content shown in Fig. 3.
- Execute *t leap* with:

```
t leap -f t leap_DNAzyme.in
```

8. Create the folders `cryst`, `equi`, and `prod` and copy the files `5xm9_RNA_solv.prmtop` and `5xm9_RNA_solv.inpcrd`, which were created with `tleap` in the previous step, into the `cryst` folder. Then copy the files provided in the SI scripts (Table 1) for `equi` and `prod` into their respective folders in your directory.

You will need to edit the files according to your computing system and workload manager configuration. Note that the files provided are intended for a bash environment with SLURM [25] running Amber19.14 with CUDA 9.2.

Explaining every line of every file is beyond the scope of this article. As to Amber, we refer to the detailed manual (Table 1) and helpful tutorials (Table 1) for further details. Here, we will explain the purpose of the files and show which parameters to edit.
9. The MD simulations start with minimizing and thermalizing the DNAzyme/RNA complex structure. This removes structural clashes within and between molecules and ions and adapts the simulation system to the ensemble conditions, i.e., temperature and/or pressure. This is handled by the scripts `cpu-par.srv`, `gpu_ser1.srv`, `gpu_ser2.srv`, `gpu_ser3.srv`, and `gpu_ser4.srv`.
  - (a) `cpu-par.srv` reads in the input files `min_ntr_h.in`, `min_ntr_l.in`, and `min_ntr_n.in` sequentially and is executed on 20 CPUs (#SBATCH --ntasks=20) in the queues named “rest-of-gpu1, rest-of-gpu2, rest-of-gpu3, and rest-of-gpu4”. If your queues for parallel CPU computing are named differently, change the line “#SBATCH --partition=rest-of-gpu1,rest-of-gpu2,rest-of-gpu3,rest-of-gpu4” according to your setup. The input files correspond to short energy minimizations using positional restraints on the DNAzyme/RNA complex with a force constant of 5 kcal mol<sup>-1</sup> Å<sup>-2</sup> (`min_ntr_h.in`), 1 kcal mol<sup>-1</sup> Å<sup>-2</sup> (`min_ntr_l.in`), and no restraints (`min_ntr_n.in`). The restraints are defined by the block:  
**Hold the nucleic acids fixed**

```
RES 1 59  
END
```

with “5.0” defining the force constant and “RES 1 59” defining the residues to be restrained. If you want to simulate a different DNAzyme/target complex, edit the line “RES 1 59” according to your system in every input file in which restraints are used (`min_ntr_h.in`, `min_ntr_l.in`, `md_nvt_ntr.in`, `md_npt_ntr_01.in`, `md_npt_ntr_02.in`, `md_npt_ntr_03.in`, `md_npt_ntr_04.in`, `md_npt_ntr_05.in`, `md_nvt_red_01.in`, `md_nvt_red_02.in`, `md_nvt_red_03.in`, `md_nvt_red_04.in`, and `md_nvt_red_05.in`). If the names of your topology (.prmtop) and input coordinate (.inpcrd) files are different, adapt the lines “PRMTOP=../cryst/5xm9\_RNA\_solv.prmtop” and “INPCRD=../cryst/5xm9\_RNA\_solv.inpcrd”, accordingly in all .srv files.

In this example, we use CPUs for the minimization as sometimes, with close contacts between atoms due to the packing, energy minimizations on GPUs can lead to errors.

At its end, `cpu-par.srv` automatically executes `gpu_ser1.srv`.

- (b) `gpu_ser1.srv` reads in the input files `md_nvt_ntr.in`, `md_npt_ntr_0[1-5].in`, and `md_nvt_red01.in` sequentially and is executed on 1 GPU (#SBATCH --ntasks=1; #SBATCH --gres=gpu:1) in the queues named “gpu-ser”. If your queues for serial GPU computing are named differently, change the line “#SBATCH --partition=gpu-ser” according to your setup. The input files correspond to short MD simulations using positional restraints on the DNAzyme/RNA complex with a force constant of 5 kcal mol<sup>-1</sup> Å<sup>-2</sup> in the NVT ensemble (constant number of particles, volume, and temperature; `md_nvt_ntr.in`, `md_nvt_red_01.in`) or NPT ensemble (constant number of particles, pressure, and temperature: `md_npt_ntr_0[1-5].in`).

Notably, `md_nvt_ntr.in` heats the system from 0 K to 100 K ( $\text{temp}0 = 100.0$ ,  $\text{temp}i = 0.0$ ) and `md_npt_ntr_01.in` heats the system from 100 K to 300.1 K ( $\text{temp}0 = 300.1$ ,

$\text{temp}i = 100.0$ ). If you simulate several replicas of the DNAzyme/RNA complex, we encourage you to alter the end temperature of 300.1 K in `md_npt_ntr_01.in` by a fraction of a Kelvin for each replica to ensure that the simulations behave differently (see **Note 4**). The input files `md_npt_ntr_0[2-5].in` are similar and ensure that the system can adapt to the environment; for each run, the previously adapted system is used as input for further adaptation.

The input file `md_nvt_red_01.in` is the first of five scripts (`md_nvt_red_0[1-5].in`) that correspond to performing 50 ns of MD simulations with decreasing restraining forces such that the ions can diffuse around and into the structure but the structure remains close to the starting conformation. The scripts `gpu_ser[1-4].srv` execute each other in order and break up the five steps so that, depending on your workload manager setup, wall time errors are avoided.

Here, `gpu_ser1.srv` executes the thermalization scripts and `md_nvt_red_01.in` as mentioned above, `gpu_ser2.srv` executes `md_nvt_red_0[2-3].in` and `gpu_ser3.srv` executes `md_nvt_red_0[4-5].in`. The script `md_nvt_red_06.in`, which is executed by `gpu_ser4.srv`, applies no restraints.

`gpu_ser4.srv` copies the final restart file (coordinate file) `md_nvt_red_06.restrt` to `../prod`, switches to the production folder `prod`, and submits the `gpu_ser.srv` script.

10. The main MD simulation is handled by the script `gpu_ser.srv`. It sequentially performs 40 subsequent NVT MD simulations of 25 ns length each, resulting in a 1  $\mu\text{s}$  long simulation. The MD simulation is executed on 1 GPU (#SBATCH --ntasks=1; #SBATCH --gres=gpu:1) in the queue named “gpu-ser”. If the simulation is stopped, e.g., due to a server crash, submit `gpu_ser.srv` again, as the script makes a copy of itself in every step and then updates all necessary parameters

after a successful execution. Resubmit the script using the command:

```
sbatch gpu_ser.srv
```

11.

After the MD simulation is finished, analyze the trajectory in two steps. First, execute a script termed `cpptraj1.in` with the content shown in Fig. 4. This will result in an RMSF plot for determining which residues to use to align the generated conformations to in the second step. This leads to a structural alignment in which the RMSF of residues with low mobility is not inflated due to the influence of high-mobility residues. For example, a loop moving from one side of the DNAzyme to the other can lead to pronounced changes in the structural alignment and increase the RMSF of the rest of the DNAzyme, although the rest barely moves.

Execute the script with command:

```
cpptraj -i cpptraj1.in
```

12.

Open the file containing the RMSF in a program capable of plotting data.

```
xmgrace 5xm9_RNA_RMSF_raw.dat
```

Then, identify regions of the DNAzyme/RNA complex with a low RMSF compared to the rest of the structure, as shown in Fig. 5, to be used as a reference for further analyses.

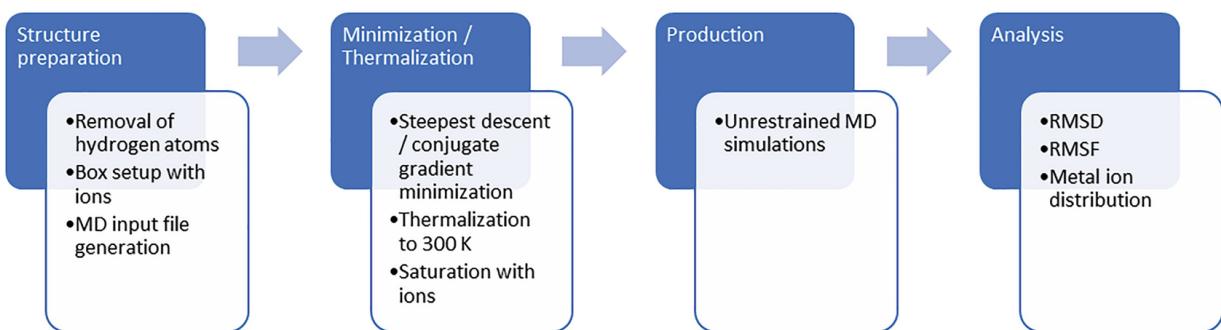
13. For the second step of the MD trajectory analysis, align the conformations via residues with low mobility as determined in the previous step to residues 5–10 and 25–41. Then generate an RMSD and RMSF plot, an average structure of the trajectory, and a density map of the  $\text{Na}^+$  ions. To this end, create the file `cpptraj2.in` with the contents shown in Fig. 6. In addition, open the file `5xm9_RNA_solv.pdb` with a text editor. The first three numbers in the first line of the PDB file are the box dimensions, which you need for the grid calculation.

Execute the script with the command:

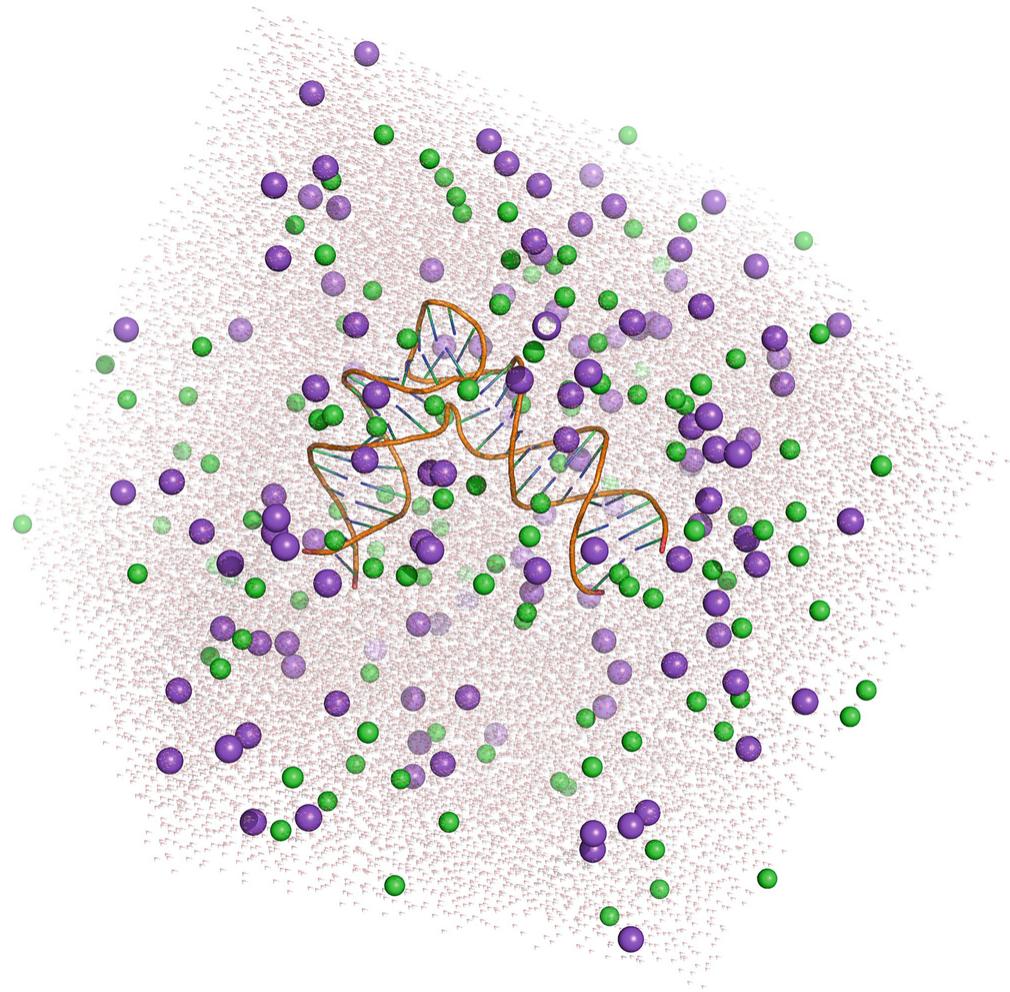
```
cpptraj -i cpptraj2.in
```

You will now be able to view the RMSF as demonstrated before by opening the file `5xm9_RNA_RMSF_final.dat` and the all-atom RMSD by opening the file `5xm9_RNA_RMSD_final.dat`. Note that the RMSD is shown for every frame. As the structures were written every 0.1 ns to the trajectory and we only considered every fifth frame, the time step between displayed data is 0.5 ns. The RMSD shows the structural deviations over time from the starting structure and should not show a steady drift towards larger values. An exemplary RMSD plot of a DNAzyme/RNA complex reaching a plateau was created with `xmgrace` and is shown in Fig. 7. The trajectory file `5xm9_RNA_ions.nc`, containing frames with time separations of 0.5 ns, can be viewed in vmd using the parameter file `5xm9_RNA_ions.prmtop`.

The  $\text{Na}^+$  ion density file `5xm9_RNA_Na_grid.dx` can be opened in *PyMOL* together with the average structure `5xm9_RNA_avg.pdb`. The density displays the occurrence frequency of  $\text{Na}^+$  ions along the MD trajectory and, therefore, in combination with the average structure, allows identifying residues that preferentially interact with the cations. A quantification of per-residue ion binding, however, must be done in several steps (see **Note 5**).



**Fig. 1** Workflow for MD simulations and analysis of a DNAzyme



**Fig. 2** The solvated system of the DNAzyme in complex with the RNA (cartoon) with water (red sticks),  $\text{Na}^+$  ions (violet spheres), and  $\text{Cl}^-$  ions (small green spheres)

```

1  logfile leap.log
2  source leaprc.water.tip3p
3  loadAmberParams frcmod.tip3p
4  loadAmberParams frcmod.ionsjc_tip3p
5  loadAmberParams frcmod.ions234lm_126_tip3p
6  loadOff DNA.OL15.lib
7  source leaprc.DNA.OL15
8  source leaprc.RNA.OL3
9  DNAzyme = loadPdb 5xm9_RNA_p4a_cut.pdb
10 addIonsRand DNAzyme Na+ 0
11 addIonsRand DNAzyme Cl- 0
12 setBox DNAzyme "vdw"
13 saveamberparm DNAzyme 5xm9_RNA_solv.prmtop 5xm9_RNA_solv.inpcrd
14 savepdb DNAzyme 5xm9_RNA_solv.pdb
15 quit

```

**Fig. 3** Input file `t leap_DNAzyme.in` (grey area) for *t leap*. Line 1: Write the input file `leap.log`. Line 2: Load the parameters for TIP3P water. Line 3–5: Load additional parameters for TIP3P water, Joung–Cheatham parameters for monovalent ions [28], and Li–Merz parameters for two-fold positively charged metal ions [29]. Line 6–8: Load parameters and libraries for DNA and RNA. Line 9: Load the input structure. Line 10, 11: Make sure the system is neutralized with  $\text{Na}^+$ - or  $\text{Cl}^-$  ions, in addition to the 150 mM NaCl added with *packmol-memgen*. Line 12: As water is already present, determine the boundary for the periodic boundary conditions by the van der Waal radii of the system. Line 13, 14: Save the input files for the MD simulations and the PDB file of the final system. Line 15: Close *t leap*

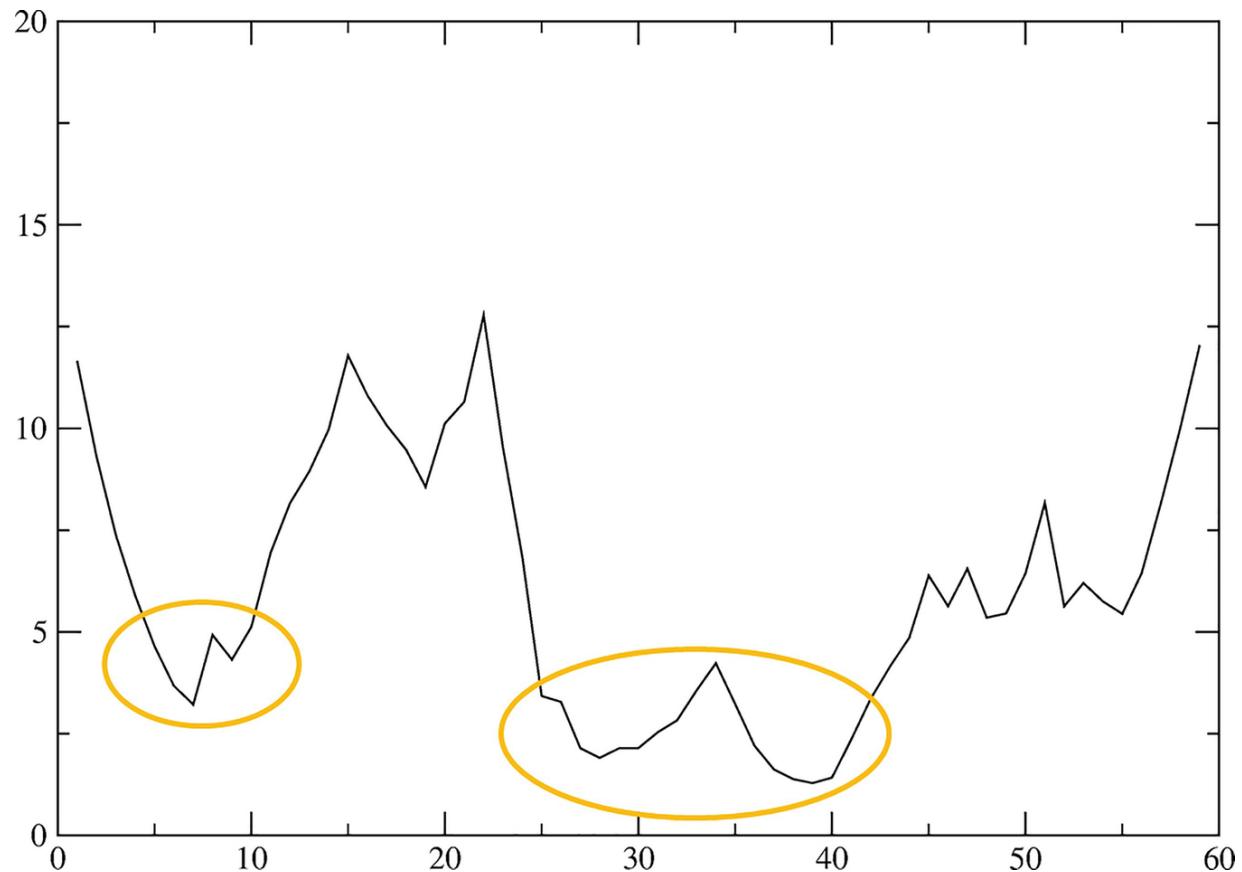
```

1  parm cryst/5xm9_RNA_solv.prmtop
2  trajin prod/*.nc
3  autoimage
4  rmsd :1-59
5  rmsf out 5xm9_RNA_RMSF_raw.dat :1-59 byres
6  go
7  exit

```

**Fig. 4** Input file `cpptraj1.in` used to find regions with low mobility. Line 1: Read in the topology file of the DNAzyme/RNA complex. Line 2: Read in the trajectory files. Line 3: Image the trajectory to the DNAzyme/RNA complex. Line 4: Fit the conformations to the starting structure (residues 1–59) using least-squares fitting. Line 5: Calculate the

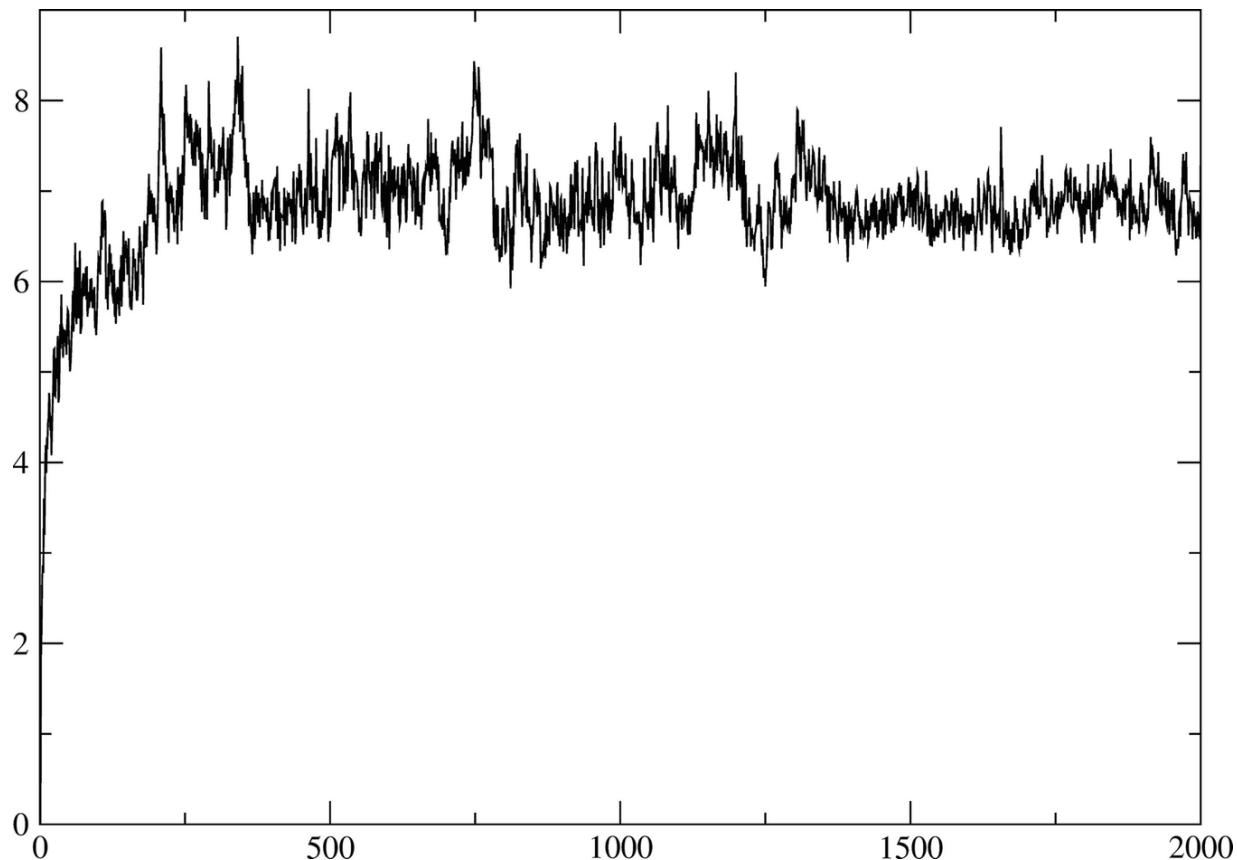
RMSF of the DNAzyme/RNA complex per residue and write it to the file  
5xm9\_RNA\_RMSF\_raw.dat. Line 6, 7: Wrap up



**Fig. 5** Example RMSF plot created with *xmgrace*. The x-axis shows the residue number and the y-axis the RMSF in Å. The orange circles show regions with a lower RMSF than the rest of the structure. These residues will be used to align the conformations for further analysis

```
1  parm cryst/5xm9_RNA_solv.prmtop
2  trajin prod/*.nc 5 last 5
3  strip :WAT
4  autoimage
5  rmsd :5-10,25-41
6  rmsd :1-59 out 5xm9_RNA_RMSD_final.dat nofit
7  rmsf out 5xm9_RNA_RMSF_final.dat :1-59 byres
8  average 5xm9_RNA_avg.pdb pdb :1-59
9  grid 5xm9_RNA_Na_grid.dx 115 0.5 115 0.5 115 0.5 :NA pdb 5xm9_RNA_Na_grid.pdb
10 trajout 5xm9_RNA_ions.nc netcdf
11 parmstrip :WAT
12 parmwrt out 5xm9_RNA_ions.prmtop
13 go
14 exit
```

**Fig. 6** Input file `cpptraj2.in` used for structural and cation binding analysis. Line 1, 2: Load parameters and trajectory. Start with the fifth frame and then only use every fifth frame until the last frame. Line 3: Remove the water from the trajectory to save storage space. Line 4: Image the trajectory to the DNAzyme/RNA complex. Line 5: Align the structure to regions of low mobility (residues 5–10 and 25–41; Fig. 5) identified in the previous step. Line 6: Write out the all-atom RMSD for the DNAzyme/RNA complex over time without changing the alignment from the previous step (`nofit`). Line 7: Write out the RMSF of the DNAzyme/RNA complex aligned on the low-mobility regions. Line 8: Generate the average structure of the DNAzyme/RNA complex over the trajectory. Line 9: Generate a density grid of the  $\text{Na}^+$  ion distribution, using the box dimensions ( $3 \times 115 \text{ \AA}$  and a grid spacing of  $0.5 \text{ \AA}$ ). Line 10: Write a complete trajectory containing only the DNAzyme/RNA complex and ions. Line 11, 12: Remove the water from the parameter file and save it. Line 13, 14: Wrap up



**Fig. 7** Example RMSD plot for a DNAzyme/RNA complex created with *xmgrace*. The x-axis shows the frame number and the y-axis the RMSD in Å. The time separation between two frames is 0.5 ns so that the 2000 frames correspond to 1 µs of simulation time. The DNAzyme/RNA complex structure fluctuates but does not show increasing deviations from the starting structure

---

## 4 Notes

1. If experimental conditions, particularly the DNAzyme concentration, are known, the box size should be adjusted accordingly. This can be achieved by changing the `--dist` flag during system generation with *packmol-memgen*. As described earlier, opening the resulting PDB file with a text editor displays the box size in the first line, which can be used to calculate the volume of the box.
2. For some DNAzymes, the presence of  $Mg^{2+}$  ions is vital for their function. Simulating  $Mg^{2+}$  ions in the presence of DNA or RNA is challenging because of the high charge density of both molecular systems and the slow exchange kinetics once a  $Mg^{2+}$  ion binds to the

nucleic acid [30]. Hence, rather than placing non-hydrated Mg<sup>2+</sup> next to the nucleic acid with *packmol-memgen*, hexahydrated Mg<sup>2+</sup> should be placed [30]. Such hexahydrated Mg<sup>2+</sup> can be generated, e.g., in *maestro* (Table 1) by placing six water molecules around one Mg<sup>2+</sup> and minimizing the complex. If stored under the name Mg\_hexa.pdb, the hexahydrated Mg<sup>2+</sup> ions can be incorporated into the simulation system during **step 4** of the setup with the following command:

```
packmol-memgen --solvate --cubic --pdb 5xm9_RNA_74.pdb \
--output 5xm9_RNA_packmol.pdb --salt --salt_c Na+ --dist
\15 --solute Mg_hexa.pdb --solute_con 0.02M \
--solute_charge +2
```

This will, in addition to 150 mM NaCl, place 20 mM of hexahydrated Mg<sup>2+</sup> ions at random places into the box, which will allow the complexes to better equilibrate around the nucleic acid structure before interactions with the phosphate backbone are formed.

3. Instead of using explicit ions, it is possible to use no ions in the MD simulations but a neutralizing plasma [31]. However, it has been shown that this leads to higher structural variability in MD simulations of DNA duplexes [32]. The presence of high local charge densities due to close contacts of phosphate backbones potentially present in DNAzymes likely will also lead to structural distortions.
4. We showed the setup for one MD simulation of 1 μs length of a DNAzyme/RNA complex. However, we recommend performing at least ten replicas while varying the starting temperatures, as shown in Subheading 3, **step 9b**. Performing independent replicas allows estimating the convergence of analyzed parameters and evaluating their statistical uncertainty.
5. The creation of cation density maps and an average structure of the DNAzyme/RNA complex can be a good means to identify ion binding sites. However, quantifying the thermodynamics or kinetics of Mg<sup>2+</sup> ion or other cation binding to specific residues requires additional

ion or other cation binding to specific residues requires additional analyses. For this, we will measure the minimal distance to any Mg<sup>2+</sup> ion throughout the simulation and store the distance over time in files for each residue of the DNAzyme/RNA complex by adding the following lines after line 9 in the script `cpptraj2.in`:

```
for residues DRNA inmask :1-59 i=1;i++
nativecontacts $DRNA :MG out
5xm9_RNA_cluster1_run1_20mM_MG_res_$i.dat mindist
byresidue skipnative
done
```

The names of the output files `5xm9_RNA_cluster1_run1_20mM_MG_res_$i.dat`, with a number “\$i” for each residue, cover several scenarios. The fragment “cluster1” denotes that several similar structural conformations, e.g., resulting from NMR experiments, are present and were subjected to MD simulations. The fragment “run1” denotes that several (ten) independent replicas were performed. The fragment “20 mM” denotes that MD simulations were performed with 20 mM of Mg<sup>2+</sup> ions.

We included two additional scripts to quickly calculate the results for each cluster and concentration, `write_MG_2col.sh` and `calc_MG_contacts.py` to be found in the folder `analysis` in the SI scripts (Table 1), which will calculate the percent of time Mg<sup>2+</sup> ions are binding with distances <4 Å for each residue. First, execute the `write_MG_2col.sh` script in the folder with the Mg<sup>2+</sup> distances in your shell. Then execute `calc_MG_contacts.py` with:

```
./calc_MG_contacts.py [output file name of your choice]
```

To adapt the scripts to your setup, you need to edit the lines:

```
for k in 1 ; do
for m in 5 10 20 ; do
for i in `seq 1 10` ; do
for j in `seq 1 59` ; do
kname=5xm9_RNA_cluster"$k"_run"$i"_"$m"mM_MG_res_"$j"
```

in the `write_MG_2col.sh` script with “`k in 1`” defining the clusters, “`m in 5 10 20`” defining the cation concentrations in mM, “`i in `seq 1 10``” defining the number of replicas, and “`j in `seq 1 59``” defining the number of residues. The variable “`kname`” reflects the name chosen in the `cpptraj` script.

To adapt the `calc_MG_contacts.py` script, the variables “clusters,” “residues,” “concentrations,” and “runs” at the beginning of the script need to be adapted as well as the variable “`kname`” in the function “`def get_results(clu, con, res):`” in line 31 of the script.

Sequentially executing both scripts will result in an output file, which contains the percentage of time a residue was closer than the minimal distance to a  $Mg^{2+}$  ion during the MD simulation for each cluster, each  $Mg^{2+}$  ion concentration, and each residue. Thus, the  $Mg^{2+}$  ion binding percentages between different experimental conditions can be quantified and compared.

## Acknowledgments

H.G. is grateful for computational support and infrastructure provided by the “Zentrum für Informations- und Medientechnologie” (ZIM) at the Heinrich Heine University Düsseldorf and the John von Neumann Institute for Computing (NIC) (user ID: HKF7), VSK33. The Center for Structural Studies is funded by the DFG (Grant number 417919780).

---

## References

1. Li Y, Sen D (1997) Toward an efficient DNAzyme. *Biochemistry* 36(18):5589–5599  
[Crossref]
2. Geyer CR, Sen D (1997) Evidence for the metal-cofactor independence of an RNA phosphodiester-cleaving DNA enzyme. *Chem Biol* 4(8):579–593  
[Crossref]
3. Torabi S-F, Wu P, McGhee CE, Chen L, Hwang K, Zheng N, Cheng J, Lu Y (2015) In vitro selection of a sodium-specific DNAzyme and its application in intracellular sensing. *Proc Natl Acad Sci U S A* 112(19):5903–5908  
[Crossref]
4. Hanke CA, Gohlke H (2015) Force field dependence of riboswitch dynamics. *Methods*

Enzymol 553:163–191

[Crossref]

5. Sponer J, Bussi G, Krepl M, Banáš P, Bottaro S, Cunha RA, Gil-Ley A, Pinamonti G, Poblete S, Jurečka P (2018) RNA structural dynamics as captured by molecular simulations: a comprehensive overview. *Chem Rev* 118(8):4177–4338  
[Crossref]
6. Salsbury AM, Lemkul JA (2021) Recent developments in empirical atomistic force fields for nucleic acids and applications to studies of folding and dynamics. *Curr Opin Struct Biol* 67:9–17  
[Crossref]
7. D.A. Case KB, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham III TE, Cruzeiro VWD, Darden TA, Duke RE, Giambasu G, Gilson MK, Gohlke H, Goetz AW, Harris R, Izadi S, Izmailov SA, Kasavajhala K, Kovalenko A, Krasny R, Kurtzman T, Lee TS, LeGrand S, Li P, Lin C, Liu J, Luchko T, Luo R, Man V, Merz KM, Miao Y, Mikhailovskii O, Monard G, Nguyen H, Onufriev A, Pan F, Pantano S, Qi R, Roe DR, Roitberg A, Sagui C, Schott-Verdugo S, Shen J, Simmerling C, Skrynnikov NR, Smith J, Swails J, Walker RC, Wang J, Wilson L, Wolf RM, Wu X, Xiong Y, Xue Y, York DM, Kollman PA (2020) AMBER 2020. University of California, San Francisco
8. Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26(16):1668–1688  
[Crossref]
9. Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, Walker RC (2013) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *J Chem Theory Comput* 9(9):3878–3888  
[Crossref]
10. Wang J, Cieplak P, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 21(12):1049–1074  
[Crossref]
11. Pérez A, Marchán I, Svozil D, Sponer J, Cheatham TE III, Laughton CA, Orozco M (2007) Refinement of the AMBER force field for nucleic acids: improving the description of  $\alpha/\gamma$  conformers. *Biophys J* 92(11):3817–3829  
[Crossref]
12. Zgarbová M, Luque FJ, Sponer J, Cheatham TE III, Otyepka M, Jurecka P (2013) Toward improved description of DNA backbone: revisiting epsilon and zeta torsion

force field parameters. *J Chem Theory Comput* 9(5):2339–2354  
[Crossref]

13. Krepl M, Zgarbová M, Stadlbauer P, Otyepka M, Banáš P, Koca J, Cheatham TE III, Jurecka P, Sponer J (2012) Reference simulations of noncanonical nucleic acids with different  $\chi$  variants of the AMBER force field: quadruplex DNA, quadruplex RNA, and Z-DNA. *J Chem Theory Comput* 8(7):2506–2520  
[Crossref]
14. Zgarbová M, Sponer J, Otyepka M, Cheatham TE III, Galindo-Murillo R, Jurecka P (2015) Refinement of the sugar–phosphate backbone torsion beta for AMBER force fields improves the description of Z-and B-DNA. *J Chem Theory Comput* 11(12):5723–5736  
[Crossref]
15. Zgarbová M, Otyepka M, Ji Š, At M, Banáš P, Cheatham TE III, Jurecka P (2011) Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J Chem Theory Comput* 7(9):2886–2902  
[Crossref]
16. Banás P, Hollas D, Zgarbová M, Jurecka P, Orozco M, Cheatham TE III, Sponer J, Otyepka M (2010) Performance of molecular mechanics force fields for RNA simulations: stability of UUCG and GNRA hairpins. *J Chem Theory Comput* 6(12):3836–3849  
[Crossref]
17. Brooks BR, Brooks CL III, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* 30(10):1545–1614  
[Crossref]
18. Bowers KJ, Chow DE, Xu H, Dror RO, Eastwood MP, Gregersen BA, Klepeis JL, Kolossvary I, Moraes MA, Sacerdoti FD (2006) Scalable algorithms for molecular dynamics simulations on commodity clusters. In: SC'06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing. IEEE, New York, pp 43–43  
[Crossref]
19. Bekker H, Berendsen H, Dijkstra E, Achterop S, Vondrumen R, Vanderspoel D, Sijbers A, Keegstra H, Reitsma B, Renardus M (1993) Gromacs: a parallel computer for molecular dynamics simulations. In: de Groot R, Nadrchal J (eds) Physics computing 92. World Scientific, Singapore
20. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E (2015)

GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1:19–25  
[Crossref]

21. Phillips JC, Hardy DJ, Maia JD, Stone JE, Ribeiro JV, Bernardi RC, Buch R, Fiorin G, Hénin J, Jiang W (2020) Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys* 153(4):044130  
[Crossref]
22. Rackers JA, Wang Z, Lu C, Laury ML, Lagardère L, Schnieders MJ, Piquemal J-P, Ren P, Ponder JW (2018) Tinker 8: software tools for molecular design. *J Chem Theory Comput* 14(10):5273–5289  
[Crossref]
23. Borggräfe J, Victor J, Rosenbach H, Viegas A, Gertzen CGW, Wuebben C, Kovacs H, Gopalswamy M, Riesner D, Steger G, Schiemann O, Gohlke H, Span I, Etzkorn M (2022) Time-resolved structural analysis of an RNA-cleaving DNA catalyst. *Nature* 601(7891):144–149. <https://doi.org/10.1038/s41586-021-04225-4>
24. Perez-Garcia P, Kobus S, Gertzen CG, Hoeppner A, Holzscheck N, Strunk CH, Huber H, Jaeger K-E, Gohlke H, Kovacic F (2021) A promiscuous ancestral enzyme’s structure unveils protein variable regions of the highly diverse metallo-β-lactamase family. *Commun Biol* 4(1):1–12  
[Crossref]
25. Yoo AB, Jette MA, Grondona M (2003) Slurm: simple linux utility for resource management. In: Workshop on job scheduling strategies for parallel processing. Springer, New York, pp 44–60  
[Crossref]
26. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14(1):33–38  
[Crossref]
27. Schott-Verdugo S, Gohlke H (2019) PACKMOL-memgen: a simple-to-use, generalized workflow for membrane-protein–lipid-bilayer system building. *J Chem Inf Model* 59(6):2522–2528  
[Crossref]
28. Joung IS, Cheatham TE III (2008) Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J Phys Chem B* 112(30):9020–9041  
[Crossref]
- 29.

- Li P, Roberts BP, Chakravorty DK, Merz KM Jr (2013) Rational design of particle mesh Ewald compatible Lennard-Jones parameters for+ 2 metal cations in explicit solvent. *J Chem Theory Comput* 9(6):2733–2748  
[Crossref]
30. Hanke CA, Gohlke H (2017) Ligand-mediated and tertiary interactions cooperatively stabilize the P1 region in the guanine-sensing riboswitch. *PLoS One* 12(6):e0179271  
[Crossref]
31. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an  $N \cdot \log(N)$  method for Ewald sums in large systems. *J Chem Phys* 98(12):10089–10092  
[Crossref]
32. Galindo-Murillo R, Cheatham TE 3rd (2019) Lessons learned in atomistic simulation of double-stranded DNA: solvation and salt concerns [article v1. 0]. *Living J Comput Mol Sci* 1(2):9974  
[Crossref]

---

## Footnotes

<sup>1</sup> During typesetting of this article, an NMR structure of a 10-23 DNAzyme in complex with RNA has been solved: PDB ID 7PDU; [23].

# **Part IV**

# **Further Biophysical Methods**

## 12. Single-Molecule Kinetic Studies of Nucleic Acids by Förster Resonance Energy Transfer

Mélodie C. A. S. Hadzic<sup>1</sup>, Roland K. O. Sigel<sup>1</sup> and Richard Börner<sup>2</sup> 

(1) Department of Chemistry, University of Zurich, Zurich, Switzerland

(2) Laserinstitut Hochschule Mittweida, University of Applied Sciences  
Mittweida, Mittweida, Germany

 **Richard Börner**  
Email: [richard.boerner@hs-mittweida.de](mailto:richard.boerner@hs-mittweida.de)

### Abstract

Single-molecule microscopy is often used to observe and characterize the conformational dynamics of nucleic acids (NA). Due to the large variety of NA structures and the challenges specific to single-molecule observation techniques, the data recorded in such experiments must be processed via multiple statistical treatments to finally yield a reliable mechanistic view of the NA dynamics. In this chapter, we propose a comprehensive protocol to analyze single-molecule trajectories in the scope of single-molecule Förster resonance energy transfer (FRET) microscopy. The suggested protocol yields the conformational states common to all molecules in the investigated sample, together with the associated conformational transition kinetics. The given model resolves states that are indistinguishable by their observed FRET signals and is estimated with 95% confidence using error calculations on FRET states and transition rate constants. In the end, a step-by-step user guide is given to reproduce the protocol with the

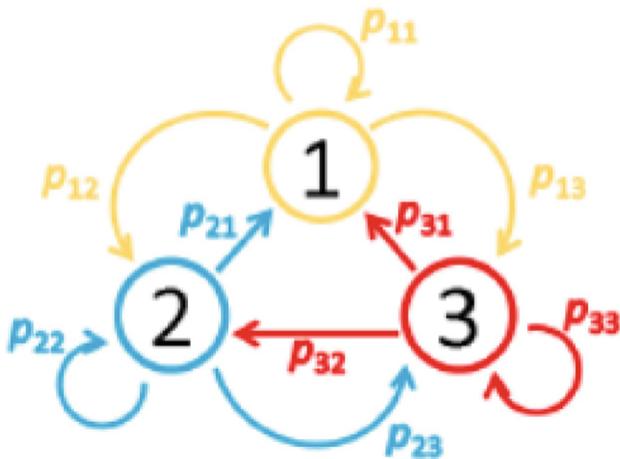
Multifunctional Analysis Software to Handle single-molecule FRET data (MASH-FRET).

**Key words** Single-molecule FRET (smFRET) – Kinetics – Conformational dynamics – Kinetic heterogeneity – Degenerate FRET states – NA folding – Structural dynamics – MASH-FRET

---

## 1 Introduction

The conformational dynamics of nucleic acids (NA) during folding [1], metabolite binding [2], or catalysis [3] are often described as a Markovian process, in which the transitions between conformational states happen stochastically (Fig. 1). Single-molecule Förster resonance energy transfer (smFRET) allows to observe these conformational transitions for individual molecules, and thus, to get a mechanistic view of the biomolecule in action. FRET is a non-radiative transfer of energy between two nearby fluorophores, a donor and an acceptor, with overlapping emission and absorption spectra [4–6]. It is particularly successful on the single-molecule level as it allows to detect short-lived or rare folding intermediates that are usually averaged out in the molecule ensemble [7–9]. In general, smFRET measurements are carried out using confocal or widefield microscopy. In the latter, total internal reflection fluorescence (TIRF) microscope illumination is used to excite surface-immobilized and fluorescently labeled single-(bio)molecules [10, 11]. The responsive fluorescence signal is recorded with a camera to obtain single-molecule videos (SMV). This method allows to observe hundreds of single molecules in parallel and over time, with a time resolution defined by the frame rate of the camera and an observation time limited by the fluorophores photobleaching probability [12]. The identification of distinct FRET states as well as the quantification of state transition rate constants from SMV is central to the data analysis in single- molecule kinetic studies [13] [4–6].



**Fig. 1** Markovian process depicted as a Trellis diagram. States are represented by circled numbers, and  $p_{ij}$  is the probability to transit to state  $j$  being in state  $i$ . Note that  $p_{ij}$  are normalized so that the sum of probabilities to leave state  $i$ , i.e., drawn with the same color, equals to one. Emission probabilities  $b_i$  are not represented here. In FRET-time traces, emission probabilities are approximated by a normal distribution centered on the discrete  $FRET_i$  values

Here, we describe a comprehensive protocol to analyze the kinetics of fluorescently labeled NAs including the discretization of FRET trajectories, i.e., the identification of states and state transitions in single-molecule FRET-time traces, the determination of a global state configuration with transition rate constants, followed by the validation of the final model using simulations. Our approach is adaptable to the experimental settings, i.e., the number of color channels, the scheme of alternating laser excitation etc., and is capable of solving degenerate state systems, i.e., systems that involve multiple states showing the same FRET signal [14, 15], and estimates with 95% confidence the error associated with the FRET states and transition rate constants of the model. The protocol is illustrated with the binding kinetics of a surface-immobilized RNA hairpin, the exon-binding site (EBS) of a group IIB intron, with its freely diffusing cognate, the intron-binding site (IBS) [15].

## 2 Materials

All steps can be performed with the MATLAB-based Multifunctional Analysis Software for Handling single-molecule FRET data (MASH-

FRET) [4, 13, 16–18]. For system requirements and installation instructions, please refer to the documentation [19].

---

### 3 Methods

Protocols for NA sample preparation in smFRET studies can be found elsewhere [11, 20]. In the present section, we summarize the necessary steps for SMV processing in the scope of TIRF-based experiments [16]. Next, we give a protocol for the determination of a global state configuration with the associated transition rate constants from a set of smFRET trajectories. The procedure is divided into four steps: (1) the determination of an observed FRET state configuration that describes the whole sample, (2) the identification of degenerate FRET states, (3) the inference of optimum transition rate constants, and (4) the validation of the model with simulation.

#### 3.1 Obtaining FRET Trajectories from SMVs

1. Single molecules are localized on the SMVs using the *houghpeaks* method, and fluorescence emission intensity trajectories are created as described elsewhere [18]. In short, the pixel values laying in a  $5 \times 5$  region around the molecule are summed up, yielding intensities in imaging counts (IC). To allow the comparison between experiments that use different sampling times, intensities are normalized to IC per second.
2. Intensity trajectories are individually corrected for the background signal using the *median* method [21]. Intensities detected in the acceptor channel are corrected for bleed-through of donor photon emission and acceptor direct excitation as explained elsewhere [22].
3. Correct labeling is ensured by selecting trajectories displaying both donor and acceptor emission above background level. In the scope of experiments using alternating laser excitation (ALEX), single dye emission levels are estimated from single-bleaching-step trajectories upon the respective dye-specific laser excitations. The fluorescence intensity trajectories selected for further analysis fulfill the following criteria:

- (a) Emission intensity  $I_{A_{\text{em}}}^{A_{\text{ex}}}(t)$  of the acceptor detected at wavelength  $A_{\text{em}}$  after acceptor excitation at wavelength  $A_{\text{ex}}$  must appear stable and correspond to single acceptor emission level along the observation time, i.e., exempted of defocusing, blinking, and interfering signals of neighboring molecules.
- (b) Summed emission intensities of donor and acceptor  $I_{A_{\text{em}}+D_{\text{em}}}^{D_{\text{ex}}}(t) = I_{D_{\text{em}}}^{D_{\text{ex}}}(t) + I_{A_{\text{em}}}^{D_{\text{ex}}}(t)$  after donor excitation at wavelength  $D_{\text{ex}}$  must appear stable and should correspond to a single donor emission level without energy transfer along the observation time.
4. If any, the photobleached portion of intensity trajectories is clipped out to prevent the contribution of artificial and aberrant FRET values.
5. The trajectories for fluorescence resonance energy transfer  $FRET(t)$  and stoichiometry  $S(t)$  are calculated according to [18]

$$FRET(t) = \frac{I_{A_{\text{em}}}^{D_{\text{ex}}}(t)}{I_{A_{\text{em}}}^{D_{\text{ex}}}(t) + \gamma I_{D_{\text{em}}}^{D_{\text{ex}}}(t)} \quad (1)$$

$$S(t) = \frac{I_{A_{\text{em}}}^{D_{\text{ex}}}(t) + \gamma I_{D_{\text{em}}}^{D_{\text{ex}}}(t)}{I_{A_{\text{em}}}^{D_{\text{ex}}}(t) + \gamma I_{D_{\text{em}}}^{D_{\text{ex}}}(t) + \beta I_{A_{\text{em}}}^{D_{\text{ex}}}(t)}. \quad (2)$$

FRET values can be corrected for possible variations in quantum yield and photon detection efficiency using the  $\gamma$ -correction in order to make the conversion to true inter-fluorophore distances possible [22–25]. In addition to the  $\gamma$ -correction, the stoichiometry can also be corrected for differences in laser intensities and fluorophore absorption cross-sections, both included in the  $\beta$ -correction [26]. The coefficients  $\gamma$  and  $\beta$  are calculated from the uncorrected  $FRET(t)$  and  $S(t)$  signals as described in the literature [26, 27]. The time traces selected for analysis exhibit both dyes, which results in an average stoichiometry of  $\bar{S} \simeq 0.5$ .

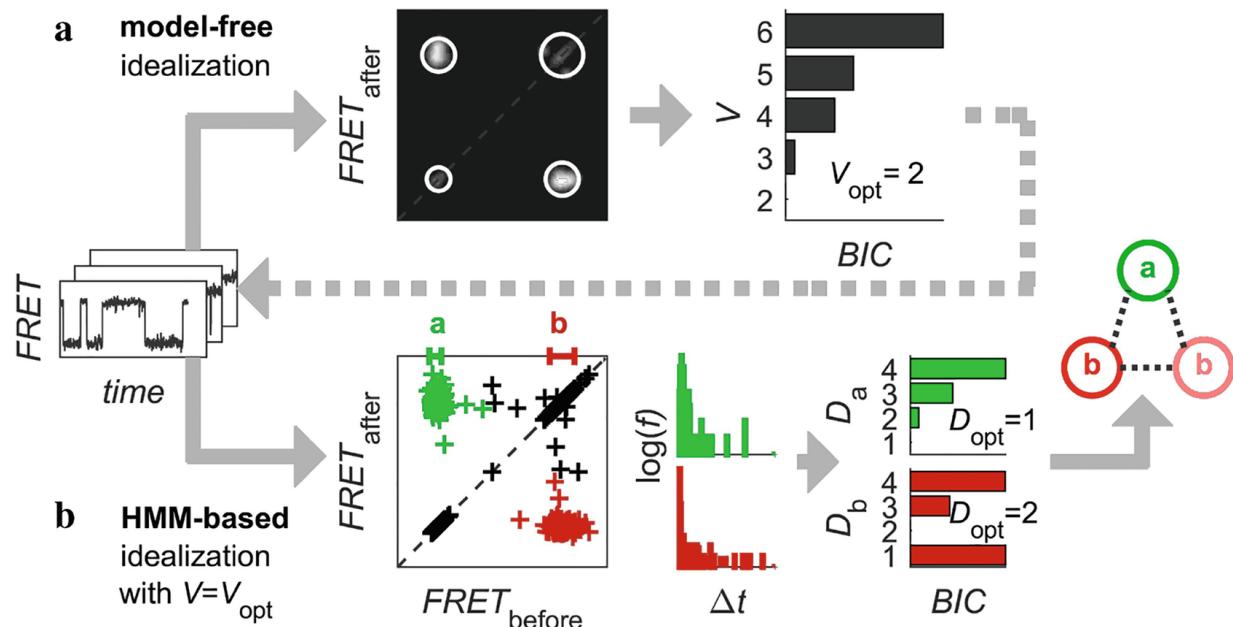
6. Selected single molecules can be classified into “dynamic” and “static” according to the presence or the absence of FRET state transitions, i.e.

according to the presence or the absence of true state transitions, i.e., detectable by eye or using state-finding algorithms as discussed below [13]. Please note that excluding the so-called static trajectories from the

analysis reduces the size of available data and induces a sampling bias, which propagates into a loss of accuracy in the transition rate constants [14].

### 3.2 Observed State Configuration and Dwell Times

In a first place, statistical processing is performed on individual FRET trajectories to get discrete FRET state sequences. The procedure is illustrated for the EBS–IBS interaction in Fig. 2a. To obtain the most reliable results, we employ an algorithm that solely relies on the signal values to detect state transitions and that has been proven to perform the best in the scope of this protocol [13]. This algorithm does not model the state transition kinetics and will later be referred to as a “model-free” method [28].



**Fig. 2** Illustration of the protocol step 3.2 for the interactions between a surface-immobilized Cy3-labeled RNA hairpin EBS with its cognate IBS, labeled with the Cy5 dye and free in solution. (a) (left) Individual FRET trajectories were first discretized with a model-free algorithm called STaSI. (middle) The TDP shows two distinct off-diagonal and two faint on-diagonal peaks that were well described by the ML-GMM method. (right) The BIC plot presents the BIC values calculated for GMMs with  $V = 2$  to 6 observed states,

yielding a minimum  $BIC$  for  $V_{\text{opt}} = 2$  observed states named “a” and “b.” (b) (left) FRET trajectories were then discretized with an HMM-based algorithm called vbFRET restricted to  $V = 2$  observed states. (middle) Trajectories’ states were subsequently hard-assigned to  $V = 2$  observed states with the ML-GMM method, allowing to build the associated accurate dwell time histograms. (right) Each histogram was analyzed by the BIC-ML-DPH algorithm: the  $BIC$  values were computed for models with  $D = 1$  to 4 degenerate states, yielding a minimum  $BIC$  value for  $D_{a,\text{opt}} = 1$  and  $D_{b,\text{opt}} = 2$  degenerate states, respectively, as depicted in the *Trellis* diagram on the left

To obtain one global state configuration, the  $O$  trajectory states are grouped based on the similarity of their FRET values. When studying large NAs, the time averaging of short-lived states combined with the flexibility of certain conformations induces a significant broadening of the FRET distribution peaks, to the extent of making them indistinguishable. To circumvent this problem, the distribution peaks are split along a second dimension—the FRET value after state transition—by building a transition density plot (TDP) [29]. The optimum peak configuration is determined by maximizing the log-likelihood  $\ell$  of 2D isotropic Gaussian mixture models (ML-GMM) for the TDP. The Gaussian centers are locked on a  $V \times V$  grid defined by the  $V$  global FRET states. The diagonal Gaussians are used to group the low-amplitude state jumps that are usually noise-induced artifacts with the trajectories’ last states to correct the former and maximize the available data with the latter. The number of FRET states is increased from  $V = 2$  to 10, and the model that renders the lowest Bayesian information criterion ( $BIC$ ), calculated with Eq. (3), is selected. From this, we get a global number of FRET states  $V_{\text{opt}}$  that shrinks the dimensionality of the problem down to one observed state configuration.

$$BIC = [V(1 + 2V) - 1] \ln(O) - 2l. \quad (3)$$

As demonstrated previously, model-free trace discretization is a powerful method to estimate the dimensions of a state configuration but shows low efficiency in resolving short-lived states, which in turn skews state kinetics [13]. For this reason, smFRET trajectories are discretized more accurately by optimizing HMMs constrained to  $V_{\text{opt}}$  observed states and using variational Bayes as shown in Fig. 2b [30]. Trajectories’ states are grouped into  $V_{\text{opt}}$  FRET states with ML-GMM using diagonal multivariate Gaussians. The errors on global FRET state values are computed by averaging the standard deviations of the respective 2D-Gaussian, each

calculated from Gaussian's co-variance matrices and averaged for state  $i$  over all  $i$ -to- $j$  transitions with  $i \neq j$ . Finally, dwell time distributions are built after hard-assigning each data point of each trajectory to the most likely global FRET state using the maximum 2D-Gaussian probabilities. The results obtained for EBS-IBS are presented in Fig. 2b.

It is important to mention that the results obtained at the end of this protocol strongly depend on the approach used to assign trajectories' states. Here, we choose BIC-ML-GMM for objective clustering, but other approaches like FRET 1D-histogram fitting, the nearest-neighbor grouping method, or ensemble HMM approaches exist, which might lead to differences in the outcome model [13, 14, 31]. The present protocol can be adapted by substituting the current state grouping method with the desired one.

### 3.3 State Degeneracy

Due to the complexity of NA molecular structures and the limited capacity of two-color FRET experiments to probe one-dimensional distances, a single observed FRET value often stands for multiple conformational states. The presence of such degenerate states usually breaks the simple exponential shape of the observed dwell time distribution, introducing sums and convolutions of multiple distributions. In this protocol, state degeneracy is estimated from the shape of ensemble dwell time histograms, which requires considerably less computational power than when inferred from single-state trajectories [14, 32, 33].

In state trajectories, the dwell times are said to be “phase-type” distributed, i.e., that each dwell time can be regarded as an absorbing Markov chain in which  $D$  degenerate states—or phases—are interconverting prior transitioning into a virtual “absorbing state” [34]. In addition, when working with discrete-time trajectories as in TIRF-based experiments, the dwell times follow a discrete phase-type distribution (DPH) defined by

$$f_v(t) \sim \text{DPH}_v(\pi, Q) = \pi Q^{t-1} (e - Qe), \quad (4)$$

where  $f_v(t)$  is the dwell time distribution for the observed state  $v$ ,  $\pi$  is the probability vector to start a dwell time in one of the  $D_v$  degenerate states, the  $D_v \times D_v$  matrix  $Q$  contains the transition probabilities between the

phases of the absorbing Markov chain, and  $e$  is a column vector filled with ones. To predict the state degeneracy that most sufficiently describes the dwell time histogram, multiple DPHs with a number of phases ranging from 1 to 4 are optimized with the expectation–maximization algorithm described elsewhere [35], using five random starting guesses of  $\pi$  and  $Q$ . The *BIC* calculated with

$$BIC_v = D_v(D_v - 1) \ln(O_v) - 2l \quad (5)$$

is used to select the optimum state degeneracy. The operation is repeated for each of the  $V$  dwell time histograms until the full global degenerate state configuration is obtained as shown in Fig. 2b for the EBS-IBS system. The total number of states  $J$  in the final state configuration sums up to all degenerate levels of all observed states, i.e.,  $J = \sum D_v$ .

### 3.4 State Transition Rate Constants

With the state degeneracies and the hard assignment of trajectories' states at hand, the kinetic model must now be completed with the determination of the transition rate constants. In conventional systems, i.e., without degenerate states, the observed dwell time histograms follow simple exponential distributions where transition rate constants can be deduced from the exponential constants. Even though the use of ensemble distributions to determine the transition rate constants has been judged less accurate than probabilistic inference [14], it remains widely employed for standard systems. Indeed, the ease of implementation and the competitively small computation costs make it an attractive choice for fast quantification of transition kinetics. In addition, when modeling the dwell time distribution with an exponential function, the time continuity of the underlying conformational dynamics is taken into account, which minimizes the underestimation of transition rate constants induced by time sampling of FRET trajectories. However, in the presence of degenerate states, the histograms lose these simple shapes to become sums and convolutions of multiple exponential distributions. In this case, the quantification of state kinetics requires a more sophisticated approach based on probabilistic inference.

#### 3.4.1 Standard State Systems

The rate constant  $k_{ij}$  that governs transitions from state  $i$  to state  $j$  depends on the lifetime of state  $i$  and on the probability  $w_{ij}$  for state  $i$  to preferably transit to state  $j$  rather than to all the other states, later referred to as the transition *repartition probability*, and calculated according to [36]

$$k_{ij} = \frac{w_{ij}}{\tau_i}. \quad (6)$$

State lifetimes  $\tau_i$  are estimated by fitting single exponential functions to the dwell time histograms with the fitting model

$$1 - F(\Delta t_i) = \exp\left(\frac{-\Delta t_i}{\tau_i}\right). \quad (7)$$

As time-binned data suffer from the absence of very short dwell times, the normalized complementary cumulative histogram  $1 - F(\Delta t)$  of dwell times  $\Delta t$  is used instead of raw counts. This minimizes the impact of the first histogram bins while preserving the overall shape. The repartition probabilities  $w_{ij}$  of transitions  $i$ -to- $j$  are calculated from transition counts  $N_{ij}$  with

$$w_{ij} = \frac{N_{ij}}{\sum_{k \neq i} N_{ik}}. \quad (8)$$

The outcomes of such analysis are single estimates of the rate constants. One way to evaluate the estimation error is to calculate the variability  $\sigma_{\tau,i}$  of  $\tau_i$  across the sample using a bootstrap-based approach [17]. The variability can then be propagated to  $k_{ij}$

$$\sigma_{k_{ij}} = \frac{\sigma_{\tau,i}}{\tau_i} \times k_{ij}, \quad (9)$$

where  $\sigma_{k_{ij}}$  is the error on rate constant  $k_{ij}$  and  $\sigma_{\tau,i}$  is the bootstrap standard deviation of parameter  $\tau_i$ . If we assume normally distributed rate constants, the value range  $[k_{ij} - 2\sigma_{k_{ij}}, k_{ij} + 2\sigma_{k_{ij}}]$  represents a confidence interval of about 95.4%.

### 3.4.2 Degenerate State Systems

Using state trajectories instead of ensemble dwell time histograms becomes indispensable when solving models with kinetic heterogeneity. This allows to keep track of the sequential order of states, and thus, to quantify forward and backward transition probabilities between all degenerate states.

Here, we apply the Baum–Welch algorithm [38, 37] to state sequences, i.e., to noiseless trajectories, in which the state assignment is inflexible. This means that the emission probabilities used to evaluate the observed state the trajectory is in are equal to either 0 or 1. Therefore, the algorithm only optimizes the vector of initial state probabilities  $\pi$  and the transition probability matrix  $P$ , respectively, defined as

$$P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1J} \\ p_{21} & p_{22} & \dots & p_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ p_{J1} & p_{J2} & \dots & p_{JJ} \end{pmatrix} \quad \text{and} \quad \pi = \begin{pmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_J \end{pmatrix} \quad (10)$$

by iterating expectation and maximization of the HMM likelihood until it eventually converges to a maximum likelihood estimator (MLE) of the model parameters. An illustration of the algorithm applied to the interaction of IBS with EBS is presented in Fig. 3a. Off-diagonal transition probabilities in  $P$  are then converted into the respective rate constants using the trajectory sampling time  $t_{\text{exp}}$

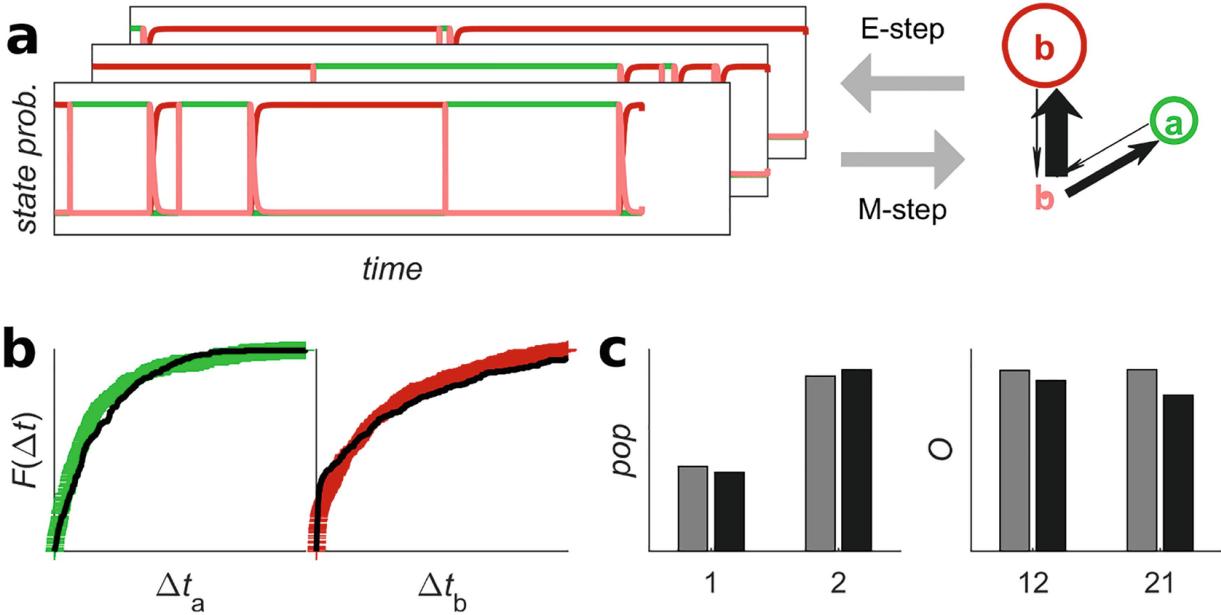
$$k_{ij} \approx \frac{p_{ij}}{t_{\text{exp}}} \quad (11)$$

Equation (11) is only an approximation because transition probabilities are calculated from discrete-time trajectories, whereas the transition rate constants describe processes that are continuous in time. As multiple conformational state transitions can happen within one recording time step, true rate constants will always be larger than their estimation from time-binned data [39]. Similarly, the state lifetimes can be approximated from the transition probabilities

$$\tau_i \approx \frac{t_{\text{exp}}}{\sum p_{ij}} \quad (12)$$

Negative and positive errors  $\sigma_{k,ij}^-$  and  $\sigma_{k,ij}^+$  on rate constants are estimated via a 95% confidence likelihood ratio test described elsewhere, giving an

estimated range delimited by the lower bound  $k_{ij} - \sigma_{k,ij}^-$  and the upper bound  $k_{ij} + \sigma_{k,ij}^+$  [33].



**Fig. 3** Illustration of the protocol steps 3.4 and 3.5 for the interactions between EBS and IBS. (a) State probability trajectories (left) were inferred via the “E-step” of the Baum–Welch algorithm using random HMM parameters (right). HMM parameters were recalculated from the probability trajectories via the “M-step” by maximizing the likelihood. The HMM includes one state with value “a” (green) and two states sharing the same value “b” but having a long (dark red) and very short (light red) lifetime, respectively. E-M iterations were repeated until convergence to the HMM presented on the right. (b) The comparison between simulated (solid black line) and experimental (color-coded crosses) cumulative dwell time distributions proves that the inferred model describes well the experimental data set as little deviation is observed. (c) The time fraction (left) spent in each observed state as well as the number of observed transitions (right) is almost identical between the simulated (dark gray) and experimental (light gray) data, which validates the use of the given model to describe the data set

### 3.5 Model Validation

To ensure the validity of the inferred model, a set of synthetic state trajectories is produced using the kinetic model parameters, i.e., the initial state probabilities and transition rate constants, and the experimental mensuration, i.e., the sample size  $N$ , the trajectory lengths  $L_n$ , and the sampling time  $t_{\text{exp}}$ , which is then compared to the experimental data set [36].

As mentioned above, the tracking over time of single-molecule conformational exchanges can be modeled as a Markov chain. In camera-based single-molecule experiments, the FRET signal is binned in time by the camera, yielding discrete-time FRET trajectories that report on the average conformation of the molecule within each time bin. However, the underlying conformational dynamics of the molecule are fundamentally continuous in time and must be represented as such in the simulation. Therefore, continuous-time Markov chains are simulated by building sequences of dwell times during each of which the chain is in a particular state. For each simulated chain, the state starting the sequence is randomly drawn from the initial state distribution  $\pi$  obtained in Subheading 3.4. For a chain being in state  $i$ , the transition probability to any other state increases exponentially with the time  $\Delta t_i$  spent in state  $i$

$$p_{i,\text{trans}}(\Delta t_i) \sim 1 - \exp\left(\frac{-\Delta t_i}{\tau_i}\right). \quad (13)$$

Accordingly, the dwell times in state  $i$  follow an exponential distribution characterized by the state lifetime  $\tau_i$  determined in Subheading 3.4.

Therefore, the dwell times used in the simulated chains are randomly drawn from such distributions using the MATLAB's built-in function *random*. The decision about which state the chain is transitioning to is stochastically determined by the transition probabilities  $w_{ij}$  introduced in the previous section. These probabilities can be calculated from transition rate constants using

$$w_{ij} = \frac{k_{ij}}{\sum_{k \neq i} k_{ik}}. \quad (14)$$

Equation (14) is obtained by combining Eq. (8) with the definition of a rate constant as the number of transitions per unit of time calculated over the total observation time  $T$  such as

$$k_{ij} = \frac{N_{ij}}{T}. \quad (15)$$

Next states and random dwell times are successively drawn until the experimental trajectory length  $L_n$  is reached. Finally, the sequences are binned in time using  $t_{\text{exp}}$  to simulate the time-averaged camera-based FRET

signals. The artifactual time-averaged states are assigned to the real state the closest in value to mimic the simplest state-finding algorithm.

Normalized dwell time histograms, observed state populations, and the number of observed transitions are calculated from the simulated state sequences and are visually compared to the experimental values. Such a comparison is exemplified in Fig. 3b, c for the EBS–IBS interaction.

## 3.6 User Guide

In the following, we give a step-by-step user guide for the execution of Subheadings 3.2, 3.3, and 3.5 with the software MASH-FRET. The procedure starts with a set of text files containing single-molecule intensity-time traces and yields a global kinetic model with error estimations of FRET state values and rate constants. The condensed procedure can be executed via a routine as described in Subheading 3.6.6.

### 3.6.1 *Observed State Configuration*

1. In the project management area, import the trace files:
  - (a) Press the “Create a new project” button: the data type to import is now requested.
  - (b) Select “Import trajectories” to open the “Experiment settings” window.
  - (c) Select the trace files, adjust the experiment settings as well as the FRET/stoichiometry calculations and set the import options according to the structure of the files.
  - (d) Press “Save” to load trajectories in MASH-FRET: the interface automatically switches to the trace processing module.
2. In panel “Find states,” configure the state-finding algorithm:
  - (a) Set “method” to the model-free algorithm “STaSI.”
  - (b) Set options “apply to” to “bottom” and “data” to “FRET” to idealize FRET time traces only.

~~Normalize FRET 1-minute traces only.~~

- (c) Set the first parameter “param” to a large value, e.g. 10, to remove any constraint on the model dimensions.
  - (d) Press “All” to apply these settings to all trajectories.
3. Process all trajectories by pressing “UPDATE ALL.”
4. Go to the “Transition analysis” module.
5. In panel “Transition density plot,” configure the TDP:
- (a) Set option “data” to “FRET” to select all FRET state trajectories.
  - (b) Set the TDP bounds to  $[-0.2;1.2]$  and the bin size to 0.01 to obtain a good compromise between the separation of transition clusters and the processing time.
  - (c) Activate options “incl. statics,” “Gaussian filter,” and “Norm.” to, respectively, include static trajectories in diagonal bins, smooth the TDP with a Gaussian filter, and display normalized TDP counts.
  - (d) Refresh the TDP counts by pressing “Update”: the TDP is now shown in the visualization area.
6. In panel “State configuration,” configure the TDP fitting or transition clustering algorithm:
- (a) Set the method to “GM” to use the ML-GMM algorithm.
  - (b) Set the maximum model complexity “ $V_{\max}$ ” and the number of model initialization “restart” to large values, e.g. 10 and 50, respectively, to test a large range of model dimensions and ensure the convergence of ML-GMM to a global maximum of the likelihood.

7.

In panel “Clusters,” configure the 2D-Gaussian clusters:

- (a) Set the “Cluster constraints” option to “matrix” to lock the Gaussian means on a  $V \times V$  grid.
- (b) Activate the “diagonal clusters” option to include the so-called static’ state trajectories and exclude the low-amplitude state jumps from the fitting procedure.
- (c) Set the cluster shapes as isotropic Gaussian by pressing the button having a perfect circle drawn on it.

8.

Press “cluster” to perform TDP fitting: the optimum GMM and state partition are shown overlaying the TDP.

9.

In panel “Results,” note down the optimum number of FRET states “ $V_{\text{opt.}}$ ”

10.

In the project management area, save the modifications to a \*.mash file by pressing the “Save project” button.

### ***3.6.2 Accurate Determination of Dwell Times and FRET States***

1.

Go to the “Trace processing” module.

2. In panel “Find states,” configure the state-finding algorithm:

- (a) Set “method” to “vbFRET” to use the variation Bayes inference of HMMs.
- (b) Set the first parameter “param” to  $V_{\text{opt}}$  to constrain the model dimensions, and parameter “deblur” to 1 to remove artifactual blur states, i.e., false identification due to time averaging of state transitions.

- (c) Press “All” to apply these settings to all trajectories.
3. Process all traces by pressing “UPDATE ALL”: accurate state trajectories are eventually obtained for all molecules.
  4. Go to the “Transition analysis” module.
  5. In panel “Transition density plot,” refresh the TDP counts by pressing “Update.”
  6. In panel “State configuration,” re-configure the clustering algorithm by setting the maximum model dimensions “ $V_{\max}$ ” to  $V_{\text{opt}}$  to limit the model complexity and computation time to what is strictly necessary.
  7. In panel “Clusters,” set the 2D-Gaussian shape to “diagonal multivariate Gaussian” by pressing the button having a diagonal ellipse drawn on it.
  8. Press “cluster” to refresh TDP fitting and perform hard assignment of trajectory states. The new GMM and partition of states are shown overlaying the TDP.
  9. In panel “Results,” select the proper observed state configuration regardless of the outcome of TDP fitting by setting the option “ $V$ ” to  $V_{\text{opt}}$  and pressing “use this config.” to use it for further analysis.

### ***3.6.3 Estimation of State Degeneracy***

1. In panel “Kinetic model,” configure the ML-DPH algorithm by setting the maximum degeneracy “ $D_{\max}$ ” to 4 and the number of model initializations “restart” to 5 to restrict the model dimensions and the computation time to reasonable values.
2. Press “Estimate” to start the estimation of state degeneracies: the visualization area switches to the *BIC* plot and shows the optimum state complexity.

### **3.6.4 Determination of State Transition Rate Constants**

#### **Standard System**

1. In panel “Dwell time histograms,” configure the fitting algorithm:
  - (a) Set the “state binning” to 0.01 to merge the observed states having very close values.
  - (b) Activate the option “exclude first and last” to exclude the dwell times truncated by the observation’s beginning and end.
  - (c) Activate the option “recalc” to remove artifactual transitions of the state sequences contained in the diagonal clusters.
  - (d) Press “Fit settings” to open the fit settings window.
  - (e) For each state in the “state value” list, activate the options “nb. of decays” and “BOBA” to perform bootstrap exponential fit and set “nb. of decays” to 1 to perform single exponential fit.
2. Press “Fit all” to start bootstrap exponential fits: the visualization area shows the fit functions together with the lowest and largest bootstrap deviations overlaying the complementary cumulative normalized dwell time histograms.
3. Go to MASH-FRET’s menu *Export → Transition analysis* to open the “Export options” window and export the results to text files:
  - (a) Activate the options “clusters,” “fitting curves & parameters (.fit)” and “BOBA FRET results (.fit)” to, respectively, export the GMM and bootstrap exponential fit parameters.
  - (b) Press “Next >>” to start exporting the files.
4. In the project management area, press the “Save project” button to save the analysis to a new .mash file.  
*Calculate the transition rate constants.*

5. Calculate the transition rate constants.

- (a) Calculate the transition probabilities  $w_{ij}$  using Eq. (8) and the cluster populations found in the .clst file.
- (b) Calculate the rate constants  $k_{ij}$  using Eq. (6) and the bootstrap mean of the exponential time constants found in the .fit files.
- (c) Calculate the error on rate constants  $\sigma_{k,ij}$  using Eq. (9) and the bootstrap standard deviation of the exponential time constants found in the .fit files.

### ***System with Degenerate States***

1. In panel “Kinetic model,” configure and execute the Baum–Welch algorithm:
  - (a) Set the number of model initializations to 5 to get a compromise between accuracy and computational cost.
  - (b) Press “Start” to start model inference: the optimum model is shown as a Trellis diagram where state transitions are depicted by arrows having a width proportional to the associated transition probability.
2. Go to MASH-FRET’s menu *Export → Transition analysis* to open the “Export options” window and export the model parameters and associated errors to a text file:
  - (a) Activate the option “model parameters (\*.mdl.txt)” to export the HMM parameters.
  - (b) Press “Next>>” to start exporting the files.
3. In the project management area, press the “Save project” button to save the analysis to a new .mash file.

### **3.6.5 Model Validation with Simulation**

1. In panel “Kinetic model,” generate a synthetic data set based on the inferred state configuration and transition rate constants:
  - (a) Set the source “model from” to “Exponential fit” for standard systems, or “Baum-Welch” for systems with degenerate states.
  - (b) Press “Simulate” to start simulation: the visualization area shows three plots comparing simulation and experimental data.
2. In the tabbed panel “Simulation” of the visualization area, validate the infer model by evaluating the deviation of the simulation dwell time histograms, state populations, and state transition from the experimental ones.

### **3.6.6 Condensed Procedure**

To perform all user guide steps at once, please execute the following:

1. Go to MASH-FRET’s menu *Routines* → *Standard kinetic analysis* → *All steps* and select the set of text files to analyze.
2. A first message box pops up: enter the number of FRET states if known or press “No” otherwise. If the number of FRET states is known, the first part of Subheading 3.2 will be skipped.
3. A second message box pops up: choose the proper noise distribution according to your data set: press “Gaussian noise” if the intensities are Gaussian-distributed in the trajectories or press “other” otherwise. If the noise is Gaussian-distributed, the second part of Subheading 3.2 will be performed to obtain accurate dwell times.
4. Once the analysis routine is completed, the analysis results are all summarized in one file [...]\_results.txt saved at the same location as the data files.

## **Acknowledgements**

Financial support from the Swiss National Science Foundation [to R.K.O.S], and the University of Zurich [to R.K.O.S] and the University of Applied Sciences Mittweida [to R.B.] is gratefully acknowledged.

---

## References

1. Paudel B, Fiorini E, Börner R, Sigel R, Rueda D (2018) Optimal molecular crowding accelerates group II intron folding and maximizes catalysis. *Proc Natl Acad Sci U S A* 115:11917–11922. <https://doi.org/10.1073/pnas.1806685115> [Crossref]
2. Zhao M, Steffen F, Börner R, Schaffer M, Sigel R, Freisinger E (2018) Site-specific dual-color labeling of long RNAs for single-molecule spectroscopy. *Nucleic Acids Res* 46:e13. <https://doi.org/10.1093/nar/gkx1100> [Crossref]
3. Steiner M, Rueda D, Sigel R (2009) Ca<sup>2+</sup> induces the formation of two distinct subpopulations of group II intron molecules. *Angew Chem Int Ed Engl* 48:9739–9742. <https://doi.org/10.1002/anie.200903809> [Crossref]
4. Börner R, Kowerko D, Miserachs HG, Schaffer MF, Sigel R (2016) Metal ion induced heterogeneity in RNA folding studied by smFRET. *Coord Chem Rev* 327:123–142. <https://doi.org/10.1016/j.ccr.2016.06.002> [Crossref]
5. Ha T, Enderle T, Ogletree D, Chemla D, Selvin P, Weiss S (1996) Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc Natl Acad Sci U S A* 93:6264–6268. <https://doi.org/10.1073/pnas.93.13.6264> [Crossref]
6. Walter N, Harris D, Pereira M, Rueda D (2001) In the fluorescent spotlight: global and local conformational changes of small catalytic RNAs. *Biopolymers* 61:224–242. <https://doi.org/10.1002/bip.10144> [Crossref]
7. Deniz A, Dahan M, Grunwell J, Ha T, Faulhaber A, Chemla D, Weiss S, Schultz P (1999) Single-pair fluorescence resonance energy transfer on freely diffusing molecules: observation of Förster distance dependence and subpopulations. *Proc Natl Acad Sci U S A* 96:3670–3675. <https://doi.org/10.1073/pnas.96.7.3670> [Crossref]
- 8.

- Blanco M, Walter N (2010) Analysis of complex single-molecule FRET time trajectories. *Methods Enzymol* 472:153–178. [https://doi.org/10.1016/S0076-6879\(10\)72011-5](https://doi.org/10.1016/S0076-6879(10)72011-5)
9. Schuler B, Hofmann H (2013) Single-molecule spectroscopy of protein folding dynamics—expanding scope and timescales. *Curr Opin Struct Biol* 23:36–47. <https://doi.org/10.1016/j.sbi.2012.10.008>  
[Crossref]
10. Lamichhane R, Solem A, Black W, Rueda D (2010) Single-molecule FRET of protein-nucleic acid and protein-protein complexes: surface passivation and immobilization. *Methods* 52:192–200. <https://doi.org/10.1016/j.ymeth.2010.06.010>  
[Crossref]
11. Zelger-Paulus S, Hadzic M, Sigel R, Börner R (2020) Encapsulation of fluorescently labeled RNAs into surface-tethered vesicles for single-molecule FRET studies in TIRF microscopy. *Methods Mol Biol* 2113:1–16. [https://doi.org/10.1007/978-1-0716-0278-2\\_1](https://doi.org/10.1007/978-1-0716-0278-2_1)  
[Crossref]
12. Holden S, Uphoff S, Hohlbein J, Yadin D, Le R, Britton O, Kapanidis A (2010) Defining the limits of single-molecule FRET resolution in TIRF microscopy. *Biophys J* 99:3102–3111. <https://doi.org/10.1016/j.bpj.2010.09.005>  
[Crossref]
13. Hadzic M, Börner R, König S, Kowerko D, Sigel R (2018) Reliable state identification and state transition detection in fluorescence intensity-based single-molecule Förster resonance energy-transfer data. *J Phys Chem B* 122:6134–6147. <https://doi.org/10.1021/acs.jpcb.7b12483>  
[Crossref]
14. Schmid S, Götz M, Hugel T (2016) Single-molecule analysis beyond Dwell times: demonstration and assessment in and out of equilibrium. *Biophys J* 111:1375–1384. <https://doi.org/10.1016/j.bpj.2016.08.023>  
[Crossref]
15. Steffen F, Khier M, Kowerko D, Cunha R, Börner R, Sigel R (2020) Metal ions and sugar puckering balance single-molecule kinetic heterogeneity in RNA and DNA tertiary contacts. *Nat Commun* 11:104. <https://doi.org/10.1038/s41467-019-13683-4>  
[Crossref]
16. Hadzic M, Kowerko D, Börner R, Zelger-Paulus S, Sigel R (2016) Detailed analysis of complex single molecule FRET data with the software MASH. *Proc SPIE* 9711:971119. <https://doi.org/10.1117/12.2211191>

[Crossref]

17. König S, Hadzic M, Fiorini E, Börner R, Kowerko D, Blanckenhorn W, Sigel R (2013) BOBA FRET: bootstrap-based analysis of single-molecule FRET data. PLoS One 8:e84157. <https://doi.org/10.1371/journal.pone.0084157>  
[Crossref]
18. Börner R, Kowerko D, Hadzic M, König S, Ritter M, Sigel R (2018) Simulations of camera-based single-molecule fluorescence experiments. PLoS One 13:e0195277. <https://doi.org/10.1371/journal.pone.0195277>  
[Crossref]
19. Hadzic M (2021) MASH-FRET: multifunctional analysis software for handling single molecule FRET data. Matlab-based. [https://rna-fretools.github.io/MASH-FRET/Getting\\_started.html](https://rna-fretools.github.io/MASH-FRET/Getting_started.html). Accessed at November 19, 2021
20. Cardo L, Karunatilaka K, Rueda D, Sigel R (2012) Single molecule FRET characterization of large ribozyme folding. Methods Mol Biol 848:227–251. [https://doi.org/10.1007/978-1-61779-545-9\\_15](https://doi.org/10.1007/978-1-61779-545-9_15)  
[Crossref]
21. Preus S, Hildebrandt L, Birkedal V (2016) Optimal background estimators in single-molecule FRET microscopy. Biophys J 111:1278–1286. <https://doi.org/10.1016/j.bpj.2016.07.047>  
[Crossref]
22. Hellenkamp B, Schmid S, Doroshenko O, Opanasyuk O, Kühnemuth R, Rezaei A, Ambrose B, Aznauryan M, Barth A, Birkedal V, Bowen M, Chen H, Cordes T, Eilert T, Fijen C, Gebhardt C, Götz M, Gouridis G, Gratton E, Ha T, Hao P, Hanke C, Hartmann A, Hendrix J, Hildebrandt L, Hirschfeld V, Hohlbein J, Hua B, Hübner C, Kallis E, Kapanidis A, Kim J, Krainer G, Lamb D, Lee N, Lemke E, Levesque B, Levitus M, McCann J, Naredi-Rainer N, Nettels D, Ngo T, Qiu R, Robb N, Röcker C, Sanabria H, Schlierf M, Schröder T, Schuler B, Seidel H, Streit L, Thurn J, Tinnefeld P, Tyagi S, Vandenberk N, Vera A, Weninger K, Wünsch B, Yanez-Orozco I, Michaelis J, Seidel C, Craggs T, Hugel T (2018) Precision and accuracy of single-molecule FRET measurements-A multi-laboratory benchmark study. Nat Methods 15:669–676. <https://doi.org/10.1038/s41592-018-0085-0>  
[Crossref]
23. Lerner E, Barth A, Hendrix J, Ambrose B, Birkedal V, Blanchard S, Börner R, Sung C, Cordes T, Craggs T, Deniz A, Diao J, Fei J, Gonzalez R, Gopich I, Ha T, Hanke C, Haran G, Hatzakis N, Hohng S, Hong S, Hugel T, Ingargiola A, Joo C, Kapanidis A, Kim H, Laurence T, Lee N, Lee T, Lemke E, Margeat E, Michaelis J, Michalet X, Myong S, Nettels D, Peulen T, Ploetz E, Razvag Y, Robb N, Schuler B,

- Soleimaninejad H, Tang C, Vafabakhsh R, Lamb D, Seidel C, Weiss S (2021) FRET-based dynamic structural biology: challenges, perspectives and an appeal for open-science practices. *elife* 10:e60416. <https://doi.org/10.7554/eLife.60416>
24. McCann J, Choi U, Zheng L, Weninger K, Bowen M (2010) Recovering absolute FRET efficiency from single molecules: comparing methods of gamma correction. *Biophys J* 98:186a–187a
  25. Hohlbein J, Craggs T, Cordes T (2014) Alternating-laser excitation: single-molecule FRET and beyond. *Chem Soc Rev* 43:1156–1171. <https://doi.org/10.1039/c3cs60233h> [Crossref]
  26. Lee N, Kapanidis A, Wang Y, Michalet X, Mukhopadhyay J, Ebright R, Weiss S (2005a) Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys J* 88:2939–2953. <https://doi.org/10.1529/biophysj.104.054114> [Crossref]
  27. Kapanidis A, Laurence T, Lee N, Margeat E, Kong X, Weiss S (2005) Alternating-laser excitation of single molecules. *Acc Chem Res* 38:523–533. <https://doi.org/10.1021/ar0401348> [Crossref]
  28. Shuang B, Cooper D, Taylor J, Kisley L, Chen J, Wang W, Li C, Komatsuzaki T, Landes C (2014) Fast step transition and state identification (STaSI) for discrete single-molecule data analysis. *J Phys Chem Lett* 5:3157–3161. <https://doi.org/10.1021/jz501435p> [Crossref]
  29. McKinney S, Joo C, Ha T (2006) Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys J* 91:1941–1951. <https://doi.org/10.1529/biophysj.106.082487> [Crossref]
  30. Bronson J, Fei J, Hofman J, Gonzalez R, Wiggins C (2009) Learning rates and states from biophysical time series: a Bayesian approach to model selection and single-molecule FRET data. *Biophys J* 97:3196–3205. <https://doi.org/10.1016/j.bpj.2009.09.031> [Crossref]
  31. Zarabi N, Schluesche P, Meisterernst M, Börsch M, Lamb D (2018) Analyzing the dynamics of single TBP-DNA-NC2 complexes using hidden Markov models. *Biophys J* 115:2310–2326. <https://doi.org/10.1016/j.bpj.2018.11.015> [Crossref]

32. Schmid S, Hugel T (2018) Efficient use of single molecule time traces to resolve kinetic rates, models and uncertainties. *J Chem Phys* 148:123312. <https://doi.org/10.1063/1.5006604>  
[Crossref]
33. Hon J, Gonzalez R (2019) Bayesian-estimated hierarchical HMMs enable robust analysis of single-molecule kinetic heterogeneity. *Biophys J* 116:1790–1802. <https://doi.org/10.1016/j.bpj.2019.02.031>  
[Crossref]
34. Ollson M (1996) Estimation of phase-type distributions from censored data. *Scand J Stat* 23:443–460
35. Bladt M, Nielsen B (2010) Estimation of phase-type distributions. In: Bladt M, Nielsen B (eds) *Matrix-exponential distributions in applied probability. Probability theory and stochastic modelling*, vol 81. Springer, Boston, pp 617–701. [https://doi.org/10.1007/978-1-4939-7049-0\\_13](https://doi.org/10.1007/978-1-4939-7049-0_13)
36. Chen J, Pyle J, Piecco K, Kolomeisky A, Landes C (2016) A two-step method for smFRET data analysis. *J Phys Chem B* 120:7128–7132. <https://doi.org/10.1021/acs.jpcb.6b05697>  
[Crossref]
37. Sammut C, Webb G (eds) (2010) *Encyclopedia of machine learning*. Springer, Boston. <https://doi.org/10.1007/978-0-387-30164-8>
38. Munro P, Toivonen H, Webb G, Buntine W, Orbanz P, Teh Y, Poupart P, Sammut C, Blockeel H, Rajnarayan D, Wolpert D, Gerstner W, Page C, Natarajan S, Hinton G (2010) *Baum-Welch algorithm*. Springer, Boston, p 74. [https://doi.org/10.1007/978-0-387-30164-8\\_59](https://doi.org/10.1007/978-0-387-30164-8_59)
39. Farooqa S, Hohlbein J (2015) Camera-based single-molecule FRET detection with improved time resolution. *17:27862-27872.* <https://doi.org/10.1039/C5CP04137F>

## 13. Chemical Dual End-Labeling of Large Ribozymes

Esra Ahunbay<sup>1</sup>, Fabio D. Steffen<sup>1</sup>, Susann Zelger-Paulus<sup>1</sup>✉ and Roland K. O. Sigel<sup>1</sup>✉

(1) Department of Chemistry, University of Zurich, Zurich, Switzerland

✉ Susann Zelger-Paulus (Corresponding author)  
Email: [susann.paulus@chem.uzh.ch](mailto:susann.paulus@chem.uzh.ch)

✉ Roland K. O. Sigel (Corresponding author)  
Email: [roland.sigel@chem.uzh.ch](mailto:roland.sigel@chem.uzh.ch)

### Abstract

Fast and efficient site-specific labeling of long RNAs is one of the main bottlenecks limiting distance measurements by means of Förster resonance energy transfer (FRET) or electron paramagnetic resonance (EPR) spectroscopy. Here, we present an optimized protocol for dual end-labeling with different fluorophores at the same time meeting the restrictions of highly labile and degradation-sensitive RNAs. We describe in detail the dual-labeling of a catalytically active wild-type group II intron as a typical representative of long functional RNAs. The modular procedure chemically activates the 5'-phosphate and the 3'-ribose for bioconjugation with a pair of fluorophores, as shown herein, or with spin labels. The mild reaction conditions preserve the structural and functional integrity of the biomacromolecule and results in covalent, dual-labeled RNA in its pre-catalytic state in yields suitable for both ensemble and single-molecule FRET experiments.

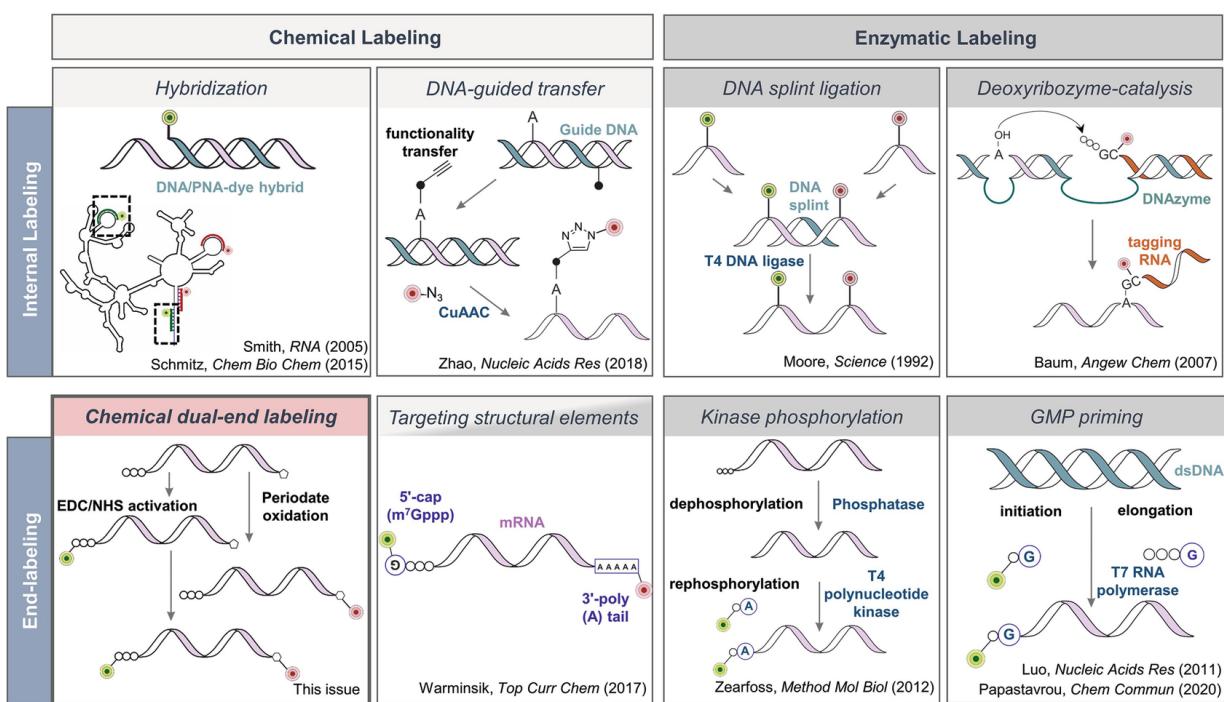
**Key words** Fluorescent RNA labeling – Ribozymes – Group II intron – Single-molecule fluorescence microscopy – FRET – Site-specific labeling – Post-transcriptional labeling

---

## 1 Introduction

Förster resonance energy transfer (FRET) has matured into a versatile technique to resolve conformational dynamics of macromolecules [1–3]. Inserting fluorophores at strategic positions on the molecule of interest is a prerequisite for any FRET experiment, but site-specific labeling of nucleic acids and especially long RNAs is challenging due to the chemical similarity of the four nucleobases. An ideal labeling method that is precise, efficient, with no perturbation of the RNA structure, and at the same time fully size-independent is still missing. Current strategies can be categorized by the position where the dye is inserted (internal or terminal), how the fluorophore is attached (non-covalent or covalent), by the coupling mechanism (chemical or enzymatic), or by the synthetic route (solid-phase synthesis or co-/post-transcriptional) of RNA preparation (Fig. 1) [4–9]. In this chapter, we focus on covalently labeling long RNAs at a defined single position where chemical synthesis with phosphoramidites is getting impractical due to decreasing yields (>100 nucleotides). Usually, long RNAs are prepared by in vitro transcription and labeled post-synthetically. One common approach is to use short, antisense oligonucleotides, such as DNA or peptide nucleic acid (PNA), carrying a fluorescent dye and annealing to internal loops of the RNA as fluorescent probes [10, 11]. With a non-covalent interaction, dye attachment is not permanent and loop regions often need to be extended to accommodate the fluorescently labeled oligonucleotide. To mitigate this issue, different DNA-guided reactions have been developed that transfer the fluorophore to the RNA by forming a stable covalent linkage [12–15]. Enzyme-assisted approaches on the other hand generate a fluorescent construct by ligating short synthetic fragments [16]. With the help of a ligase or an engineered DNase [17], riboswitches, and ribozymes with lengths up to about 200 bases can be prepared in this way [18, 19]. T4 RNA ligase and T4 polynucleotide kinase can also add a single nucleotide to an in vitro transcribed RNA at the 3'- or 5'-terminus, respectively [20]. Alternatively, a labeled guanosine monophosphate (GMP) may be used to initiate transcription with T7 RNA

polymerase to introduce a fluorophore at the 5'-end [21]. Finally, eukaryotic mRNA can be site-specifically targeted at their 5'-cap [22] and the 3'- poly(A) tail [23]. The downside of all these latter approaches is the restriction of the labeling sites to the two termini limiting the information content of FRET measurements. However, end-labeling can be complemented with internal site-specific covalent attachment of fluorophores [13]. Apart from this limitation, end-labeling is a straightforward and efficient way of visualizing nucleic acids of any size with minimal structural perturbation.



**Fig. 1** Overview of site-specific labeling strategies for long RNAs. The methods can be subdivided into internal- (top), and end-labeling (bottom) approaches and categorized as either enzymatic (dark gray) or chemical (light gray). Except for GMP priming, all modifications occur after transcription. Hybridization methods can also be performed with nucleic acid variants, e.g., peptide nucleic acids. The application and RNA structure determine the choice of the respective strategy. We present a covalent method that expands the site-specific fluorescent labeling repertoire of ribozymes (highlighted in red)

Here, we revisit a strategy for dual end-labeling of long RNAs that does not rely on any enzyme nor requires synthesis of any precursors or DNA hybridization probes. Instead, the method described uses chemical reactions which have been described for modifying carboxylic acids of proteins [24] or diols of carbohydrates [25] and were repurposed for labeling the 5'-

phosphate [26, 27] or the 3'-ribose of nucleic acids [27–30], respectively. While diol cleavage with subsequent 3'-labeling is comparatively well established, we advocate here an alternative route for 5'-end labeling which employs an *N*-hydroxysuccinimide (NHS) phosphate intermediate instead of the more commonly used phosphorimidazolide [27, 31, 32] to increase efficiency at lower temperatures (Fig. 1). We optimize reaction conditions, improve purification, and outline the quality control steps needed to assess coupling efficiency, RNA integrity and function. We apply our protocol to one of the largest known ribozymes, the wild-type group II intron Sc.ai5y of *Saccharomyces cerevisiae* [33, 34] carrying two flanking exon sequences with a total length of 915 nucleotides. Self-splicing of this intron at 42 °C in the presence of millimolar amounts of Mg<sup>2+</sup> [35] poses additional challenges to the labeling process and precludes the use of metal ion-dependent enzymes at elevated temperatures. Chemical end-labeling is thus a valid replacement for enzymatic ligation with terminal fluorescent tags and applicable to a wide variety of non-coding RNAs.

---

## 2 Materials

All buffers and solutions are prepared with deionized and filtered water and chemicals of at least *puriss. p.a.* grade. To minimize RNA degradation, all solutions are sterile-filtered (0.2 µm pore size), and experiments with sodium periodate and fluorescent dyes are performed in the dark to prevent photodegradation. For the whole procedure, basic laboratory equipment like a thermoshaker, vortexer, and centrifuges are needed.

### 2.1 Fluorescent Labeling at the 5'-End

1. 250 pmol in vitro transcribed and polyacrylamide gel electrophoresis (PAGE)-purified RNA in 100 µL water (corresponds to 2.5 µM of RNA, *see Note 1*).
2. 1.5 mg 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC-HCl) (*see Note 2*).
3. 2 mg *N*-hydroxysuccinimide (NHS) (*see Note 3*).
4. 50 mM sodium acetate (NaOAc) buffer, pH 6.0 (*see Note 4*).

5. 100 mM 3-morpholinopropane-1-sulfonic acid (MOPS) buffer, pH 7.5, stored at 4 °C, protected from light.
6. 5 µL of 2 mM sulfonated Cyanine3 amine (sCy3-amine, Lumiprobe) solution: Dissolve the sCy3-amine powder in water (*see Note 5*) and determine its concentration by UV-Vis spectroscopy.

## 2.2 Fluorescent Labeling at the 3'-End

1. 250 pmol in vitro transcribed and PAGE-purified RNA in 100 µL water (corresponds to 2.5 µM of RNA).
2. 20 mM sodium meta-periodate (NaIO<sub>4</sub>) solution in water, freshly prepared, keep protected from light.
3. 50 mM NaOAc buffer, pH 5.5.
4. 50% aqueous glycerol solution.
5. 50 mM NaOAc buffer, pH 6.0.
6. 5 µL of 2 mM Cyanine 5 hydrazide (Lumiprobe) solution: Dissolve Cy5-hydrazide powder in dimethyl sulfoxide (DMSO) (*see Note 6*) and determine its concentration by UV-Vis spectroscopy.

## 2.3 Chemical Dual End-Labeling

1. All materials used for the 5'- and 3'-end labeling methods (except for NaIO<sub>4</sub>).
2. 10 mM NaIO<sub>4</sub> solution in water, freshly prepared, keep protected from light.
3. Blocking buffer: 100 mM Tris–HCl, pH 7.5.

## 2.4 RNA Purification

1. 100% and 70% EtOH (ethanol, v/v in water), ice-cold.
2. 5 M NaCl or 3 M NaOAc, pH 5.5.
3. Vivaspin® centrifugal filtration unit (MWCO 50.000, PES, 500 µL).

## 2.5 Analytical Size-Exclusion HPLC

1. HPLC system (e.g., UltiMate 3000 equipped with a diode array detector for UV-Vis detection of the RNA (max. 260 nm), sCy3 (max. 550 nm) and Cy5 (650 nm), Thermo Fisher Scientific).
2. Size-exclusion column: Zenix SEC-300 column (4.6 × 300 mm, 300 Å, 3 µm, Sepax Technologies).
3. Mobile phase: 150 mM sodium phosphate buffer, pH 7.0, filtered and degassed (see **Note 7**).

## 2.6 Native Agarose Gel Electrophoresis

1. 0.5 µg single- and dual-labeled RNA samples.
2. 2% agarose gel with 1× tris(hydroxymethyl)aminomethane (Tris)-acetate-ethylenediaminetetraacetic acid (EDTA) (TAE) buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA).
3. 6× Gel Loading Dye, Purple.
4. Biomolecular imager (e.g., Typhoon FLA 9500 laser scanner, GE Healthcare).

## 2.7 Ensemble FRET Spectroscopy

1. 500 mM KCl solution.

2. 100 mM MgCl<sub>2</sub> solution.
  3. Fluorescence spectrophotometer (e.g., Fluorolog FL3-222 equipped with dual-grating excitation and emission monochromators and a PDD-900 photodiode detector, Horiba).
  4. Fluorescence sub-micro quartz cuvette.
- 

### 3 Methods

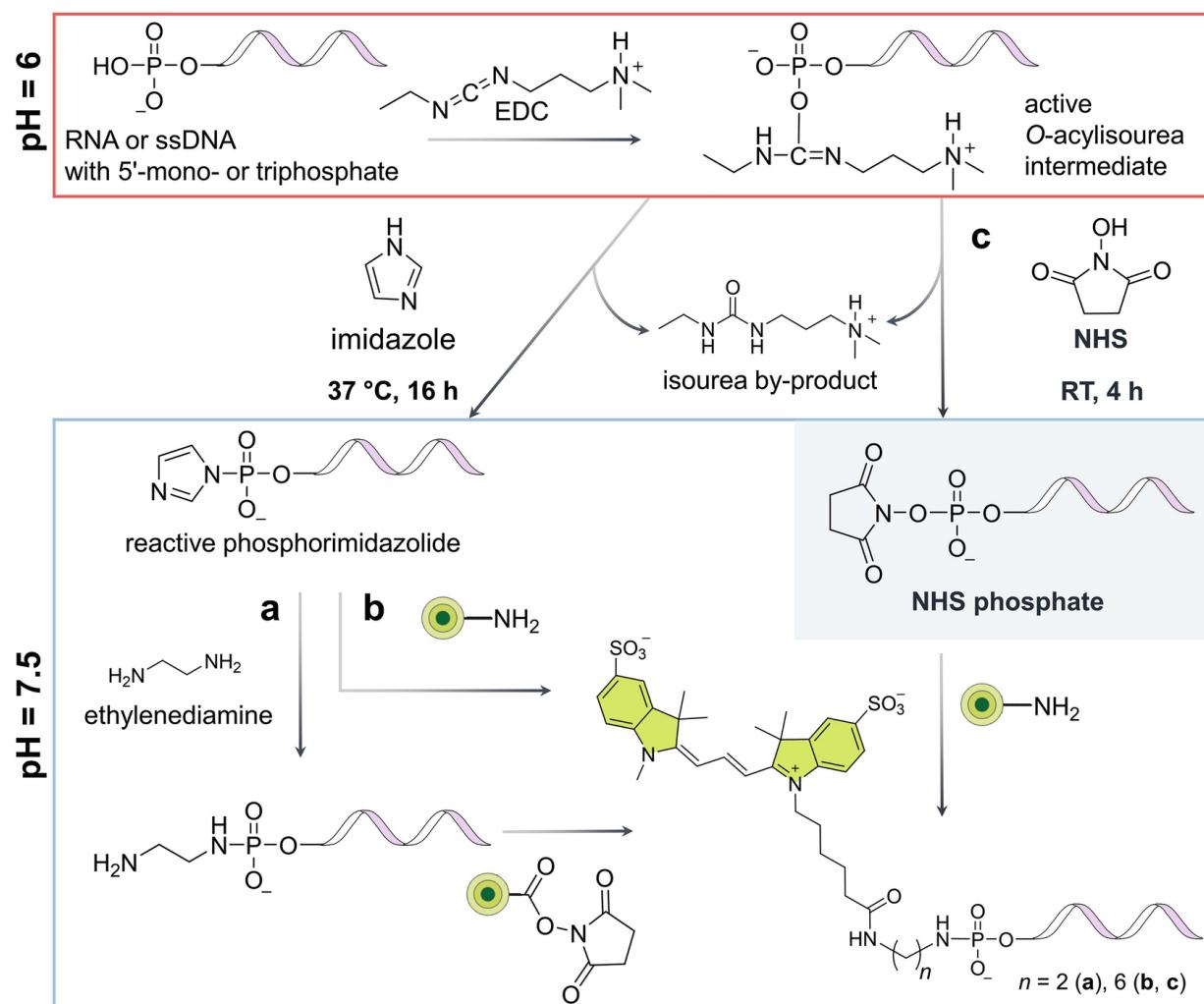
The group II intron ribozyme is in vitro transcribed by T7 RNA polymerase and purified by 5% PAGE under denaturing conditions (*see Note 8*) followed by electroelution and ethanol (EtOH) precipitation [36].

#### 3.1 EDC/NHS Activation and Dye Coupling at the RNA 5'-End

In this protocol, we adapted the 5'-end labeling protocol to use EDC/NHS activation instead of EDC/imidazole as previously described to increase efficiency and speed of the reaction (Fig. 2) (*see Note 9*). In a first step, the RNA of interest is activated at its 5'-phosphate (*see Note 10*) in the presence of EDC to an O-acylisourea. This active intermediate then reacts spontaneously with NHS, to form a NHS phosphate derivative, which can be precipitated in EtOH, allowing for buffer exchange and pH control (*see Note 11*). In a second step, an amine functionalized fluorophore of choice, here an sCy3-amine, is coupled to the activated 5'-end.

1. Freshly dissolve 1.5 mg EDC and 2 mg NHS in water.
2. Prepare a solution of 2.5 μM purified RNA in 50 mM NaOAc buffer, pH 6.0, with a final volume of 100 μL.
3. Mix the two solutions and incubate for 4 h at 25 °C and 500 rpm (*see Note 9*).
4. Pellet the activated RNA by EtOH precipitation.

5. Resuspend the RNA pellet in MOPS buffer to a final volume of 100  $\mu$ L.
6. Add a 40-fold excess (100  $\mu$ M) of sCy3-amine solution.
7. Incubate for 16 h at 25 °C and 500 rpm.
8. Perform at least two rounds of EtOH precipitation to remove the unbound dyes.
9. Resuspend in water and remove any remaining unbound dyes by filtration with an appropriate molecular weight cut-off. Elute the colored sample in water.



**Fig. 2** Performance enhancement of the 5'-end ssDNA and RNA labeling via phosphoramidation by EDC/NHS chemistry. Carbodiimides, such as EDC, activate the 5'-end phosphate of nucleic acids under acidic conditions to form amine-reactive intermediates. The phosphorimidazolide can react at higher pH (**a**) with ethylenediamine, followed by NHS-ester coupling, or (**b**) with an amine-fluorophore in a single step. The low stability of the phosphorimidazolide complicates pH control, which in turn decreases efficiency. We replaced the imidazole with (**c**) NHS. In this way, the reactive intermediate molecule, an NHS phosphate (shaded in blue), is stable to be purified from EDC and by-products before increasing the pH allowing for efficient labeling at shorter times and lower incubation temperatures

### 3.2 Periodate Oxidation and Dye Coupling at the RNA 3'-End

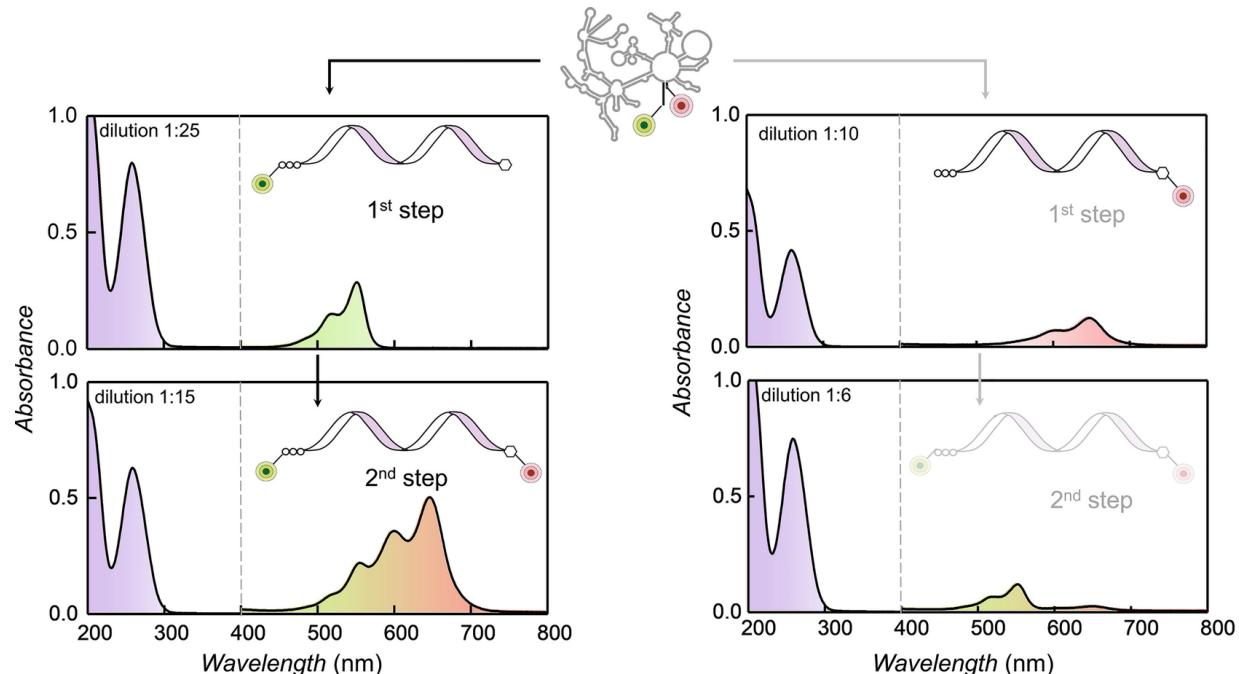
First, periodate oxidizes the two adjacent hydroxyl residues of the 3'-terminal ribose (see **Note 12**) to aldehydes. The dialdehyde is subsequently functionalized with the hydrazine derivative of a fluorophore of choice, here a Cy5-hydrazide, to form a stable hydrazone bond.

1. Prepare a solution of 2.5  $\mu$ M purified RNA with 20 mM NaIO<sub>4</sub> in 50 mM NaOAc buffer, pH 5.5 in a final volume of 100  $\mu$ L, and incubate for 2 h at 25 °C and 500 rpm.
2. Quench the excess of periodate with 30  $\mu$ L of glycerol solution (see **Note 13**) and continue the incubation for 30 min at 25 °C and 500 rpm.
3. Purify the oxidized RNA from the supernatant containing the quenched periodate by EtOH precipitation.
4. Resuspend the RNA pellet in 50 mM NaOAc, pH 6.0 to a final reaction volume of 100  $\mu$ L.
5. Add a 40-fold excess of Cy5-hydrazide dye solution (100  $\mu$ M) in DMSO and incubate the mixture for 16 h at 25 °C and 500 rpm.
6. After the completion of the reaction, avoid the use of reducing agents with cyanine dyes (see **Note 14**).
7. Stop the reaction by EtOH precipitation and proceed analogous to **steps 8 and 9** in Subheading **3.1**.

### 3.3 Chemical Dual End-Labeling

We explored both labeling orders ( $3' \rightarrow 5'$  and  $5' \rightarrow 3'$ ) and found that only the second resulted in high labeling densities of both sCy3 and Cy5 (Fig. 3). In this order, the reaction conditions of the  $5'$ - and  $3'$ -labeling are compatible with each other (see Note 15).

1. Start with labeling the  $5'$ -end (see Subheading 3.1).
2. To inactivate any unreacted *N*-succinimidyl phosphates from the  $5'$ -labeling step, incubate the RNA in Tris–HCl blocking buffer for 2 h at 25 °C and 500 rpm (see Note 16).
3. Stop the reaction by EtOH precipitation, resuspend in water, and remove any remaining free amine dyes by filtration (see Note 17).
4. Proceed with the  $3'$ -end labeling, but only use an aqueous 10 mM NaIO<sub>4</sub> solution (see Note 18).

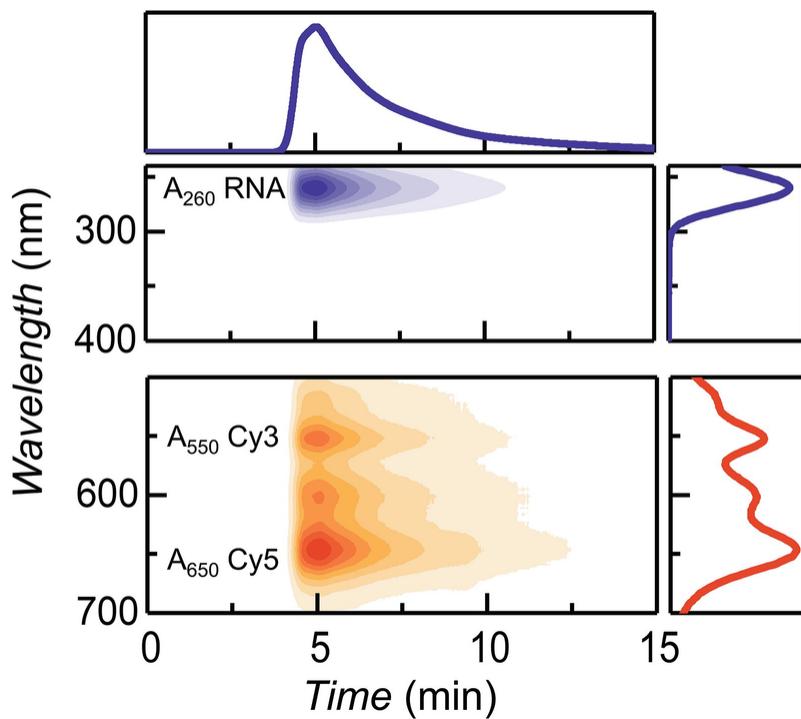


**Fig. 3** UV-Vis spectra of single and dual-labeled group II introns with sCy3 and Cy5. Each panel is composed of two spectra (separated by grey dashed line): one recorded from 200 to 400 nm with an indicated dilution factor and a second from 400 to 800 nm without

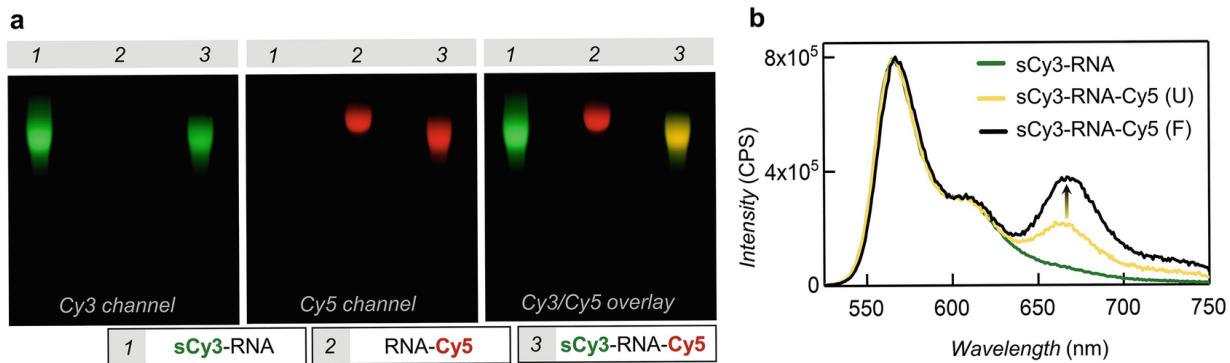
dilution. The dilution factor accounts for the very different extinction coefficients of the RNA ( $\sim 10,000 \text{ mM}^{-1} \text{ cm}^{-1}$ ) and the dyes (sCy3:  $162 \text{ mM}^{-1} \text{ cm}^{-1}$  and Cy5:  $250 \text{ mM}^{-1} \text{ cm}^{-1}$ ). Dual-labeling is much more efficient if started with the 5'-end, i.e.,  $5' \rightarrow 3'$  dual end-labeled RNA (see **Note 15**)

### 3.4 Quality Checks

1. To confirm that dyes are successfully coupled to the RNA, inject the labeled RNA sample onto an analytical size-exclusion HPLC column and elute in sodium phosphate buffer at a flow rate of 0.4 mL/min. The RNA (detected at 260 nm) and the conjugated dyes (absorption at 550/650 nm) should co-elute (Fig. 4).
2. To check the integrity of the RNA after labeling is complete, run a polyacrylamide or agarose gel. Here we load 1.5–2.0 pmol (equivalent to about 0.5  $\mu\text{g}$ ) labeled RNA onto a 2% agarose gel. Run the gel with 1 $\times$  TBE buffer in the dark. Visualize the gel on a biomolecular imager by excitation at 532 nm and 635 nm using a bandpass (BPG1, 560–580 nm) and a longpass filter (LPR, >665 nm) for the Cy3 and Cy5 channel, respectively.
3. To assess whether one RNA molecule is in fact labeled with both sCy3 and Cy5 dyes, record a fluorescence spectrum by exciting the donor dye at 515 nm (Fig. 5). If the dyes are in proximity (2–10 nm, i.e., on the same molecule) energy transfer from the donor to acceptor should lead to appearance of a FRET peak. In our case, folding of the group II intron brings the 5'- and 3'-exons closer together which is reflected in an increase of the FRET peak. In the specific case of the group II intron Sc.ai5 $\gamma$ , RNA folding is induced by incubation with 500 mM KCl at 70 °C for 3 min, followed by lowering the temperature to 42 °C for 5 min before adding 100 mM MgCl<sub>2</sub>.



**Fig. 4** Coupling of the both sCy3 and Cy5 dyes to the RNA is confirmed by analytical size-exclusion chromatography with UV-Vis detection. Co-elution of the RNA with absorption at  $A_{260}$  (purple) and the coupled sCy3 at  $A_{550}$  and Cy5  $A_{650}$  (orange) with the same retention time confirms the presence of the dye-RNA conjugate



**Fig. 5** Assessing RNA structural integrity by gel electrophoresis and ensemble FRET. **(a)** Fluorescently labeled RNA is visualized on an analytical 2% agarose gel by illumination at 532 nm and 635 nm. The absence of any degradation products suggests that RNA integrity is preserved upon labeling. Overlaying the Cy3 (green) and Cy5 (red) channel shows a co-localization (yellow) of both fluorophores in the dual-labeled sample. **(b)** Fluorescence emission spectra of the donor-only (green) and dual-labeled RNA (yellow, U unfolded) with energy transfer happening in the presence of the acceptor. The FRET peak increases upon folding (black, F folded) which confirms that both fluorophores are coupled to the RNA and highlights the strength of the labeling method in maintaining RNA structure and

function. All spectra are rescaled to the donor-only labeled sample at the Cy3 emission maximum

---

## 4 Notes

1. We recommend starting with a concentrated RNA sample, as also suggested by others [37]. The 2.5  $\mu\text{M}$  used herein are considered reasonably high for such a long RNA. High concentrations also make it easier to visually estimate the labeling density through the color of the pellet.
2. Among the many carbodiimide derivatives, EDC is used because of its high water-solubility.
3. Here we chose to work with NHS as an economic alternative to *N*-hydroxysulfosuccinimide (sulfo-NHS), which may be used if solubility is an issue.
4. Avoid the use of phosphate buffers during 5'-labeling since the buffer may compete with the terminal phosphate for EDC. After labeling is complete, phosphate buffer can be used for purification and analysis.
5. The sulfonated form of the Cy3-amine dye is preferred because of its high solubility in water, which obviates the need for organic solvents (DMSO, DMF (Dimethylformamide), or DCM (Dichloromethane)). Particularly, DMSO should be avoided in the presence of EDC because it reacts with the carbodiimide via Pfitzner–Moffatt oxidation.
6. Cy5-hydrazide is only moderately soluble in water and instead polar organic solvents (DMSO or DMF) can be used for 3'-end labeling to dissolve the fluorophore. The sulfonated form of the dye is not used in our case due to its limited commercial availability.
7. Because group II intron folding and catalysis is promoted in the presence of  $\text{K}^+$ , here  $\text{Na}^+$  is used as a cation.
8. This method keeps the ribozyme inactive during the entire process, from RNA preparation until dual-labeling is complete, to prevent the

group II intron from cleaving its exons. However, if native working conditions are desired, refer to adjustments in [38].

9. Using NHS reduces incubation times and temperatures compared to imidazole. Reaction via phosphorimidazolide on the other hand works best at 37 °C with incubation times up to 16 h (other tested temperatures were 25 °C, 30 °C, and 50 °C).
10. Transcribed RNAs are known to possess 5'-ends with a mixture of mono-, di-, and triphosphates. The labeling protocol itself is not affected by this heterogeneity. However, if clean ends are desired, the RNA may be dephosphorylated using alkaline phosphatase or calf intestinal phosphatase and subsequently re-mono-phosphorylated with kinase and adenosine triphosphate (ATP). The conditions used for such Mg<sup>2+</sup>-dependent enzymes can lead to simultaneous intron splicing. New classes of Mg<sup>2+</sup>-independent enzymes such as RNA 5'-pyrophosphohydrolase can convert 5'-triphosphates directly to 5'-monophosphates. However, the quantity of enzymes needed for such treatments is not economical for micromolar scale RNA labeling.
11. pH control of the EDC activation is crucial. At pH > 8 carbodiimides react with surface-accessible guanines and uracils. This reactivity has been exploited for in-line probing [39, 40]. Prior to amination at higher pH, EDC and the isourea by-product have to be removed.
12. In classical run-off in vitro transcription, as performed here, the T7 RNA polymerase is known to add one to three unspecific nucleotides at the 3'-end [36, 41, 42]. For short RNAs, the difference in size is visible on a denaturing PAGE and the correct transcript can be selected. However, for large RNAs, such size differences are not distinguishable by gel electrophoresis, resulting in a heterogeneous RNA sample. One way to achieve homogeneous 3'-ends is to add flanking *cis*-acting small ribozyme into the DNA template, e.g., the hammerhead or hairpin ribozyme [36, 41, 42]. For active group II intron constructs, this plasmid design is not possible, as the ionic conditions and temperature used for cleavage also induce intron splicing.

13. The reaction is stopped by quenching the excess of periodate with a diol. Glycerol is preferred to ethylene glycol or 2-mercaptopropanol due to lower toxicity.
14. Several protocols describing aldehyde-hydrazide coupling reactions end with a final reduction step by sodium borohydride ( $\text{NaBH}_4$ ) to increase product stability [30, 31]. Here we refrain from using  $\text{NaBH}_4$  to avoid reducing the dyes to non-fluorescent hydrocyanines, which have gained attention in sensing reactive oxygen species and in super-resolution imaging [43, 44].
15. The order  $5' \rightarrow 3'$  is enforced because any remaining free aldehydes on the 3'-terminal ribose can neither be reduced to alcohols (see Note 14) nor oxidized to carboxylic acids (reactive towards EDC), nor treated with methylhydrazine or *N,N*-dimethylhydrazine (because of toxicity and explosivity).
16. Tris (or any other buffer containing primary amines, such as glycine, or ammonium) should be avoided during the previous steps since it competes with the amine fluorophore, and is added only at the end to quench any unreacted *N*-succinimidyl phosphates.
17. This intermediate filtration step is crucial as remaining amine dyes might cross-react with the 3'-aldehydes via a Schiff's base reaction forming an imine. Although amine-aldehyde coupling is reversible, the reaction efficiency of the hydrazide could be diminished.
18. Oxidation may bleach the sCy3 dye attached to the RNA 5'-end. The periodate concentration in dual-labeling is therefore decreased by a factor of two. By the end, a purple-colored pellet indicates labeling with both sCy3 and Cy5 dyes.

## Acknowledgements

Financial support from the Swiss National Science Foundation [200020\_192153 to RKOS], the UZH Forschungskredit [FK-20-081 to

EA], the University of Zurich, and the Graduate School of Chemical and Molecular Sciences Zurich (CMSZH) is gratefully acknowledged.

---

## References

1. Panja S, Hua B, Zegarra D et al (2017) Metals induce transient folding and activation of the twister ribozyme. *Nat Chem Biol* 13(10):1109–1114. <https://doi.org/10.1038/nchembio.2459>
2. Borgia A, Borgia MB, Bugge K et al (2018) Extreme disorder in an ultrahigh-affinity protein complex. *Nature* 555(7694):61–66. <https://doi.org/10.1038/nature25762>
3. Suddala KC, Price IR, Dandpat SS et al (2019) Local-to-global signal transduction at the core of a Mn<sup>2+</sup> sensing riboswitch. *Nat Commun* 10(1):4304. <https://doi.org/10.1038/s41467-019-12230-5>
4. Steffen FD, Börner R, Freisinger E et al (2019) Stick, Flick, click: DNA-guided fluorescent labeling of long RNA for single-molecule FRET. *Chimia* 73(4):257–261. <https://doi.org/10.2533/chimia.2019.257>
5. Hanspach G, Trucks S, Hengesbach M (2019) Strategic labelling approaches for RNA single-molecule spectroscopy. *RNA Biol* 16(9):1119–1132. <https://doi.org/10.1080/15476286.2019.1593093>
6. Solomatin S, Herschlag D (2009) Methods of site-specific labeling of RNA with fluorescent dyes. In: Herschlag D (ed) Biophysical, chemical, and functional probes of RNA structure approaches to RNA structure, interactions and folding, vol 469, 1st edn. Elsevier, San Diego, pp 47–68.
7. Liu Y, Sousa R, Wang Y-X (2016) Specific labeling: an effective tool to explore the RNA world. *BioEssays* 38(2):192–200. <https://doi.org/10.1002/bies.201500119>
8. Paredes E, Evans M, Das SR (2011) RNA labeling, conjugation and ligation. *Methods* 54(2):251–259. [https://doi.org/10.1016/j.ymeth.2011.02.008](https://doi.org/10.1016/jymeth.2011.02.008)
9. Klöcker N, Weissenboeck FP, Rentmeister A (2020) Covalent labeling of nucleic acids. *Chem Soc Rev* 49(23):8749–8773. <https://doi.org/10.1039/d0cs00600a>
10. Smith GJ, Sosnick TR, Scherer NF et al (2005) Efficient fluorescence labeling of a large RNA through oligonucleotide hybridization. *RNA* 11(2):234–239. <https://doi.org/10.1261/rna.7180305>
11. Schmitz AG, Zelger-Paulus S, Gasser G et al (2015) Strategy for internal labeling of

- large RNAs with minimal perturbation by using fluorescent PNA. *Chembiochem* 16(9):1302–1306. <https://doi.org/10.1002/cbic.201500180>
12. Egloff D, Oleinich IA, Zhao M et al (2016) Sequence-specific post-synthetic oligonucleotide labeling for single-molecule fluorescence applications. *ACS Chem Biol* 11(9):2558–2567. <https://doi.org/10.1021/acscchembio.6b00343>
  13. Zhao M, Steffen FD, Börner R et al (2018) Site-specific dual-color labeling of long RNAs for single-molecule spectroscopy. *Nucleic Acids Res* 46(3):e13. <https://doi.org/10.1093/nar/gkx1100>
  14. Jahn K, Olsen EM, Nielsen MM et al (2011) Site-specific chemical labeling of long RNA molecules. *Bioconjug Chem* 22(1):95–100. <https://doi.org/10.1021/bc100422k>
  15. Onizuka K, Taniguchi Y, Sasaki S (2009) Site-specific covalent modification of RNA guided by functionality-transfer oligodeoxynucleotides. *Bioconjug Chem* 20(4):799–803. <https://doi.org/10.1021/bc900009p>
  16. Moore MJ, Sharp PA (1992) Site-specific modification of pre-mRNA: the 2'-hydroxyl groups at the splice sites. *Science* 256(5059):992–997.
  17. Baum DA, Silverman SK (2007) Deoxyribozyme-catalyzed labeling of RNA. *Angew Chem Int Ed* 46(19):3502–3504. <https://doi.org/10.1002/anie.200700357>
  18. Lang K, Micura R (2008) The preparation of site-specifically modified riboswitch domains as an example for enzymatic ligation of chemically synthesized RNA fragments. *Nat Protoc* 3(9):1457–1466. <https://doi.org/10.1038/nprot.2008.135>
  19. Ghaem Maghami M, Scheitl CPM, Höbartner C (2019) Direct in vitro selection of trans-acting ribozymes for posttranscriptional, site-specific, and covalent fluorescent labeling of RNA. *J Am Chem Soc* 141(50):19546–19549. <https://doi.org/10.1021/jacs.9b10531>
  20. Zearfoss NR, Ryder SP (2012) End-labeling oligonucleotides with chemical tags after synthesis. In: Conn GL (ed) Recombinant and in vitro RNA synthesis: methods and protocols, vol 941. Humana Press, New York, pp 181–193.
  21. Papastavrou N, Bande O, Marlière P et al (2020) Vitamin-guanosine monophosphate conjugates for in vitro transcription priming. *Chem Commun* 56(18):2787–2790. <https://doi.org/10.1039/c9cc09427j>
  22. Warminski M, Sikorski PJ, Kowalska J et al (2017) Applications of phosphate modification and labeling to study (m)RNA caps. *Top Curr Chem* 375(1):16. <https://doi.org/10.1007/s41061-017-0106-y>
  - 23.

Westerich KJ, Chandrasekaran KS, Gross-Thebing T et al (2020) Bioorthogonal mRNA labeling at the poly(A) tail for imaging localization and dynamics in live zebrafish embryos. *Chem Sci* 11(11):3089–3095. <https://doi.org/10.1039/C9SC05981D>

24. Hoare DG, Koshland DE (1966) A procedure for the selective modification of carboxyl groups in proteins. *J Am Chem Soc* 88(9):2057–2058. <https://doi.org/10.1021/ja00961a045>
25. Bobbitt JM (1956) Periodate oxidation of carbohydrates. In: Wolfrom ML, Tipson RS (eds) *Advances in carbohydrate chemistry*, vol 11. Academic Press, Cambridge, pp 1–41.
26. Chu BC, Wahl GM, Orgel LE (1983) Derivatization of unprotected polynucleotides. *Nucleic Acids Res* 11(18):6513–6529. <https://doi.org/10.1093/nar/11.18.6513>
27. Qin PZ, Pyle AM (1999) Site-specific labeling of RNA with fluorophores and other structural probes. *Methods* 18(1):60–70. <https://doi.org/10.1006/meth.1999.0757>
28. Whitfeld PR (1954) A method for the determination of nucleotide sequence in polyribonucleotides. *Biochem J* 58(3):390–396. <https://doi.org/10.1042/bj0580390>
29. Zamecnik PC, Stephenson ML, Scott JF (1960) Partial purification of soluble RNA. *Proc Natl Acad Sci U S A* 46(6):811–822. <https://doi.org/10.1073/pnas.46.6.811>
30. Proudnikov D, Mirzabekov A (1996) Chemical methods of DNA and RNA fluorescent labeling. *Nucleic Acids Res* 24(22):4535–4542. <https://doi.org/10.1093/nar/24.22.4535>
31. Hermanson GT (2013) *Bioconjugate techniques*, 3rd edn. Elsevier, London.
32. Wickramathilaka MP, Tao BY (2019) Characterization of covalent crosslinking strategies for synthesizing DNA-based bioconjugates. *J Biol Eng* 13:63. <https://doi.org/10.1186/s13036-019-0191-2>
33. Pyle AM (2016) Group II intron self-splicing. *Annu Rev Biophys* 45:183–205. <https://doi.org/10.1146/annurev-biophys-062215-011149>
34. Peebles CL, Perlman PS, Mecklenburg KL et al (1986) A self-splicing RNA excises an intron lariat. *Cell* 44(2):213–223. [https://doi.org/10.1016/0092-8674\(86\)90755-5](https://doi.org/10.1016/0092-8674(86)90755-5)
35. Su LJ, Brenowitz M, Pyle AM (2003) An alternative route for the folding of large RNAs: apparent two-state folding by a group II intron ribozyme. *J Mol Biol* 334(4):639–652. <https://doi.org/10.1016/j.jmb.2003.09.071>
36. Gallo S, Furler M, Sigel RKO (2005) In vitro transcription and purification of RNAs

of different size. Chimia 59(11):812–816. <https://doi.org/10.2533/000942905777675589>

37. Greenfeld M, Herschlag D (2013) Fluorescently labeling synthetic RNAs. In: Lorsch J (ed) Laboratory methods in enzymology: RNA, vol 530, 1st edn. Elsevier, Amsterdam, pp 281–297.
38. Awwad DA, Rahmouni AR, Aboul-ela F (2020) Protocol for efficient fluorescence 3' end-labeling of native noncoding RNA domains. MethodsX 7:101148. <https://doi.org/10.1016/j.mex.2020.101148>
39. Wang PY, Sexton AN, Culligan WJ et al (2019) Carbodiimide reagents for the chemical probing of RNA structure in cells. RNA 25(1):135–146. <https://doi.org/10.1261/rna.067561.118>
40. Mitchell D, Renda AJ, Douds CA et al (2019) In vivo RNA structural probing of uracil and guanine base-pairing by 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC). RNA 25(1):147–157. <https://doi.org/10.1261/rna.067868.118>
41. Price SR, Ito N, Oubridge C et al (1995) Crystallization of RNA-protein complexes. I. Methods for the large-scale preparation of RNA suitable for crystallographic studies. J Mol Biol 249(2):398–408. <https://doi.org/10.1006/jmbi.1995.0305>
42. Ferré-D'Amaré AR, Doudna JA (1996) Use of cis- and trans-ribozymes to remove 5' and 3' heterogeneities from milligrams of in vitro transcribed RNA. Nucleic Acids Res 24(5):977–978. <https://doi.org/10.1093/nar/24.5.977>
43. Vaughan JC, Jia S, Zhuang X (2012) Ultrabright photoactivatable fluorophores created by reductive caging. Nat Methods 9(12):1181–1184. <https://doi.org/10.1038/nmeth.2214>
44. Sadlowski CM, Maity S, Kundu K et al (2017) Hydrocyanines: a versatile family of probes for imaging radical oxidants in vitro and in vivo. Mol Syst Des Eng 2(3):191–200. <https://doi.org/10.1039/C7ME00014F>

## 14. Spin Labeling of Long RNAs Via Click Reaction and Enzymatic Ligation

Maria Francesca Vicino<sup>1</sup>, Christine Wuebben<sup>1</sup>, Mark Kerzhner<sup>2</sup>,  
Michael Famulok<sup>2</sup> and Olav Schiemann<sup>1</sup>✉

- (1) Institute of Physical and Theoretical Chemistry, Rheinische Friedrich Wilhelms University, Bonn, Germany  
(2) Life & Medical Sciences Institute (LIMES), Chemische Biologie, c/o Kekulé-Institut für organische Chemie, Bonn, Germany

✉ **Olav Schiemann**  
Email: [schiemann@pc.uni-bonn.de](mailto:schiemann@pc.uni-bonn.de)

### Abstract

Electron paramagnetic resonance (EPR) is a spectroscopic method for investigating structures, conformational changes, and dynamics of biomacromolecules, for example, oligonucleotides. In order to be applicable, the oligonucleotide has to be labeled site-specifically with paramagnetic tags, the so-called spin labels. Here, we provide a protocol for spin labeling of long oligonucleotides with nitroxides. In the first step, a short and commercially available RNA strand is labeled with a nitroxide via a copper-(I)-catalyzed azide–alkyne cycloaddition (CuAAC), also referred to as “click” reaction. In the second step, the labeled RNA strand is fused to another RNA sequence by means of enzymatic ligation to obtain the labeled full-length construct. The protocol is robust and has been shown experimentally to deliver high yields for RNA sequences up to 81 nucleotides, but longer strands are in principle also feasible. Moreover, it sets the path to label, for example, long riboswitches, ribozymes, and

DNAzymes for coarse-grained structure determination and enables to investigate mechanistical features of these systems.

**Key words** Azide–alkyne cycloaddition – Click chemistry – Enzymatic ligation of RNA – Electron paramagnetic resonance (EPR) spectroscopy – Electron spin resonance (ESR) spectroscopy – In vitro RNA labeling – Spin-labeled RNA – Spin labeling

---

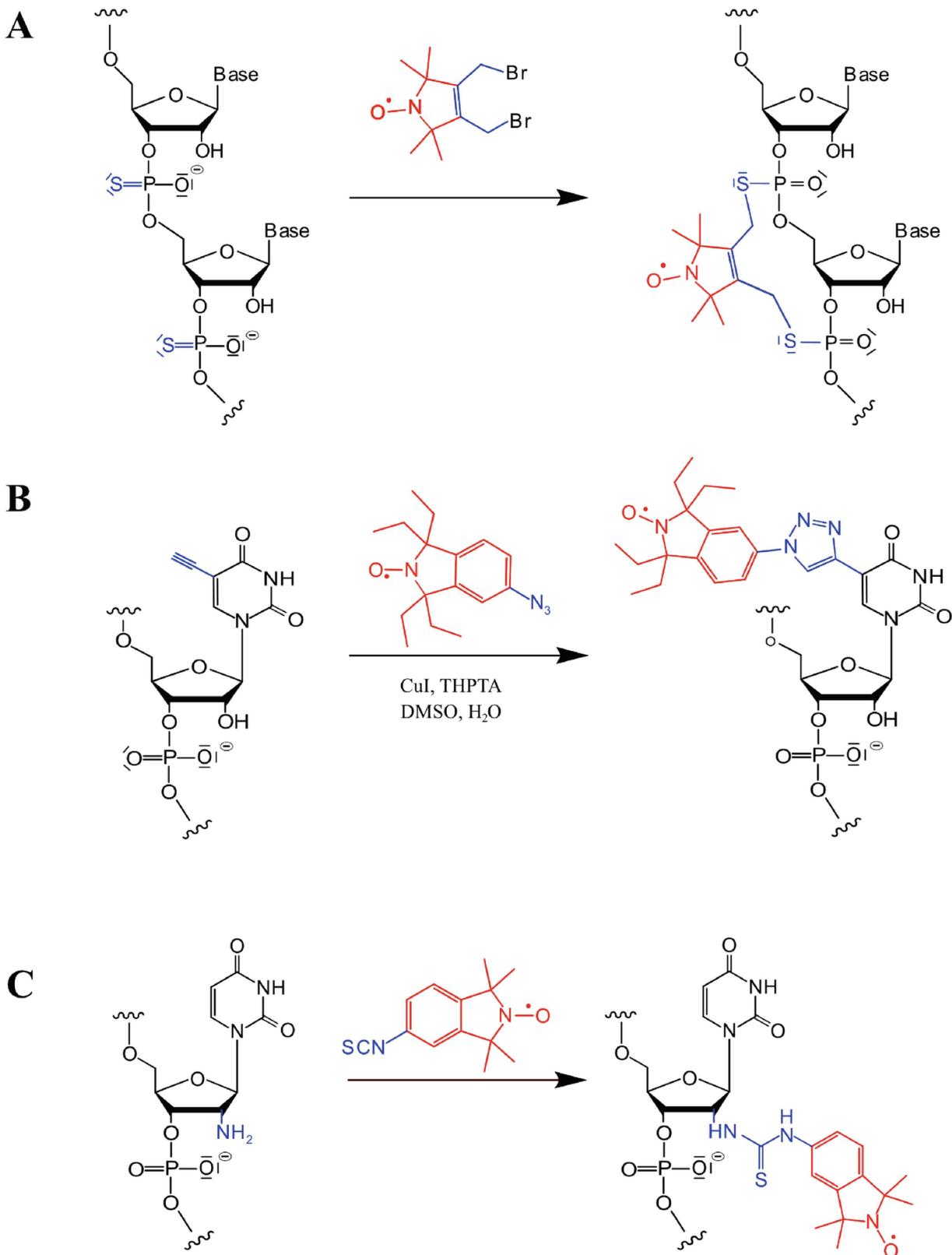
## 1 Introduction

Like proteins, the function of oligonucleotides is governed not only by the primary sequence but also by their secondary and tertiary structures. For example, it is the three-dimensional fold that leads to the formation of binding pockets for small molecules or ions, provides interaction sites for proteins or other RNA stands, and even cleavage sites for other RNA strands [1]. Often, this substrate binding or interaction is accompanied by changes in the fold of the RNA. Riboswitches for example regulate gene expression on a transcriptional and/or translational level via ligand binding–induced conformational changes [1, 2]. Ribozymes, like the hammerhead ribozymes catalyze not only phosphodiester bond cleavage but also phosphodiester bond ligation after conformational changes [3–6].

DNAzymes cleave and ligate RNA and can even be used to modify proteins [7–9]. Resolving the initial and final structures, monitoring conformational changes and internal dynamics as well as observing structural heterogeneity is vital for an understanding of the underlying principles of oligonucleotide structure–function relationships. Thus, biophysical methods are needed that can provide such information on the relevant length and time scale.

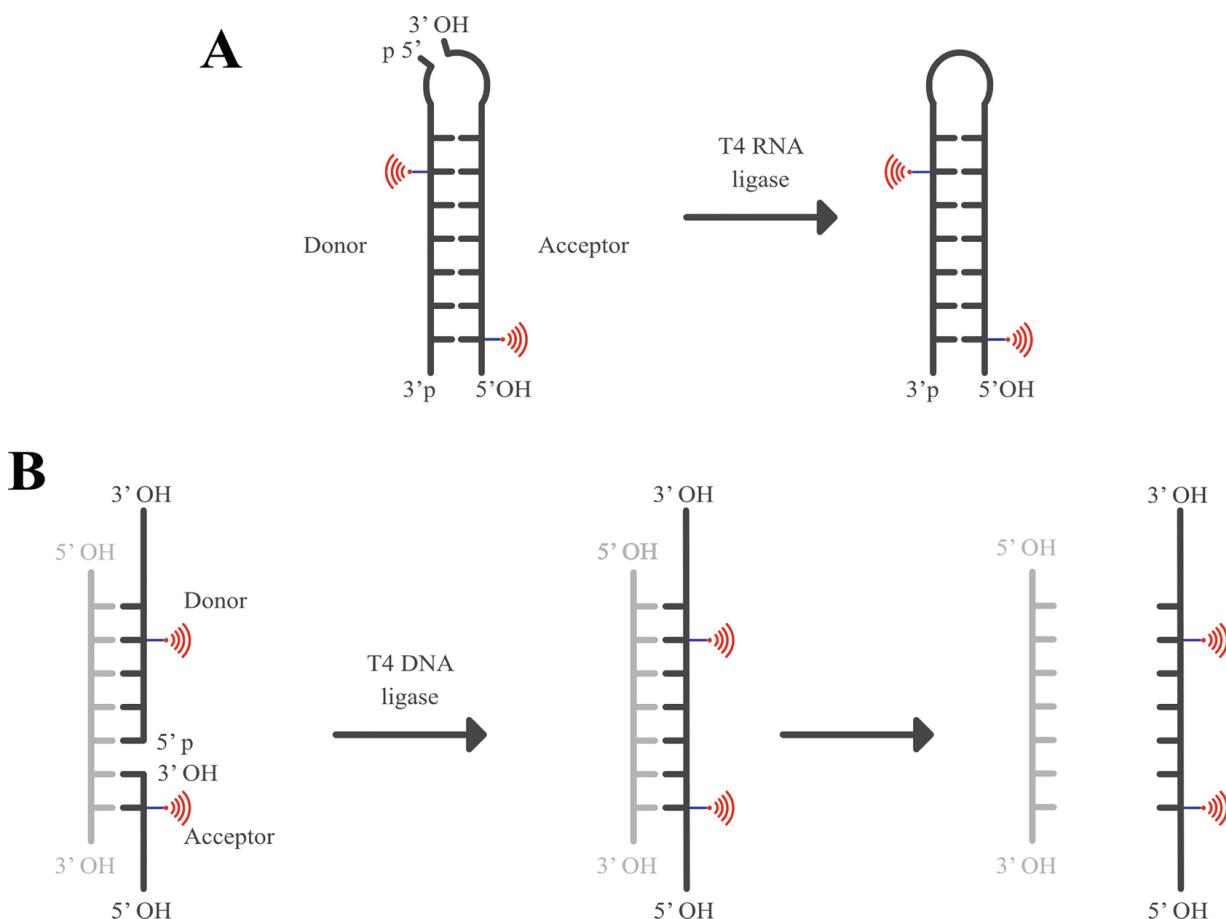
Electron paramagnetic resonance (EPR) spectroscopy [10, 11] belongs to the family of high precision biophysical methods providing structural information from the angstrom to nanometer scale with a time resolution ranging from nanoseconds to hours [12–15]. EPR has no size limit regarding the biomolecule [16–18], can be performed in liquid [19, 20] or frozen solutions [16–18, 21–23], within membranes [24, 25], cells [26], or crystals [27] and down to nanomolar concentrations [26]. However, EPR spectroscopic methods require that the biomolecule contains unpaired electrons. Since biomolecules are usually diamagnetic, its application does require labeling with unpaired electrons. This can either be achieved by

means of a paramagnetic for diamagnetic metal ion exchange, for example,  $Mg^{2+}$  for  $Mn^{2+}$  [28–30] or by site-directed spin labeling with stable organic radicals [26, 31–34] or paramagnetic metal complexes [35, 36]. In the case of oligonucleotides, the general approach is to modify the phosphate backbone [37, 38], the base [39–42] or the sugar [43, 44] with a unique functional group and to specifically label this site with a spin label bearing the corresponding functional group. For selected examples see Fig. 1. The labeling methods mentioned above require a unique functional group that can be introduced during solid phase synthesis of RNA. This limits the length of the oligonucleotide to less than 50 nucleotides (nt) [45]. Ways to overcome this limitation are: (1) synthesis of a short, labeled strand that binds to the complementary sequence in the long strand [46], (2) genetic alphabet extension transcription [47], and (3) enzymatic ligation of a short, labeled strand to the remaining sequence [22, 23]. The ligation approach can be divided into two main strategies: one makes use of T4 RNA and the other of T4 DNA ligase, schematically shown in Fig. 2. Both are adenosine triphosphate (ATP) dependent enzymes, which ligate the oligonucleotide via phosphodiester linkage. The main difference being that T4 RNA ligase has a preference for single-stranded RNA, whereas T4 DNA ligase is restricted to the ligation of double-stranded oligonucleotides, which are generated in vitro through annealing of a splint to the two oligonucleotide strands [48, 49].



**Fig. 1** Different nitroxide spin labeling approaches for RNA. The spin label is highlighted in red and the linker in blue. **(A)** Modification of the phosphate backbone is introduced

during solid-phase RNA synthesis in two consecutive nucleotides in a nucleotide independent manner. The labeling is performed after RNA synthesis, yielding a very rigid spin label attachment [37]. (B) Click reaction used in this protocol for RNA labeling. The spin labeling at the base is performed here by introduction of 5-ethynyl-2'-dU into RNA, which reacts with an azide functionalized spin label catalyzed by Cu<sup>+</sup> in a cycloaddition. Cu<sup>+</sup> forms a complex with tris(1-hydroxy-propyl-1*H*-1,2,3-triazol-4-yl(methyl)amine) (THPTA) and the reaction takes place in an aqueous solution [40]. (C) RNA with a 2'-amino group at the sugar reacts with a spin label, carrying an isothiocyanate moiety, yielding a stable thiourea linkage [43]



**Fig. 2** Comparison between ligation strategies utilizing (A) T4 RNA ligase and (B) T4 DNA ligase. Blue sticks denote the linker, as also in Fig. 1 and the label is schematically shown in red. (A) For ligation with T4 RNA ligase, the donor strand requires a 5'-phosphorylated end and the acceptor a 3'-hydroxy moiety. (B) Ligation with T4 DNA ligase requires double stranded templates. This is achieved by formation of an active complex between the acceptor, the donor strand and the splint (colored in light gray). This complex needs to be formed without any bulges or gaps. The donor needs to be 5'-phosphorylated and the acceptor requires a 3'-hydroxy moiety. This scheme is based on [48]

Here we present in the first step a labeling protocol for short oligonucleotides based on “click”-chemistry (Fig. 1b). This reaction provides minor by-products, produces high yields, is easy to perform in aqueous solution, and purification of the labeled RNA by means of high-performance liquid chromatography (HPLC) is straightforward [50, 51]. In the second step we outline the T4 DNA ligase-based ligation for obtaining longer RNAs. In contrast to the ligation with T4 DNA ligase, the ligation using T4 RNA ligase strongly depends on the position of the ligation site in the three-dimensional RNA structure. The RNA strands need to be folded in a way that the to-be ligated ends lie close to each other, for example through a hairpin loop (see Fig. 2a). In the T4 DNA ligase-based ligation this is ensured by the splint, which bridges both ends (see Fig. 2b), making this approach easier [47]. The protocol was established for RNA but also works for DNA strands and is based on our recent publications [40, 45, 52, 53].

---

## 2 Materials

**Very important:** Before you begin to work with RNA, make sure to set up an RNase-free environment. Exclusively use autoclaved pipette tips and bottles, clean your bench before starting, and wear gloves! Except where noted, use purified deionized water, filtered through a Milli-Q® Direct water purification system and autoclaved afterward. For HPLC purification, use HPLC-grade solvents.

### 2.1 RNA Labeling

1. 70% ethanol (technical) or an RNase decontamination solution.
2. 2.5 nmol dried RNA with one or more 5-ethynyl-2'-deoxyuridine modifications as labeling positions. To obtain it, dissolve 2.5 nmol of RNA in water in a 1.5 mL reaction tube and dry the RNA via lyophilization or vacuum concentrator, operated without heating (see **Note 1**).
3. Vacuum concentrator or freeze dryer.
4. Thermomixer.

5. Diethylpyrocarbonate (DEPC, ≥97%)-treated water; 0.1–0.2 mL DEPC are added to 100 mL MilliQ water. The solution is shaken and incubated overnight in a fume hood. Its bottle cap should not be fully closed to ensure that CO<sub>2</sub> can escape. The solution is autoclaved the next day in order to remove remaining DEPC.
6. Copper(I)iodide (≥98%).
7. 250 mM tris((1-hydroxy-propyl-1*H*-1,2,3-triazol-4-yl)methyl)amine (THPTA, 95%) in dimethyl sulfoxide (DMSO). Vortex the solution prior use.
8. DMSO (≥99.5%, stored at ambient conditions).
9. 100 mM spin label bearing an azide moiety (we use the 1,1,3,3-tetraethyl-5-azidoisoindolin-2-oxyl spin label in DMSO, which we synthesized ourselves, *see* Fig. 1b) [52]. Vortex the solution prior use.
10. Table-top centrifuge.
11. Ultrafiltration centrifugal device with a nominal molecular weight limit (NMWL) suited for your sample (we usually use a filter with 3.000 NMWL).
12. Spectrophotometer for determination of the RNA concentration.

## 2.2 Reverse Phase HPLC Purification

1. High-performance liquid chromatography (HPLC) purification system.
2. Reverse phase column, heatable (e.g., Zorbax 300SB-C18—4.6 × 150 mm for RNA sequences below 50 nt or Zorbax 300SB-C18—9.4 × 250 mm for longer RNA sequences).
3. Hamilton syringe (50 μL).

4. Buffer A: acetonitrile (HPLC grade).
5. Buffer B: 0.1 M triethylammonium acetate (TEAA, 1 M stock solution pH 7), dilute 100 mL 1 M stock solution with 900 mL MilliQ water.
6. Vacuum concentrator or freeze dryer.
7. Tabletop centrifuge.
8. Ultrafiltration centrifugal device with a suited NMWL (we usually use a filter with 3000 NMWL) or a disposable DNA purification column (e.g., NAP<sup>TM</sup>-25 gravity column, GE Healthcare).
9. Spectrophotometer for determination of RNA concentrations.
10. Optional: Liquid chromatography–mass spectrometry (LC-MS) in combination with an HPLC system using a Zorbax Narrow Bore SB C18 (2.1 × 50 mm, 5 µm) column (Agilent Technologies).
11. Buffer C (for LC-MS): 10 mM triethylamine (1 M stock solution pH 7, LC-MS grade), dilute 10 mL stock solution with 990 mL MilliQ water.
12. Buffer D (for LC-MS): 0.1 M hexafluoroisopropanol (LC-MS grade), dilute 100 mL hexafluoroisopropanol with 900 mL MilliQ water.
13. Optional: continuous wave (cw-) EPR spectrometer (e.g., EMX nano, Bruker), for spin count/labeling efficiency determination.

## 2.3 Enzymatic Ligation

1. 70% ethanol (99%, technical) or an RNase decontamination solution.
2. Thermomixer.
3. DEPC water, *see* Subheading 2.1, item 5 for preparation.

4. RNA acceptor strand bearing a 3'- and 5'-hydroxyl moiety, which may be spin-labeled or not.
5. 5'-phosphorylated RNA donor strand bearing a 3'-hydroxyl moiety, which may be spin-labeled or not.
6. Splint (*see Note 2*).
7. T4 DNA ligase and T4 DNA ligase buffer.
8. Polyethylene glycol, PEG 4000 (100%).
9. Tris–boric acid–disodium ethylenediaminetetraacetate dihydrate (TBE) buffer: dissolve 108 g Tris ( $\geq 99.9\%$ , p.a.), 55 g boric acid (99.999%), and 7.4 g disodium ethylenediaminetetraacetate dihydrate ( $\geq 99\%$ ) in 1 L MilliQ water.
10. 4% Agarose gel: mix 0.34 g agarose and 0.83 g Synergel and wash it with 50 mL 100% ethanol absolute ( $\geq 99.5\%$ , *see Note 3*). Wait 5 min and remove all ethanol. Dissolve the mixture in 50 mL TBE buffer by heating the mixture up until boiling. When everything is dissolved, use 25 mL warm ( $30\text{--}40\text{ }^{\circ}\text{C}$ ) agarose gel mixture, mix with 2.5  $\mu\text{L}$  1% ethidium bromide solution in the fume hood, and pour the gel.
11. Low-range ssRNA ladder or another RNA ladder suited for your samples.
12. 2 $\times$  RNA loading buffer.
13. Tabletop centrifuge.
14. Phenol ( $\geq 99.5\%$ ).
15. Chloroform ( $\geq 99.5\%$ ). Instead of phenol and chloroform, you can also use the ROTI phenol–chloroform–isoamyl alcohol ready-to-use mix from Carl Roth GmbH.
16. 5 M sodium acetate (NaOAc,  $\geq 98.5\%$ ) previously cooled on ice, (dissolve 10.2 g NaOAc in 25 mL water, adjust the pH to 5.2 and

filter the solution through a 0.2 µm filter device.

17. Ice-cold ethanol ( $\geq$ 99.5%, absolute).
18. Ice-cold 70% ethanol absolute, dissolve 35 mL pure ethanol absolute with 15 mL water.

## 2.4 Anion-Exchange HPLC

1. High-performance liquid chromatography (HPLC) purification system.
2. Anion-exchange column (e.g., Dionex DNAPac PA-100 (9 × 250 mm)).
3. Buffer E: dissolve 360 g (6 mol) urea ( $\geq$ 99.5%) and 100 mL 250 mM Tris–HCl ( $\geq$ 99%) buffer, pH 8 in 1 L water. Filter the solution through a cellulose acetate filter (0.2 µm pore size).
4. Buffer F: dissolve 360 g urea ( $\geq$ 99.5%), 70.2 g (500 mmol) sodium perchlorate monohydrate ( $\geq$  99%) and 100 mL of 250 mM Tris–HCl ( $\geq$ 99%) buffer, pH 8 in 1 L MilliQ water. Filter the solution through a cellulose acetate filter (0.2 µm pore size).
5. Storage Buffer: dissolve 42.8 g (0.8 M) ammonium chloride ( $\geq$ 99.7%), 3.1 g (25 mM) Tris–HCl ( $\geq$ 99%), and 100 mL (10%) acetonitrile in 900 mL MilliQ water.

---

## 3 Methods

### 3.1 RNA Labeling

1. Clean your work bench and tools with 70% ethanol or an RNase decontamination solution.
2. Preheat the thermomixer to 60 °C.
3. Resuspend one or two aliquots of dried RNA in 4.4 µL DEPC treated

4. water in an 1.5 mL reaction tube (scale up *see Note 1*).
- Weigh in 1–2 mg CuI ( $X$ ) in a 1.5 mL reaction tube (*see Note 4*). Calculate the volume of DMSO ( $Y$ ) needed to obtain a 50 mM  $\text{Cu}^+$  solution:
- $$\frac{X \text{ [mg]} \times 100}{0.952 \left[ \frac{\text{g}}{\text{L}} \right]} = Y \text{ [\mu L]}$$
- Dissolve CuI by adding  $Y$   $\mu\text{L}$  DMSO with a pipette, to obtain a yellow solution.
5. Now, **immediately**, do the complexation of  $\text{Cu}^+$  with THPTA by pipetting into another 1.5 mL reaction tube: 20  $\mu\text{L}$  DMSO, 8  $\mu\text{L}$  of the 50 mM  $\text{Cu}^+$  solution, and 3.6  $\mu\text{L}$  of a 250 mM THPTA solution. Incubate for 5 min at room temperature (*see Note 5*).
6. Prepare the labeling reaction by pipetting 2  $\mu\text{L}$  of the spin label stock solution (100 mM in DMSO) and 3.6  $\mu\text{L}$  of the  $\text{Cu}^+$ -THPTA-solution from **step 5** into the resuspended RNA. Incubate the mixture for 30 min at 300 rpm and 60 °C in the thermomixer (*see Note 6*).
7. Add 480  $\mu\text{L}$  water and transfer into an ultrafiltration centrifugal device for purification.
8. After loading the sample onto the device, spin for 30 min at 14,000  $\times g$  at room temperature. You should be left with 50–100  $\mu\text{L}$  RNA solution.
9. Pipet 500  $\mu\text{L}$  water onto the ultrafiltration centrifugal device and spin again for 30 min at 14,000  $\times g$  at room temperature.
10. Repeat **step 9** a second time. The volume left should be around 50–100  $\mu\text{L}$ .
11. To recover the RNA, turn the filter device around and put it upside-down into a new ultrafiltration centrifugal device. Spin 2 min at 1000 to 2000  $\times g$  at room temperature.
12. Wash the ultrafiltration centrifugal device with 50  $\mu\text{L}$  water. Take care

to wash each of the two filters by pipetting up and down without damaging them. Put the filter device upside-down into the collector tube from **step 11** and spin again for 2 min at 1000 to 2000  $\times g$  at room temperature.

13. Repeat **step 12** two times. Afterward, the sample volume should be around 200  $\mu\text{L}$ .
- 14.

Use the spectrophotometer to determine the absorption of the RNA solution at 260 nm ( $A(260)$ ) and to calculate the amount of RNA ( $n(\text{RNA})$ ). For the calculation you need the volume of the RNA solution ( $V(\text{RNA})$ ) and the extinction coefficient of the RNA ( $\varepsilon(\text{RNA})$ ), which is provided by the vendor of the RNA or can be calculated by free accessible calculators, for example, OligoAnalyzer Tool by IDT [54].

$$\frac{A(260 \text{ nm}) \left[ \frac{1}{\text{cm}} \right]}{\varepsilon(\text{RNA}) \left[ \frac{\text{L}}{\text{cm mol}} \right]} \times V(\text{RNA}) [\text{L}] = n(\text{RNA}) [\text{mol}]$$

15. Concentrate the sample in the vacuum concentrator to a volume of 25–30  $\mu\text{L}$ . This volume can vary, depending on which HPLC-injection loop and HPLC column is used in the next step (see **Note 7**).

## 3.2 HPLC Purification of the Labeled RNA

Perform a reverse phase HPLC purification of your sample. A general workflow and exemplary chromatograms are shown in Fig. 3.

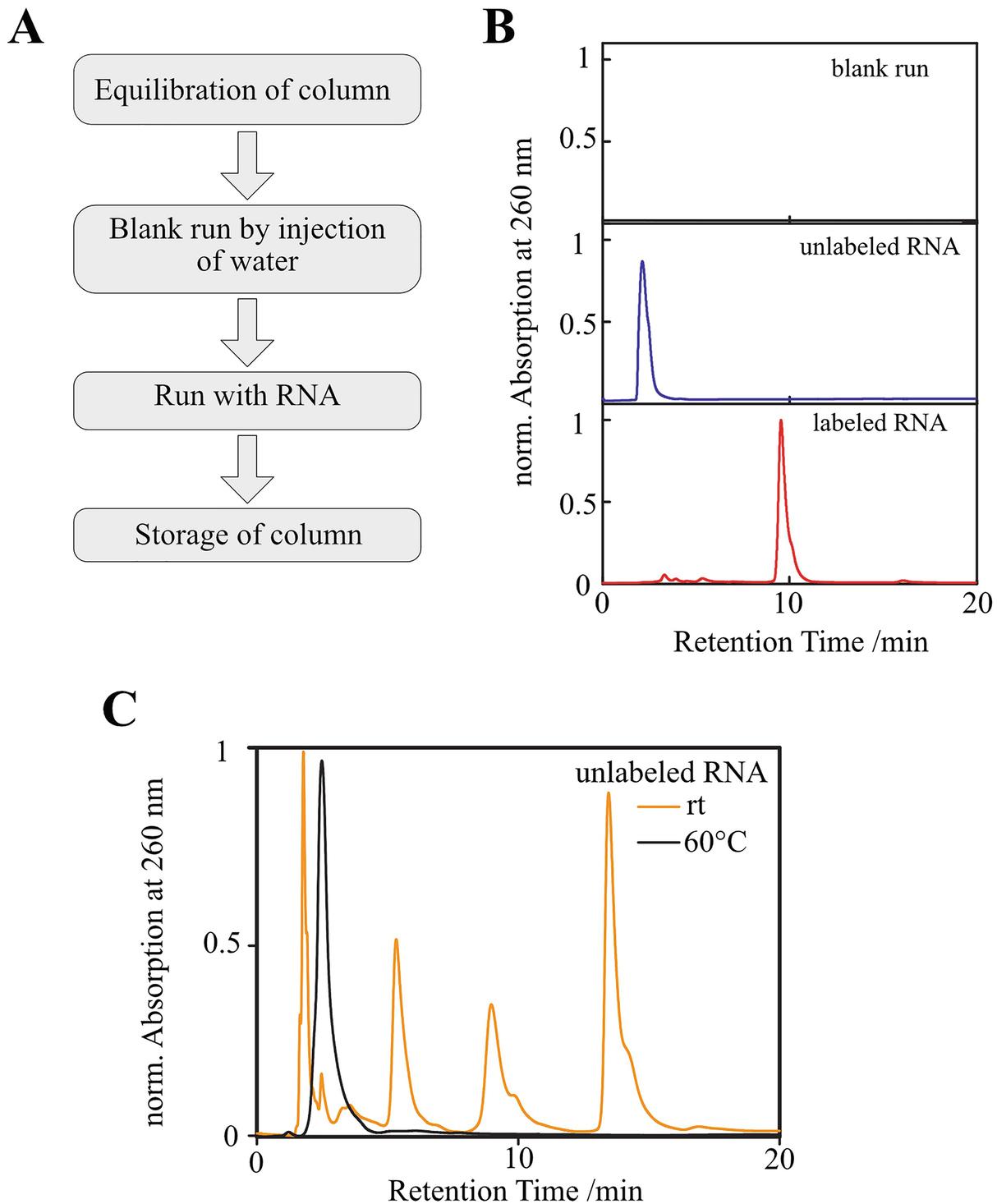
1. Equilibrate the HPLC column by starting a method, which rebuffers the column to 8% buffer A.
2. Run a blank run, load 30–50  $\mu\text{L}$  MilliQ water with a Hamilton syringe into the injection loop and start the same method used later for RNA purification. The HPLC running method should be optimized for each new RNA sequence, but as a first try, a gradient of 8%  $\rightarrow$  25% buffer B in 20 min with a flow rate of 1.5 mL/min can be used for RNA sequences shorter than 50 nt and should yield good results. For longer

RNA strands, we routinely run the purification with the preparative Zobrax 300SB-C18, 9.4 × 50 mm column by elution of 8% acetonitrile for 10 min and then applying a gradient of 8% → 23% in 55 min with a constant flow rate of 2 mL/min. Moreover, one should consider heating the column. We routinely heat it to 60 °C to avoid the formation of secondary structures (see Fig. 3c) and we recommend to heat the column already during the blank run.

3. Perform the RNA purification, load the resuspended RNA from a reaction onto the HPLC column, by injecting 25–30 µL with a Hamilton syringe into the loading loop 100 µL). Do not inject more than 2.5 nmol RNA, to avoid overloading the column specified above.
4. Run the HPLC method (see step 2).
5. Collect the fractions containing unlabeled and labeled RNA manually into 2 mL reaction tubes or use a fraction collector. We collect each peak, which exceeds an absorbance of 15 mAU.
6. Perform the next RNA purification, or store the HPLC column by starting a method, which rebuffers the column to 80% buffer A and 20% water.
7. Concentrate the collected fractions in the vacuum concentrator and pool them to yield a solution of about 500 µL for each HPLC peak containing labeled and unlabeled RNA, separately.
8. Desalt the RNA solution with an ultrafiltration centrifugal device, by following the manufacturer's instructions, and as already described in Subheading 3.1, steps 8–13 (see Notes 8 and 9).
9. Concentrate your sample in the vacuum concentrator from about 100 µL to the volume or concentration you need.
10. Determine the concentration of the labeled RNA by using the spectrophotometer as described in Subheading 2.1, step 14. We usually get yields of about 50–70% with respect to the amount of starting RNA. The losses occur to large extent during the purification steps.

11.

Determine the identity and purity of the labeled RNA (no spin label reduction or degradation) by LC-MS. Here we use a gradient of 5% → 20% buffer D in 20 min with a constant flow rate of 0.4 mL/min. Determine the labeling efficiency by using a *continuous wave (cw)*-EPR spectrometer. For this, the “spin count” routine from Bruker can be used. Alternatively, the spectrum can be double integrated, and the number of spins present in the sample calculated in relation to a reference sample. Usually, the labeling is quantitative.



**Fig. 3** HPLC-purification of labeled RNA. **(A)** Scheme of the general HPLC-workflow. Before loading the RNA, the column needs to be equilibrated with the buffer in which the purification is performed. We recommend to run a blank at 60 °C, by injecting water into the column. This enables to detect possible impurities on the column, originating from previous runs. After RNA purification, the column is brought to storage conditions. **(B)** Exemplary normalized chromatograms of the blank (black), unmodified RNA (blue), and

labeled RNA (red) runs. Due to the decrease of the RNA's polarity upon labeling, the retention time increases with respect to the unmodified RNA. (C) Normalized chromatograms of a 16 nt long RNA sequence performed at room temperature (orange) and at 60 °C (black). Upon column heating, the formation of secondary structures is inhibited, leading to a single peak

This RNA can be used for further EPR experiments or it can be ligated to a second RNA strand, which can be labeled or unlabeled.

### 3.3 Enzymatic Ligation

1. Clean your work bench and tools with 70% ethanol or an RNase decontamination solution.
2. Preheat the thermomixer to 90 °C.
3. Prepare the annealing mix by pipetting 20 µL of the spin-labeled RNA (acceptor strand containing a free 3'-hydroxy group, 2 nmol), 20 µL of a spin-labeled or unmodified RNA (5-phosphorylated donor strand, 2 nmol), 20.8 µL splint oligonucleotide (2.8 nmol), and 79.2 µL DEPC-treated water into a 1.5 mL Eppendorf tube (*see Note 10*).
4. Assemble the complex, heat this mixture for 2 min to 90 °C in the thermomixer and allow the mixture to cool down to room temperature for at least 15 min (*see Note 11*).
5. Set the thermomixer to 35 °C (*see Note 12*).
6. Prepare the reaction mixture for enzymatic ligation by pipetting 350 µL Annealing mix, 20 µL 10× T4 DNA ligase buffer, 20 µL PEG 4000, and 20 µL T4 DNA ligase in an 1.5 mL reaction tube.
7. Incubate the reaction mixture for 3.5 h at 35 °C in the thermomixer.
8. Run a 4% agarose gel to track the stage of ligation product formation. For this, mix 5 µL of the reaction solution with 5 µL 2× RNA loading buffer and heat the mixture in the thermomixer to 95 °C for 5 min. For the ssRNA ladder, mix 2 µL ladder with 3 µL MilliQ water and 5 µL 2× RNA loading dye. Heat this mixture in the thermomixer to 90 °C

for 2 min and let it cool on ice for 5 min. Load the RNA ladder and your sample onto the gel and run the gel at 125 V for 40 min (*see Note 13*).

- Note 13.** If the ligation product is not fully formed, you can incubate for a longer period.
9. Perform a phenol–chloroform extraction to separate the RNA from T4 DNA ligase:
  10. Place absolute ethanol, 3 M NaOAc, and 70% ethanol (25 mL each in Falcon tubes) on ice and allow the solutions to cool down for at least 30 min.
  11. Add 500  $\mu$ L phenol to the reaction mixture, vortex the Eppendorf tube well and centrifuge for 3 min at  $14,000 \times g$ .
  12. Take the upper (aqueous) phase and mix with 2 mL chloroform, vortex vigorously, transfer this mixture in two 1.5 mL Eppendorf tubes and centrifuge for 3 min at  $14,000 \times g$ .
  13. Take the upper (aqueous) phase, transfer in a 2 mL Eppendorf tube and add 1/10 volume of 3 M NaOAc and three volumes of EtOH absolute. Incubate the mixture for 20 min at  $-80^{\circ}\text{C}$ .
  14. Centrifuge the Eppendorf tube at  $14,000 \times g$  for 30 min.
  15. Carefully remove the solution with a pipette.
  16. Wash the pellet with 100  $\mu$ L 70% ethanol, centrifuge again for 15 min at  $14,000 \times g$ .
  17. Remove the supernatant and dissolve the RNA in 30  $\mu$ L DEPC treated water (*see Note 7*).
  18. Check the formation and purity of the ligation product on a 4% agarose gel, as described in **step 8**.

### 3.4 Anion-Exchange HPLC of the Ligated RNA

Perform an anion-exchange HPLC purification of your samples. The general workflow is the same as described in Subheading 3.2 and the elution method should be optimized for each RNA strand individually.

1. Begin by equilibrating the anion-exchange HPLC column and by performing a blank run at 80 °C, with the same method used for your RNA samples. As a guide, one may start with the following run: begin with an elution for 6 min with buffer E, then use a gradient of 0% → 35% buffer F in 12 s, followed by a gradient of 35% → 50% buffer F within 17 min, stay at 50% buffer F for 2 min, increase the gradient to 50% → 100% of buffer F within 12 s, stay at 100% buffer F for 7 min and then decrease buffer F within a gradient of 100% → 0% in 12 s.
2. Perform the RNA purification with the same method as described in step 1; however, the run should be performed at 80 °C.
3. Collect the fractions containing RNA manually into 2 mL reaction tubes or use a fraction collector. We collect each peak, which exceeds an absorbance of 15 mAU.
4. Perform the next purification or store the HPLC column by starting a method, which rebuffers the column with the storage buffer.
5. Proceed with desalting and rebuffering your RNA, as described in Subheading 3.2, steps 7–10.

For the enzymatic ligation, we obtained yields of 30–40%.

---

## 4 Notes

1. Scale up: We labeled many RNA sequences of variable length ranging from 16 nt to 59 nt and routinely perform two reactions of 2.5 nmol RNA in parallel. In principle, 5 nmol or more are also possible. However, scaling up the reaction might require further optimizations of the reaction conditions.
2. The crucial step in using T4 DNA ligase is the design of the splint and the formation of the acceptor-RNA/splint/donor-RNA complex (from

now on called active RNA complex, see Fig. 2b). In terms of base pairing, the splint needs to be paired strong enough to the RNAs to break other intramolecular base pairs. However, at the same time, the splint oligonucleotide also needs to be removed from the ligation product, meaning that the base pairing should not be too strong. Moreover, by assembling the active RNA complex, the strands should match to each other without forming gaps or bulges. Thus, it is important to carefully design all RNA strands and the splint [48].

3. 4% agarose is difficult to dissolve in water. Thus, it is important to first wash the agarose–Synergel mixture with ethanol absolute. Mix this in circular motions and wait 5–10 min. This will allow the agarose and Synergel to set on the bottom of, for example, your beaker and will allow an easy decant of the ethanol.
4. The **steps 4** and **5** should follow immediately. Cu<sup>+</sup> is instable in solution and is quickly oxidized by atmospheric oxygen to Cu<sup>2+</sup>. This can be followed by the color of the Cu<sup>+</sup> solution: First, the solution presents a light-yellow color, which gets darker the more CuI dissolves/the longer it is stirred. Once the CuI is completely dissolved, the yellow solution is ready to be used for the complexation (**step 5**). Within about 5–15 min the solution will turn from yellow to dark green and finally brown and should not be used anymore.
5. This step enables the formation of the THPTA-Cu<sup>+</sup> complex and inhibits possible Cu<sup>+</sup>-mediated hydrolysis of RNA during labeling. Here it is important to wait 5 min for complete Cu<sup>+</sup> complexation! THPTA should be used in large excess to enable full complexation.
6. The reaction time needs to be adjusted for each RNA strand individually. If the reaction time is too short, the conversion of unlabeled RNA will not be complete. If the reaction time is too long partial reduction of the spin label will occur, reducing the yield of spin-labeled RNA. Note that the retention time of labeled RNA containing a reduced spin label is slightly shorter compared to spin-labeled RNA (see Fig. 3b). The appropriate reaction time can be determined by analyzing 1 or 2 µL of reaction mixture on reverse HPLC every 5–10 min. Note that every sample should be previously

desalting using an ultrafiltration centrifugal device.

7.

We use a 100 µL HPLC-injection loop and load up to a maximum of 50 µL. In general, the sample volume should not exceed 50% of the loop volume to avoid possible sample loss.

8.

Instead of using an ultrafiltration centrifugal device, a disposable DNA purification column may also be used to desalt the sample. It is recommended to check for RNA loss during both desalting methods to determine which one works best. We observed for dried, labeled RNA strands of >40–50 nt length, that these cannot be resolubilized in water. Instead, we observed a white precipitate. Thus, it is preferable to directly rebuffer this RNA into the desired buffer, instead of consecutively desalting and concentrating it.

9.

When labeling a new RNA strand, we recommend checking every peak occurring during HPLC purification via LC-MS.

10.

We use the splint in excess to ensure the annealing of the splint as a bridge between donor and acceptor RNA strand.

11.

The complex between donor RNA, acceptor RNA, and splint needs to be formed before adding the T4 DNA ligase. Here it is important to wait at least 15 min for complete complexation.

12.

The optimal temperature for the T4 DNA ligase used in this protocol is 35 °C [36], but this depends on the vendor. The temperature needs to be tested and adjusted for each new sequence. Take care that the temperature is not too high, because the T4 DNA ligase used here is inactivated at 65 °C.

13.

The running time of the agarose gel needs to be adjusted, depending on the length of your RNA sequence. We run an agarose gel for 40 min for an RNA sequence length of 81 nt.

---

## References

1. Meyers RA (ed) (2014) RNA regulation: advances in molecular biology and medicine. Wiley-VCH-Verl, Weinheim
2. Lau MWL, Ferré-D'Amaré AR (2016) Riboswitches and ribozymes. In: Encyclopedia of Cell Biology. Elsevier, Amsterdam, pp 375–383  
[Crossref]
3. Canny MD, Jucker FM, Pardi A (2007) Efficient ligation of the *Schistosoma* hammerhead ribozyme. Biochemistry 46(12):3826–3834. <https://doi.org/10.1021/bi062077r>  
[Crossref][PubMed]
4. Schultes EA, Bartel DP (2000) One sequence, two ribozymes: implications for the emergence of new ribozyme folds. Science 289(5478):448–452. <https://doi.org/10.1126/science.289.5478.448>  
[Crossref][PubMed]
5. Talini G, Branciamore S, Gallori E (2011) Ribozymes: flexible molecular devices at work. Biochimie 93(11):1998–2005. <https://doi.org/10.1016/j.biochi.2011.06.026>  
[Crossref][PubMed]
6. Lilley DMJ, Eckstein F (2007) Ribozymes and RNA catalysis. Royal Society of Chemistry, Cambridge  
[Crossref]
7. Flynn-Charlebois A, Wang Y, Prior TK et al (2003) Deoxyribozymes with 2'-5' RNA ligase activity. J Am Chem Soc 125(9):2444–2454. <https://doi.org/10.1021/ja028774y>  
[Crossref][PubMed]
8. Silverman SK (2015) Pursuing DNA catalysts for protein modification. Acc Chem Res 48(5):1369–1379. <https://doi.org/10.1021/acs.accounts.5b00090>  
[Crossref][PubMed][PubMedCentral]
9. Rosenbach H, Victor J, Etzkorn M et al (2020) Molecular features and metal ions that influence 10-23 DNAzyme activity. Molecules 25(13):3100. <https://doi.org/10.3390/molecules25133100>  
[Crossref][PubMedCentral]
10. Goldfarb D, Stoll S (2018) EPR spectroscopy: fundamentals and methods. In: EMagRes Books. John Wiley & Sons Ltd, Chichester, West Sussex
11. Schweiger A, Jeschke G (2001) Principles of pulse electron paramagnetic resonance.

Oxford University Press, Oxford

12. Ward RJ, Schiemann O (2012) EPR based distance measurements in oligonucleotides. *Struct Bond* 152:249–282. [https://doi.org/10.10007/430\\_2012\\_76](https://doi.org/10.10007/430_2012_76) [Crossref]
13. Drescher M, Jeschke G (2012) EPR Spectroscopy, vol 321. Springer, Berlin Heidelberg, Berlin, Heidelberg [Crossref]
14. Abdullin D, Schiemann O (2020) Pulsed dipolar EPR spectroscopy and metal ions: methodology and biological applications. *ChemPlusChem* 85(2):353–372. <https://doi.org/10.1002/cplu.201900705> [Crossref][PubMed]
15. Weber S (2007) Transient EPR. In: Harris RK, Wasylishen RL (eds) eMagRes. John Wiley & Sons, Ltd, Chichester, UK, pp 255–270
16. Malygin AA, Krumkacheva OA, Graifer DM et al (2019) Exploring the interactions of short RNAs with the human 40S ribosomal subunit near the mRNA entry site by EPR spectroscopy. *Nucleic Acids Res* 47(22):11850–11860. <https://doi.org/10.1093/nar/gkz1039> [Crossref][PubMed][PubMedCentral]
17. Pliotas C, Ward R, Branigan E et al (2012) Conformational state of the MscS mechanosensitive channel in solution revealed by pulsed electron-electron double resonance (PELDOR) spectroscopy. *Proc Natl Acad Sci U S A* 109(40):E2675–E2682. <https://doi.org/10.1073/pnas.1202286109> [Crossref][PubMed][PubMedCentral]
18. Constantinescu-Aruxandei D, Petrovic-Stojanovska B, Schiemann O et al (2016) Taking a molecular motor for a spin: helicase mechanism studied by spin labeling and PELDOR. *Nucleic Acids Res* 44(2):954–968. <https://doi.org/10.1093/nar/gkv1373> [Crossref][PubMed]
19. Yang Z, Liu Y, Borbat P et al (2012) Pulsed ESR dipolar spectroscopy for distance measurements in immobilized spin labeled proteins in liquid solution. *J Am Chem Soc* 134(24):9950–9952. <https://doi.org/10.1021/ja303791p> [Crossref][PubMed][PubMedCentral]
20. Kuzhelev AA, Krumkacheva OA, Shevelev GY et al (2018) Room-temperature distance measurements using RIDME and the orthogonal spin labels trityl/nitroxide. *Phys Chem Chem Phys* 20(15):10224–10230. <https://doi.org/10.1039/C8CP01093E> [Crossref][PubMed]

21. Peter MF, Tuukkanen AT, Heubach CA et al (2019) Studying conformational changes of the Yersinia type-III-secretion effector YopO in solution by integrative structural biology. *Structure* 27(9):1416–1426.e3. <https://doi.org/10.1016/j.str.2019.06.007> [Crossref][PubMed]
22. Duss O, Michel E, Yulikov M et al (2014) Structural basis of the non-coding RNA RsmZ acting as a protein sponge. *Nature* 509(7502):588–592. <https://doi.org/10.1038/nature13271> [Crossref][PubMed]
23. Duss O, Yulikov M, Jeschke G et al (2014) EPR-aided approach for solution structure determination of large RNAs or protein-RNA complexes. *Nat Commun* 5:3669. <https://doi.org/10.1038/ncomms4669> [Crossref][PubMed]
24. Orwick-Rydmark M, Lovett JE, Graziadei A et al (2012) Detergent-free incorporation of a seven-transmembrane receptor protein into nanosized bilayer Lipodisq particles for functional and biophysical studies. *Nano Lett* 12(9):4687–4692. <https://doi.org/10.1021/nl3020395> [Crossref][PubMed]
25. Ward R, Pliotas C, Branigan E et al (2014) Probing the structure of the mechanosensitive channel of small conductance in lipid bilayers with pulsed electron-electron double resonance. *Biophys J* 106(4):834–842. <https://doi.org/10.1016/j.bpj.2014.01.008> [Crossref][PubMed][PubMedCentral]
26. Fleck N, Heubach CA, Hett T et al (2020) SLIM: a short-linked, highly redox-stable Trityl label for high-sensitivity in-cell EPR distance measurements. *Angew Chem Int Ed* 59(24):9767–9772. <https://doi.org/10.1002/anie.202004452> [Crossref]
27. Shilovskikh VV, Timralieva AA, Belogub EV et al (2020) Radical activity of binary melamine-based hydrogen-bonded self-assemblies. *Appl Magn Reson* 51(9-10):939. <https://doi.org/10.1007/s00723-020-01254-6> [Crossref]
28. Hunsicker-Wang L, Vogt M, DeRose VJ (2009) EPR methods to study specific metal-ion binding sites in RNA. In: Herschlag D (ed) Biophysical, chemical, and functional probes of RNA structure, interactions and folding: part A. Methods in enzymology, vol 468. Elsevier, San Diego, Burlington, London, pp 335–367 [Crossref]
29. Kisseeleva N, Khvorova A, Westhof E et al (2005) Binding of manganese(II) to a

- tertiary stabilized hammerhead ribozyme as studied by electron paramagnetic resonance spectroscopy. *RNA* 11(1):1–6. <https://doi.org/10.1261/rna.7127105> [Crossref][PubMed][PubMedCentral]
30. Woody AY, Eaton SS, Osumi-Davis PA et al (1996) Asp537 and Asp812 in bacteriophage T7 RNA polymerase as metal ion-binding sites studied by EPR, flow-dialysis, and transcription. *Biochemistry* 35(1):144–152. <https://doi.org/10.1021/bi952037f> [Crossref][PubMed]
  31. Heinz M, Erlenbach N, Stelzl LS et al (2020) High-resolution EPR distance measurements on RNA and DNA with the non-covalent spin label. *Nucleic Acids Res* 48(2):924–933. <https://doi.org/10.1093/nar/gkz1096>
  32. Reginsson GW, Shelke SA, Rouillon C et al (2013) Protein-induced changes in DNA structure and dynamics observed with noncovalent site-directed spin labeling and PELDOR. *Nucleic Acids Res* 41(1):e11. <https://doi.org/10.1093/nar/gks817> [Crossref][PubMed]
  33. Haugland MM, Lovett JE, Anderson EA (2018) Advances in the synthesis of nitroxide radicals for use in biomolecule spin labelling. *Chem Soc Rev* 47(3):668–680. <https://doi.org/10.1039/c6cs00550k> [Crossref][PubMed]
  34. Shelke SA, Sigurdsson ST (2013) Site-directed Nitroxide spin Labeling of biopolymers. In: Timmel CR, Harmer JR (eds) Structural information from spin-labels and intrinsic paramagnetic centres in the biosciences, vol 152. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 121–162  
[Crossref]
  35. Goldfarb D (2014) Gd<sup>3+</sup> spin labeling for distance measurements by pulse EPR spectroscopy. *Phys Chem Chem Phys* 16(21):9685–9699. <https://doi.org/10.1039/c3cp53822b> [Crossref][PubMed]
  36. Ghosh S, Lawless MJ, Brubaker HJ et al (2020) Cu<sup>2+</sup>-based distance measurements by pulsed EPR provide distance constraints for DNA backbone conformations in solution. *Nucleic Acids Res* 48(9):e49. <https://doi.org/10.1093/nar/gkaa133> [Crossref][PubMed][PubMedCentral]
  37. Nguyen PH, Popova AM, Hideg K et al (2015) A nucleotide-independent cyclic nitroxide label for monitoring segmental motions in nucleic acids. *BMC Biophys* 8:6. <https://doi.org/10.1186/s13628-015-0019-5> [Crossref][PubMed][PubMedCentral]

38. Qin PZ, Haworth IS, Cai Q et al (2007) Measuring nanometer distances in nucleic acids using a sequence-independent nitroxide probe. *Nat Protoc* 2(10):2354–2365. <https://doi.org/10.1038/nprot.2007.308> [Crossref][PubMed][PubMedCentral]
39. Sicoli G, Wachowius F, Bennati M et al (2010) Probing secondary structures of spin-labeled RNA by pulsed EPR spectroscopy. *Angew Chem Int Ed* 49(36):6443–6447. <https://doi.org/10.1002/anie.201000713> [Crossref]
40. Kerzhner M, Abdullin D, Więcek J et al (2016) Post-synthetic spin-Labeling of RNA through click chemistry for PELDOR measurements. *Chem Eur J* 22(34):12113–12121. <https://doi.org/10.1002/chem.201601897> [Crossref][PubMed]
41. Schiemann O, Piton N, Plackmeyer J et al (2007) Spin labeling of oligonucleotides with the nitroxide TPA and use of PELDOR, a pulse EPR method, to measure intramolecular distances. *Nat Protoc* 2(4):904–923. <https://doi.org/10.1038/nprot.2007.97> [Crossref][PubMed]
42. Piton N, Mu Y, Stock G et al (2007) Base-specific spin-labeling of RNA for structure determination. *Nucleic Acids Res* 35(9):3128–3143. <https://doi.org/10.1093/nar/gkm169> [Crossref][PubMed][PubMedCentral]
43. Saha S, Jagtap AP, Sigurdsson ST (2015) Site-directed spin labeling of 2'-amino groups in RNA with isoindoline nitroxides that are resistant to reduction. *Chem Commun* 51(66):13142–13145. <https://doi.org/10.1039/c5cc05014f> [Crossref]
44. Edwards TE, Sigurdsson ST (2007) Site-specific incorporation of nitroxide spin-labels into 2'-positions of nucleic acids. *Nat Protoc* 2(8):1954–1962. <https://doi.org/10.1038/nprot.2007.273> [Crossref][PubMed]
45. Kerzhner M, Matsuoka H, Wuebben C et al (2018) High-yield spin Labeling of long RNAs for electron paramagnetic resonance spectroscopy. *Biochemistry* 57(20):2923–2931. <https://doi.org/10.1021/acs.biochem.8b00040> [Crossref][PubMed]
46. Babaylova ES, Malygin AA, Lomzov AA et al (2016) Complementary-addressed site-directed spin labeling of long natural RNAs. *Nucleic Acids Res* 44(16):7935–7943. <https://doi.org/10.1093/nar/gkw516>

[Crossref][PubMed][PubMedCentral]

47. Domnick C, Eggert F, Wuebben C et al (2020) EPR distance measurements on long non-coding RNAs empowered by genetic alphabet expansion transcription. *Angew Chem Int Ed* 59(20):7891–7896. <https://doi.org/10.1002/anie.201916447>  
[Crossref]
48. Lang K, Micura R (2008) The preparation of site-specifically modified riboswitch domains as an example for enzymatic ligation of chemically synthesized RNA fragments. *Nat Protoc* 3(9):1457–1466. <https://doi.org/10.1038/nprot.2008.135>  
[Crossref][PubMed]
49. Hartmann RK, Bindereif A, Schön A et al (2014) *Handbook of RNA biochemistry*. Wiley-VCH Verlag GmbH & Co, KGaA, Weinheim, Germany  
[Crossref]
50. Hein JE, Fokin VV (2010) Copper-catalyzed azide-alkyne cycloaddition (CuAAC) and beyond: new reactivity of copper(I) acetylides. *Chem Soc Rev* 39(4):1302–1315.  
<https://doi.org/10.1039/b904091a>  
[Crossref][PubMed][PubMedCentral]
51. Rostovtsev VV, Green LG, Fokin VV et al (2002) A stepwise Huisgen cycloaddition process: copper(I)-catalyzed Regioselective “ligation” of Azides and terminal alkynes. *Angew Chem Int Ed* 41(14):2596–2599. [https://doi.org/10.1002/1521-3773\(20020715\)41:14<2596:AID-ANIE2596>3.0.CO;2-4](https://doi.org/10.1002/1521-3773(20020715)41:14<2596:AID-ANIE2596>3.0.CO;2-4)  
[Crossref]
52. Wuebben C, Blume S, Abdullin D et al (2019) Site-directed spin Labeling of RNA with a gem-Diethylisoindoline spin label: PELDOR, relaxation, and reduction stability. *Molecules* 24(24):4882. <https://doi.org/10.3390/molecules24244482>  
[Crossref]
53. Wuebben C, Vicino MF, Mueller M et al (2020) Do the P1 and P2 hairpins of the guanidine-II riboswitch interact? *Nucleic Acids Res* 48(18):10518–10526. <https://doi.org/10.1093/nar/gkaa703>  
[Crossref][PubMed][PubMedCentral]
54. IDT OligoAnalyzer Tool. <https://eu.idtdna.com/pages/tools/oligoanalyzer>. Accessed 17 Sep2020

## 15. Preparation of Site-Specifically Spin-Labeled RNA by in Vitro Transcription Using an Expanded Genetic Alphabet

Lisa Bornewasser<sup>1</sup> and Stephanie Kath-Schorr<sup>1</sup> 

(1) Department of Chemistry, University of Cologne, Köln, Germany

 **Stephanie Kath-Schorr**  
Email: [skaths@uni-koeln.de](mailto:skaths@uni-koeln.de)

### Abstract

Recent advances in pulsed electron paramagnetic resonance (EPR) spectroscopy enable studying structure and folding of nucleic acids. An efficient introduction of spin labels at specific positions within the oligonucleotide sequence is a prerequisite. We here present a step-by-step guide to synthesize long RNA oligonucleotides bearing spin labels at specific positions within the sequence. RNA preparation is achieved enzymatically via in vitro transcription using an expanded genetic alphabet. Highly structured, several hundred nucleotides long RNAs with two nitroxide spin labels at specific positions can be prepared by this method.

**Key words** Spin labeling – RNA – Electron paramagnetic resonance (EPR) – Pulsed electron–electron double resonance (PELDOR) – Expanded genetic alphabet – Unnatural base pairs – In vitro transcription

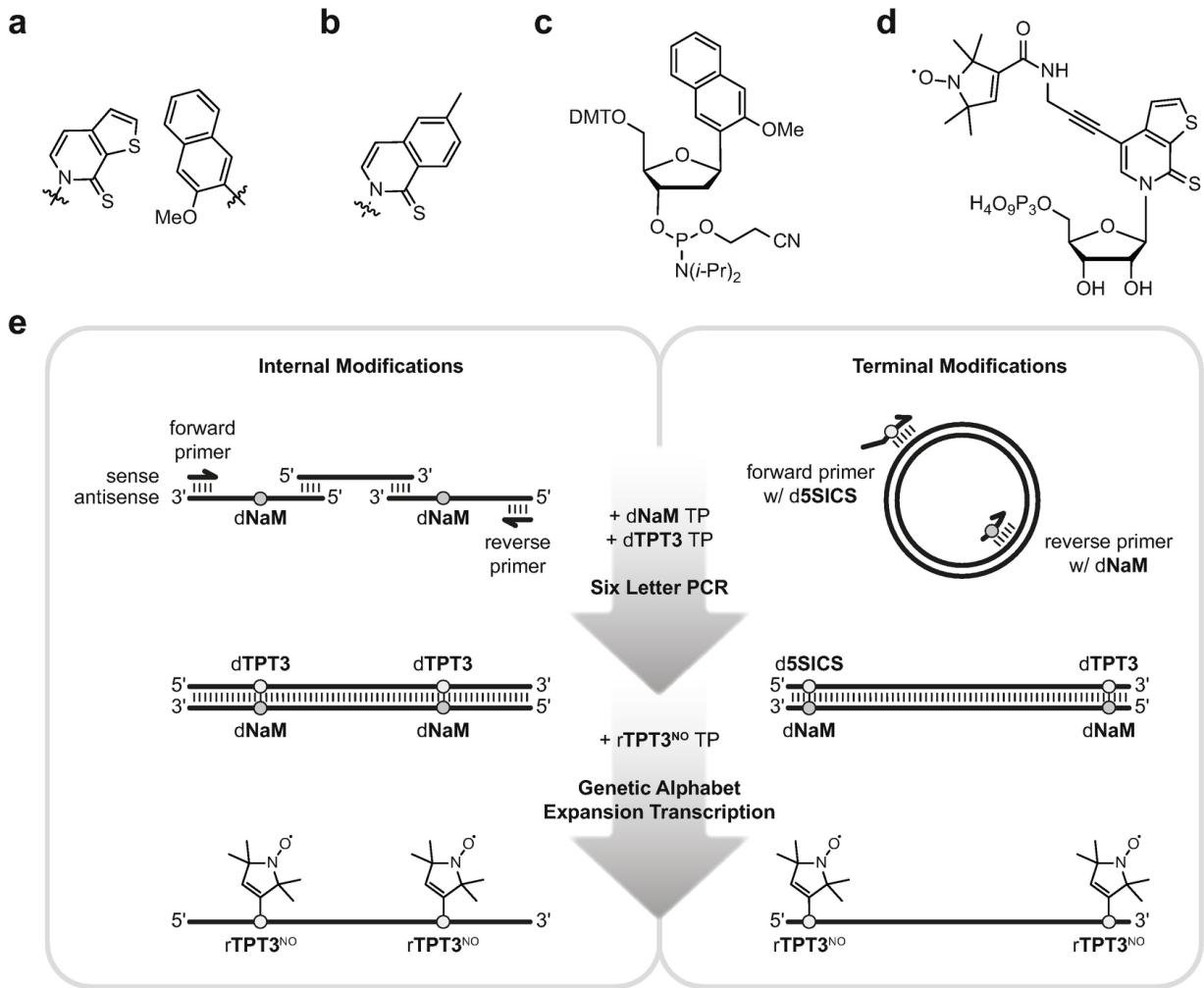
---

### 1 Introduction

Pulsed electron–electron double resonance (PELDOR) is a powerful approach to study the global fold of biomolecules and has been applied to measure nanometer distances in both DNA and RNA [1–12]. The efficient, site-specific introduction of spin labels, usually nitroxides, into oligonucleotides is a critical step for successful PELDOR experiments.

The introduction of nitroxyl spin labels into long, complexly folded RNA molecules has remained a challenge [1, 13–19]. Common sizes of most naturally occurring noncoding RNA molecules are several hundred to several thousand nucleotides in lengths, thus, these RNA molecules are not accessible by direct solid phase synthesis [17–20]. With the recent progress in identifying regulatory long noncoding RNA molecules and studying their cellular distribution and function, a particular focus has been set on the investigation of structure and function of those oligonucleotides [21].

Here we provide a step-by-step approach for using an expanded genetic alphabet to introduce nitroxyl spin labels at specific positions into long RNA oligonucleotides during in vitro transcription [22]. For this, unnatural hydrophobic base pairs [23–30], are introduced into the DNA template to direct the site-specific introduction of unnatural, spin-labeled ribonucleoside triphosphates during RNA transcription [31–39]. The preparation of double-stranded DNA templates bearing two unnatural base pairs [30], the thieno substituted pyridine-2(1H)-thione derivative dTPT3 and the 3-methoxy-2-naphthyl derivative dNaM (Fig. 1a), for in vitro transcription is achieved either by assembly from several short synthetic oligonucleotides using an overlapping PCR approach in the presence of the four canonical and two unnatural nucleoside triphosphates [37]. An alternative approach for template preparation presented in this chapter uses a plasmid containing the sequence of interest as template and PCR primers which introduce the unnatural base pairs via mismatches with the plasmid sequence [22] (Fig. 1).



**Fig. 1** Preparation of site-specifically spin-labeled RNA by in vitro transcription using an expanded genetic alphabet. **(a)** Unnatural base pair dTPT3:dNaM developed by Romesberg and coworkers [30] (sugar residues omitted for clarity). **(b)** Unnatural base d5SICS [40] (sugar residue omitted for clarity). d5SICS also pairs with dNaM. **(c)** dNaM cyanoethyl phosphoramidite for solid-phase synthesis of unnatural base modified DNA [40]. **(d)** rTPT3<sup>NO</sup> triphosphate (TP) for spin labeling of RNA by in vitro transcription [22]. **(e)** Schematic approach to generate spin-labeled RNA with either internal modification (left panel) or modifications close to the 5' and 3' terminal sites (right panel)

The method allows the facile preparation of large, spin-labeled RNA molecules via standard T7 in vitro transcription in nanomole quantities and provides a valuable alternative to chemical solid phase synthesis of large spin-labeled RNAs which requires complex ligation strategies [17, 18, 20] to prepare these functionalized oligonucleotides.

---

## 2 Materials

Prepare all solutions using ultrapure water. Perform all working steps at room temperature and store reagents at  $-20\text{ }^{\circ}\text{C}$  (unless indicated otherwise).

### 2.1 DNA Primers and Oligonucleotides

1. Phosphoramidites for solid phase DNA synthesis: dNaM cyanoethyl phosphoramidite and d5SICS cyanoethyl phosphoramidite (*see Note 1*): Can be purchased for example from Berry & Associates Inc., USA.
2. dNaM- and d5SICS-modified DNA primers and oligonucleotides: Design according to the sequence of interest and prepare by solid phase synthesis. Employing commercial oligonucleotide synthesis services (*see Note 2*), primers and oligonucleotides will be supplied either lyophilized or in aqueous solution. If lyophilized, dissolve in ultrapure water to a concentration of  $100\text{ }\mu\text{M}$ . Store dissolved primers and oligonucleotides at  $-20\text{ }^{\circ}\text{C}$ .
3. Unmodified DNA primers and oligonucleotides: Need to be designed according to the sequence of interest and prepared by solid phase synthesis (*see Note 3*).

### 2.2 DNA Template Preparation by Six Letter PCR

Depending on the desired positions for spin labeling within the RNA sequence of interest, unnatural base modifications can be introduced internally or into the 3'- or 5'-region (within 10–20 nucleotides (nt) from each terminal site) of the DNA template. For internal modifications, a two-step six letter fusion PCR approach is applied using overlapping dNaM-modified and unmodified DNA oligonucleotides and primers [30]. To introduce unnatural nucleotides within the 3'- and 5'-region of the RNA (here referred to as “terminal” modification sites), a six letter PCR approach is performed using dNaM-modified forward and d5SICS-modified reverse primers and an unmodified DNA template for PCR amplification [22, 30]. A mix of both approaches to combine internal and terminal modification sites can also be employed.

1. dNaM nucleoside (e.g., from Berry & Associates Inc., USA).
2. dNaM triphosphate (TP) (e.g., from MyChem LLC, USA).  
10 mM dNaM TP in ultrapure water, freeze in aliquots of 50 µL at -20 °C.
3. dTPT3 TP: Needs to be synthesized according to literature [30].  
10 mM dTPT3 TP in ultrapure water, freeze in aliquots of 50 µL at -20 °C.
4. Canonical dNTP Mix: 25 mM each dATP, dTTP, dGTP, dCTP in aqueous solution).
5. 5 U/µL OneTaq® DNA Polymerase.
6. 5× OneTaq® Standard Reaction Buffer: 20 mM Tris–HCl pH 8.9, 22 mM NH<sub>4</sub>Cl, 22 mM KCl, 1.8 mM MgCl<sub>2</sub>, 0.06% IGEPAL® CA-630, 0.05% Tween® 20.
7. Ultrapure water.
8. Autoclaved plastic ware: PCR tubes, pipette tips (free from human DNA, DNase, and RNase contamination).
9. Manual pipettes.
10. Thermal cycler.
11. Vortex mixer.

### ***2.2.1 For Six Letter Fusion PCR (for Internal Modification Sites)***

1. 100 µM dNaM-modified antisense (=template) oligonucleotide(s).
2. 100 µM unmodified sense oligonucleotide(s).

3. 100 µM forward DNA primer.
4. 100 µM reverse DNA primer with two 2'-O-methoxy modified nucleotides at the 5'-end (*see Note 4*).
5. 100 mM MgCl<sub>2</sub> (*see Note 5*).

### ***2.2.2 Alternatively, for Six Letter PCR from Plasmid DNA (for Terminal Modification Sites)***

1. 100 µM dNaM-modified forward DNA primer.
2. 100 µM d5SICS-modified reverse DNA primer with double 2'-O-methoxy-modification at 5'-end (*see Note 4*).
3. 50 ng/µL unmodified template DNA, for example plasmid DNA.

## **2.3 Analysis and Purification of PCR Products**

1. High Resolution Agarose.
2. 0.5× TBE (44.5 mM Tris, 44.5 mM boric acid, 1 mM EDTA) buffer.
3. Microwave oven.
4. 50 mL Conical tubes.
5. 1% Ethidium bromide (EtBr) solution.
6. 6× Agarose loading dye.
7. DNA ladder for reference (e.g. Low Range ssRNA Ladder) or GeneRuler Ultra Low Range DNA Ladder).

8. Gel casting system.
9. Agarose gel electrophoresis system.
10. Gel documentation system (*see Note 11*).
11. NucleoSpin® Gel and PCR cleanup kit.
12. Ultrapure water.
13. Absolute ethanol.
14. Autoclaved plastic ware: 1.5 mL microcentrifuge tubes, 2.5 µL, 10 µL, 200 µL, 1000 µL pipette tips (free from human DNA, DNase, and RNase contamination).
15. Centrifuge with fixed angle rotor for 1.5 mL microcentrifuge tubes.
16. Manual pipettes.
17. Heating block (70 °C).
18. Vortex mixer.
19. Centrifuge for microcentrifuge tubes.
20. UV-Vis spectrophotometer.
21. Oligo Calc: Oligonucleotide Properties Calculator [43].

## 2.4 Genetic Alphabet Expansion Transcription

1. Purified dNaM- and/or d5SICS-modified dsDNA template (*see Note 6*).

2. 200 mM Tris–HCl pH 7.9 (*see Notes 7 and 8*).
3. 100 mM MgCl<sub>2</sub>.
4. 100 mM dithiothreitol (DTT) solution (*see Note 9*).
5. Canonical rNTPs Mix: 25 mM each rATP, rUTP, rGTP, rCTP in aqueous solution.
6. Nitroxide modified TPT3 derivative rTPT3<sup>NO</sup> TP: Needs to be synthesized according to literature [22].  
10 mM rTPT3<sup>NO</sup> TP in ultrapure water, freeze in aliquots of 50 µL at -20 °C.
7. 40 U/µL RNasin Plus® RNase inhibitor.
8. 0.1 U/µL Inorganic pyrophosphatase (iPP).
9. 50 U/µL T7 RNA polymerase (self-made [37], *see Note 10*).
10. DNase (2 U/µL, RNase free).
11. DNase buffer (100 mM Tris–HCl pH 7.6, 25 mM MgCl<sub>2</sub>, 5 mM CaCl<sub>2</sub>) (*see Note 7*).
12. Ultrapure water.
13. Autoclaved plastic ware: 1.5 mL microcentrifuge tubes, 2.5 µL, 10 µL, 200 µL, 1000 µL pipette tips (free from human DNA, DNase, and RNase contamination).
14. Manual pipettes.
15. Heating block (37 °C and 95 °C).
16. Vortex mixer.

## 2.5 RNA Purification and Concentration Determination

1. High Resolution Agarose (*Carl Roth*).
2. 0.5× TBE (44.5 mM Tris, 44.5 mM Boric acid, 1 mM EDTA) buffer.
3. Microwave oven.
4. 50 mL Conical tubes.
5. 1% Ethidium bromide solution.
6. 6× Agarose loading dye.
7. RNA or DNA ladder for reference (for example: Low Range ssRNA Ladder or GeneRuler Ultra Low Range DNA Ladder).
8. Gel casting system.
9. Agarose gel electrophoresis system.
10. Gel documentation system (*see Note 11*).
11. Scalpel.
12. Micro scales.
13. Heating block (50 °C and 70 °C).
14. NucleoSpin® Gel and PCR cleanup kit (Macherey-Nagel).
15. Binding-Buffer NTC (Macherey-Nagel) (*see Note 12*).

16. Ultrapure water.
17. Absolute ethanol.
18. Autoclaved plastic ware: 1.5 mL microcentrifuge tubes, 2.5  $\mu$ L, 10  $\mu$ L, 200  $\mu$ L, 1000  $\mu$ L pipette tips (free from human DNA, DNase, and RNase contamination).
19. Centrifuge with fixed angle rotor for 1.5 mL microcentrifuge tubes.
20. Manual pipettes.
21. Vortex mixer.
22. UV-Vis spectrophotometer.
23. Oligo Calc: Oligonucleotide Properties Calculator [43].

## 2.6 Sample Preparation for EPR Measurements

1. 100 mM MgCl<sub>2</sub> in D<sub>2</sub>O (sterile filtered).
2. PELDOR buffer (145 mM NaCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 10 mM NaH<sub>2</sub>PO<sub>4</sub>, in D<sub>2</sub>O (99.9%), sterile filtered).
3. Ethylene glycol-d<sub>6</sub>.
4. Heating block (70 °C, gradient 70 → 18 °C).
5. Autoclaved plastic ware: 1.5 mL microcentrifuge tubes, 2.5  $\mu$ L, 10  $\mu$ L, 200  $\mu$ L, 1000  $\mu$ L pipette tips (free from human DNA, DNase, and RNase contamination).
6. Vortex mixer.
- 7.

- .. Manual pipettes.
  - 8. Freeze-dryer.
  - 9. Amicon® Ultra-0.5 3K Centrifugal Filter Device.
  - 10. Centrifuge with fixed angle rotor for 1.5 mL microcentrifuge tubes.
- 

### 3 Methods

#### 3.1 Preparation of DNA Templates for T7 In Vitro Transcription by Six Letter PCR

In accordance to the positions for the introduction of spin labels within the target RNA sequence, unnatural base modifications can be introduced internally or into the 3'- or 5'-region (within 10–20 nt from each terminal site, referred to as “terminal” modification) of the DNA template. For internal modifications, a two-step six letter fusion PCR approach is applied using overlapping dNaM-modified antisense and unmodified sense DNA oligonucleotides and primers. The protocol for internal modifications described in the following was optimized for a 185 nt long RNA sequence which was divided into two dNaM-modified antisense and one unmodified sense DNA oligonucleotide. For 3'- and 5'-terminal modification sites, a six letter PCR approach is performed using a single dNaM-modified forward primer and a single d5SICS-modified reverse primer in combination with an unmodified template DNA. The protocol for the introduction of spin labels within the 3'- and 5'-region of the RNA described in the following was optimized for a 377 nt long RNA sequence which was amplified from unmodified plasmid DNA introducing the unnatural base pair modifications and the T7 promotor sequence simultaneously during PCR. A mix of both approaches to combine internal and terminal modification sites is also possible but is not described here.

1. Design DNA oligonucleotides and primers (for internal modification sites) or only DNA primers (for modifications sites close to the 5' and 3' terminal sites) for your sequence of interest. Choose labeling sites within your sequence of interest (for internal modification sites) or

close to the 5' and 3' terminal sites of your sequence of interest (for terminal modifications sites). The modification sites should only minimally perturb the RNA's structure and function. Introduce dNaM modifications in the DNA template preferably by point mutations of either dA or dT (see **Note 13**). For internal modification sites, divide your sequence of interest into shorter segments to enable solid phase synthesis of dNaM modified DNA oligonucleotides (see **Note 14**). Include overlapping sequences between sense and antisense DNA strands. For terminal modification sites, introduce dNaM modifications in the reverse primer and d5SICS modifications in the forward primer. The DNA template needs to comprise the T7 promoter sequence (see **Note 15**) for subsequent T7 in vitro transcription using T7 RNA polymerase. Therefore, include the T7 promoter sequence, if it is not already present in your DNA template sequence, at the 5'-end of the first sense or at the 3'-end of the last antisense strand (for internal modification sites, reading direction 5' → 3') or at the 5'-end of your forward primer as an overhang sequence (for terminal modification sites). Also, your DNA template should include one to two 2'-O-methoxy-modified nucleotides at the 5'-end of the antisense strand (=template strand). Hence, include these modifications at the 5'-end of the first antisense strand and at the 5'-end of your reverse primer (for internal modification sites, reading direction 5' → 3') or only at the 5'-end of your reverse primer (for terminal modification sites).

2. Pipet PCR reaction mix into a sterile PCR tube. Prepare a 100 µL PCR mix of 20 µL 5× OneTaq® Standard Reaction Buffer, 1.5 µL canonical dNTP mix, 2 µL dNaM TP, 2 µL dTPT3 TP, and 0.5 µL OneTaq® DNA polymerase. For introduction of internal modification sites, add 1 µL MgCl<sub>2</sub> (see **Note 5**), 1 µL of each dNaM-modified antisense oligonucleotide, 0.5 µL unmodified sense oligonucleotide, and 70.5 µL of ultrapure water. For introduction of terminal modification sites, add 1 µL d5SICS-modified forward primer, 1 µL of dNaM-modified reverse primer, 1 µL template plasmid DNA, and 71 µL of ultrapure water. Add the polymerase to the PCR mix last and pipet gently up and down for mixing. Spin down briefly and place PCR tube in the thermal cycler.
3. For introduction of terminal modification sites, run the following PCR

program: 2 min at 94 °C for initial denaturation, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 45–68 °C depending on the melting temperatures of the unnatural base modified primers (see **Notes 15** and **16**) for 40 s and elongation at 68 °C for 1 min, followed by a final elongation step at 68 °C for 3 min.

4. For introduction of internal modification sites, a two-step PCR approach is performed. Run the following program as first PCR step: 2 min at 95 °C for initial denaturation, followed by five cycles of denaturation at 95 °C for 15 s, annealing at 45–68 °C depending on the melting temperatures of the overlapping sequences of the designed oligonucleotides (see **Note 16**) for 20 s and elongation at 72 °C for 1 min.
5. Take 3 µL of the obtained dsDNA from the first PCR step and submit it to a 100 µL PCR reaction mix for amplification during the second PCR step. Add 20 µL 5× OneTaq® Standard Reaction Buffer, 1.2 µL MgCl<sub>2</sub> (see **Note 5**), 2 µL forward primer, 2 µL reverse primer, 1.5 µL canonical dNTP mix, 2 µL dNaM TP, 2 µL dTPT3 TP, and 0.5 µL OneTaq® DNA polymerase to 65.8 µL of ultrapure water. Add the polymerase last and pipet gently up and down for mixing.
6. Spin down briefly and place PCR tube in the thermal cycler. Run the following program for the second PCR step: 20 cycles of denaturation at 95 °C for 15 s, annealing at 45–68 °C depending on the melting temperatures of the overlapping sequences of the designed oligonucleotides (see **Note 16**) for 20 s and elongation at 72 °C for 1 min. Optionally the PCR program can be repeated to obtain higher amounts of DNA template. Therefore, add another 0.5 µL OneTaq® DNA polymerase afterward and repeat the program.
7. Store crude PCR products at –20 °C or proceed with purification.

### 3.2 Six Letter PCR (for Terminal Modification Sites)

1. Design DNA primers for your sequence of interest. Choose labeling sites close to the 5' and 3' terminal sites of your sequence of interest. Introduce dNaM modifications in the reverse primer and d5SICS

modifications in the forward primer preferably by point mutations of either dA or dT. The DNA template needs to comprise the T7 promoter sequence (see **Note 15**) for subsequent T7 in vitro transcription.

Therefore, include the T7 promoter sequence at the 5'-end of your forward primer as an overhang sequence, if it is not already included in your DNA template sequence. Also, the DNA template should include one to two 2'-O-methoxy-modified nucleotides at the 5'-end of the antisense strand (=template strand) to produce transcripts with a homogeneous 3'-end. Hence, include these modifications at the 5'-end of your reverse primer.

2. Pipet PCR reaction mix into a sterile PCR tube. Prepare a 100 µL PCR mix by adding 20 µL 5× OneTaq® Standard Reaction Buffer, 1.5 µL canonical dNTP mix, 2 µL dNaM TP, 2 µL dTPT3 TP, 1 µL d5SICS-modified forward primer, 1 µL of dNaM-modified reverse primer, 1 µL template plasmid DNA, and 0.5 µL OneTaq® DNA polymerase to 71 µL of ultrapure water. Add the polymerase last and pipet gently up and down for mixing. Spin down briefly and place PCR tube in the thermal cycler.
3. Run the following PCR program: 2 min at 94 °C for initial denaturation, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 45–68 °C depending on the melting temperatures of the unnatural base modified primers (see **Notes 16** and **17**) for 40 s and elongation at 68 °C for 1 min, followed by a final elongation step at 68 °C for 3 min.
4. Store crude PCR products at –20 °C or proceed with purification.

### 3.3 Analysis and Purification of DNA Templates

First, perform an analytical agarose gel electrophoresis to check length and purity of the PCR product. Then purify the dsDNA template using a silica column-based cleanup kit.

1. Prepare a 2% (w/v) agarose solution in 0.5× TBE buffer. Therefore, add 2 g high resolution agarose to 100 mL 0.5× TBE buffer and microwave until the agarose is completely dissolved. Store agarose solution at 60 °C for up to 2 weeks.

2. Cast an agarose gel supplemented with 1 mg/mL EtBr. Add 5  $\mu$ L EtBr to 50 mL agarose solution in a 50 mL conical tube. Mix by inverting the tube several times, pour solution into a gel tray and let set. Transfer the agarose gel to a gel electrophoresis system filled with 0.5 $\times$  TBE running buffer.
3. Take 1–5  $\mu$ L crude PCR product and mix with agarose loading buffer and load onto the agarose gel. Include a lane loaded with suitable DNA ladder for reference.
4. Run agarose gel electrophoresis program, for, 15–20 min at 150 V const.
5. Visualize the PCR product under UV-light irradiation using a gel documentation system.
6. You should observe a clean PCR product of the expected length (*see Note 18*).
7. Clean up the crude PCR product by using the NucleoSpin® Gel and PCR cleanup kit. Proceed according to the manufacturer's protocol for *PCR cleanup*.
8. Determine the concentration of cleaned template DNA by measuring the absorption at 260 nm ( $A_{260}$ ) using a UV-Vis spectrophotometer. Calculate the DNA concentration using the *OligoCalc* software for the respective sequences incorporating canonical bases for modified DNA oligonucleotides.
9. Store dsDNA template at –20 °C or proceed with transcription.

### 3.4 Genetic Alphabet Expansion Transcription of Spin-Labeled RNA

1. Pipet in vitro transcription (IVT) reaction mix into a sterile reaction tube. Prepare an IVT reaction mix by adding 20  $\mu$ L Tris–HCl (*see Note*

- 7), 25  $\mu$ L MgCl<sub>2</sub>, 5  $\mu$ L DTT, 10  $\mu$ L rNTP mix, 8  $\mu$ L rTPT3<sup>NO</sup> TP, 1.2  $\mu$ L RNasin Plus<sup>®</sup>, 0.3  $\mu$ L iPP, and 5  $\mu$ L T7 RNA polymerase. The optimal amount of dsDNA template needs to be determined experimentally and for each individual sequence separately (*see Note 19*). Adjust the IVT reaction mix to a final volume of 100  $\mu$ L with ultrapure water (*see Note 20*). Add the polymerase last to the mix and pipet gently up and down for mixing. Spin down briefly and place reaction tube in a 37 °C heating block.
2. Incubate IVT reaction(s) at 37 °C for 4 h (*see Note 21*).
  3. Add 11  $\mu$ L DNase I reaction buffer to the crude IVT reaction and mix by vortexing. Spin down briefly.
  4. Add 2  $\mu$ L DNase I and pipet gently up and down for mixing. Spin down briefly and place reaction tube back in a 37 °C heating block.
  5. Incubate DNase digestion reaction at 37 °C for 30 min.
  6. Store crude RNA IVT product at -20 °C or proceed with purification.

### 3.5 Purification of Spin-Labeled RNA

First, perform preparative agarose gel electrophoresis to purify the spin-labeled RNA from truncated fragments of the IVT reaction, then extract and purify the spin-labeled RNA using a silica column-based cleanup kit.

1. Prepare a 2% (w/v) agarose solution in 0.5× TBE buffer. Therefore, add 2 g High Resolution Agarose to 100 mL 0.5× TBE buffer and microwave until the agarose is completely dissolved. The agarose solution can be stored at 60 °C for up to 2 weeks.
2. Cast an agarose gel supplemented with 1 mg/mL EtBr. Add 5  $\mu$ L EtBr to 50 mL agarose solution in a 50 mL conical tube. Mix by inverting the tube several times, pour solution into a gel tray and let set. Use a preparative comb to generate a broad pocket for higher sample volumes. Transfer the agarose gel to a gel electrophoresis system filled with 0.5× TBE running buffer.

3. Mix crude IVT product with agarose loading buffer and load onto the agarose gel. Include a lane loaded with suitable single-stranded RNA ladder for reference.
4. Run agarose gel electrophoresis program, for 15–20 min at 150 V const (see Note 22).
5. Visualize the IVT product under UV-light irradiation using a gel documentation system (see Note 23). Excise the spin-labeled RNA band from the agarose gel.
6. Extract and purify the spin-labeled RNA by using the Nucleo-Spin® Gel and PCR cleanup kit. Proceed according to the manufacturer's protocol for *RNA extraction from agarose gels (Buffer NTC)* (see Note 12). Notwithstanding, perform an additional third washing step with NT3 buffer to improve RNA purity during the cleanup process. Elute the RNA in two steps using 25 µL ultrapure water each to improve the yield of recovered RNA (see Note 24).
7. Determine the concentration of cleaned spin-labeled RNA by measuring the absorption at 260 nm ( $A_{260}$ ) using a UV-Vis spectrophotometer. Calculate the RNA concentration using the *OligoCalc* software for the respective sequences incorporating canonical bases for modified DNA oligonucleotides.
8. Store the spin-labeled RNA at –20 °C or proceed with sample preparation for EPR.

## 3.6 Sample Preparation for EPR Measurements

To prepare spin-labeled RNA samples for EPR measurements, the RNA needs to be dissolved in phosphate-buffered D<sub>2</sub>O. Therefore, the RNA can either be freeze-dried and dissolved in D<sub>2</sub>O buffer afterward or kept solubilized applying a buffer exchange using Amicon® Ultra Centrifugal Filters. The latter approach may be preferred for more sensitive RNA and samples with higher volumes.

### **3.6.1 Preparation by Lyophilization**

1. Add 0.5 µL MgCl<sub>2</sub> in D<sub>2</sub>O to the cleaned RNA sample.
2. Freeze-dry the RNA sample (*see Note 25*).
3. Resolve the RNA in 50 µL PELDOR buffer. Mix by vortexing until fully dissolved.
4. Hybridize the RNA by heating to 70 °C for 5 min, followed by a gradient 70 → 18 °C with a cooling rate of 2 °C min<sup>-1</sup>.
5. Add 12.5 µL ethylene glycol-d<sub>6</sub>. Mix gently by pipetting up and down several times, spin down briefly.
6. Store spin-labeled RNA at -20 °C until EPR measurements.

### **3.6.2 Preparation by Buffer Exchange**

1. Concentrate your RNA samples(s) using an Amicon® Ultra 3K Centrifugal Filter according to the manufacturer's protocol (*see Note 26*). Spin for 30 min.
2. Perform buffer exchange by adding 450 µL PELDOR buffer to the RNA concentrate and repeat concentration step (*see Note 26*). Spin for 30 min.
3. Repeat **step 2**.
4. Recover the RNA concentrate according to the manufacturer's protocol.
5. Transfer the RNA concentrate to a fresh reaction tube.
6. Add 0.5 µL MgCl<sub>2</sub> in D<sub>2</sub>O to the RNA concentrate.
7. Adjust the final sample volume to 50 µL by adding PELDOR buffer.

8. Hybridize the RNA by heating to 70 °C for 5 min, followed by a gradient 70 → 18 °C with a cooling rate of 2 °C min<sup>-1</sup>.
9. Add 12.5 µL ethylene glycol-d<sub>6</sub>. Mix gently by pipetting up and down several times, spin down briefly.
10. Store spin-labeled RNA at -20 °C until EPR measurements.

Prior to EPR studies, the incorporation efficiency of TPT3<sup>NO</sup> TP during in vitro transcription, purity, and correct folding should be determined for each RNA sample using standard methods including HPLC-ESI-MS, room temperature continuous-wave electron paramagnetic resonance (cw-EPR) spectroscopic analysis, analytical denaturing and native gel electrophoresis, UV melting experiments, and circular dichroism (CD) spectroscopy.

---

## 4 Notes

1. Please note that currently only d5SICS cyanoethyl phosphoramidite, as part of unnatural base pair dNaM:d5SICS, is commercially available [40]. dTPT3 cyanoethyl phosphoramidite is currently not commercially available. For our studies we applied a d5SICS-modified forward primer but used dTPT3 TP in the PCR reaction mix.
2. Solid phase DNA synthesis and purification of dNaM- and d5SICS-modified DNA primers was performed by Ella Biotech, Germany. Syntheses for modified primers were done in 200 nmol scale.
3. Solid phase DNA synthesis and purification of unmodified DNA primers was performed by Biomers.net, Germany.
4. When designing reverse primers for PCR, 2'-O-methoxy-modifications (2'-OMe) of the final two nucleotides at the primer's 5'-end are recommended. This generates a DNA template strand with two 2'-OMe modified nucleotides at the 5'-end and promotes the formation of homogeneous 3'-ends in the transcript avoiding unspecific 3'-extension during in vitro transcription.  
Depending on the convenience addition of extra MgCl<sub>2</sub> with a final

5. Depending on the sequence, addition of extra MgCl<sub>2</sub> with a final concentration of 1 mM to the PCR reaction mix can increase the yield. The MgCl<sub>2</sub> is supplemented additionally to MgCl<sub>2</sub> already included in the OneTaq® Standard Buffer and can improve the polymerase's activity.
6. If T7 in vitro transcription from a single-stranded DNA template is performed, a short complementary oligonucleotide (T7 primer) needs to be annealed to the template strand generating a double-stranded T7 promoter region necessary for T7 RNA polymerase to initiate transcription. Annealing can be performed by heating template and T7 primer in Tris-HCl buffer containing MgCl<sub>2</sub> to 95 °C, followed by cooling to 4 °C with a cooling rate of 5 °C min<sup>-1</sup>.
7. When preparing catalytically active glmS ribozymes, note that Tris buffer leads to extensive self-cleavage during in vitro transcription. As the glmS ribozyme [41] requires glucosamine-6-phosphate (GlcN6P) for its activity with its amino group acting in general acid-base catalysis, the presence of Tris would partially activate the self-cleaving mechanism during in vitro transcription. Therefore, HEPES buffer at pH 7.9 (100 mM) was used instead of Tris-HCl for in vitro transcriptions of the glmS ribozyme. Likewise, HEPES buffer at pH 7.6 supplemented with 25 mM MgCl<sub>2</sub> and 5 mM CaCl<sub>2</sub> was employed for subsequent DNase digestion.
8. Depending on the sequence, we found it sometimes helpful to add spermidine with a final concentration of 2 mM to the in vitro transcription mix. This may increase the yield of transcribed RNA.
9. Gently warm up DTT solution at room temperature. Do not use a heating block to thaw DTT solutions as DTT is heat sensitive.
10. Home-made T7 RNA polymerase was used for in vitro transcription. The amino acid sequence conforms with GenBank [42]: AY264774.1. T7 RNA polymerase is also commercially available from various suppliers.
11. Remember to use the UV safety installation of the gel documentation system or wear a personal UV face shield to protect face and eyes.

System or wear a personal protective face mask to protect face and eyes from UV radiation damage.

12. For purification of RNA and single-stranded DNA using the NucleoSpin® Gel and PCR cleanup kit, the special NTC buffer needs to be purchased from Macherey-Nagel. It has to be used (instead of the NTI buffer for purification of double-stranded DNA which is included in the kit) to avoid loss of RNA or single-stranded DNA during the purification process.
13. Mutating dC or dG is also possible but probably has a stronger effect on structure and stability of both DNA and RNA.
14. Efficient solid phase synthesis of DNA oligonucleotides is limited to an approximate length of 150 nt with reduced yields for extended length.
15. The T7 promoter sequence is 5'-TAA TAC GAC TCA TAG-3'. Transcriptions starts at the underlined G. To improve transcription yields, adding one or two more G at the start of the RNA sequence is recommended.
16. The annealing temperature during PCR cycling should be chosen five degrees lower than the melting temperature(s) of the overlapping oligonucleotide sequences or primer sequences. Design your corresponding oligonucleotide overlapping sequences and primer pair sequences to have the same melting temperature. If this condition cannot be met, keep the melting temperature difference as low as possible and calculate the annealing temperature using the average melting temperature.
17. We observed decreased melting temperatures for unnatural base modified primers. In our experience, point mutations of dA or dT with either dNaM or d5SICS within the primer sequences will reduce the melting temperature by approximately 5 °C per single mutation in comparison to the unmodified, fully complementary primers. A temperature gradient PCR in advance is recommended to determine an optimal annealing temperature for the PCR reaction.

18. If PCR side products or other impurities are observed, try improving the PCR reaction mix and cycling conditions. If a PCR product of the wrong length is observed or no PCR product is obtained, reassess your oligonucleotide and primer design.
19. For in vitro transcriptions of the 185 nt long RNA construct with internal modifications a final concentration of 150 nM dsDNA template was used, for the 377 nt long RNA construct with terminal modifications 5 µg/mL dsDNA template was employed.
20. Depending on the sequence, adding spermidine with a final concentration of 2 mM to the IVT reaction mix can increase the RNA yield.
21. Depending on the RNA sequence, the yield of a 100 µL IVT reaction may vary. Considering a loss of RNA during subsequent cleanup, it may be necessary to run multiple 100 µL IVT reaction mixes in parallel to achieve sufficient RNA amounts for EPR measurements. We do not recommend to increase the volume of a single IVT mix beyond 100 µL.
22. Avoid higher voltage during gel electrophoresis to avoid degradation of the spin-labels in RNA. For purification of spin-labeled RNA, we found that agarose gel electrophoresis is preferred over polyacrylamide gel electrophoresis (PAGE) with respect to preserve the nitroxide spin-labels.
23. Keep the exposure time as short as possible as the UV radiation may cause damage to the RNA.
24. When using several columns during RNA purification to obtain sufficient yields of spin-labeled RNA, an additional centrifugation step is recommended to pellet silica carryover. Perform RNA extraction and purification according to the protocol, combine the eluates and centrifuge at full speed for 10 min. Take the supernatant and discard the pellet.

Before connecting your sample to the freeze-dryer freezing the RNA

- ~~25. Before connecting your sample to the freeze dryer, freezing the RNA sample in liquid nitrogen first is recommended. Therefore, prepare a 1.5 mL reaction tube with a pierced cap by pricking holes into the cap using a small canula. Transfer your RNA sample into the reaction tube with pierced cap, place it in a floating foam tube rack and freeze it in a dewar flask filled with liquid nitrogen.~~
26. Do not yet perform the final step to recover the concentrated RNA solution, keep the concentrate within the Amicon® Ultra device. Discard the filtrate(s) during concentration, place the Amicon® Ultra device with concentrated RNA solution back to the microcentrifuge tube.

## Acknowledgments

This work has received financial support from the Boehringer Ingelheim Stiftung and the Deutsche Forschungsgemeinschaft (KA 3699/6-1).

---

## References

1. Kamble NR, Granz M, Prisner TF, Sigurdsson ST (2016) Noncovalent and site-directed spin labeling of duplex RNA. *Chem Commun (Camb)* 52(100):14442–14445. <https://doi.org/10.1039/c6cc08387k> [Crossref]
2. Endeward B, Marko A, Denysenkov VP, Sigurdsson ST, Prisner TF (2015) Advanced EPR methods for studying conformational dynamics of nucleic acids. *Methods Enzymol* 564:403–425. <https://doi.org/10.1016/bs.mie.2015.06.007> [Crossref][PubMed]
3. Krstic I, Hansel R, Romainczyk O, Engels JW, Dotsch V, Prisner TF (2011) Long-range distance measurements on nucleic acids in cells by pulsed EPR spectroscopy. *Angew Chem Int Ed Engl* 50(22):5070–5074. <https://doi.org/10.1002/anie.201100886> [Crossref][PubMed]
4. Schiemann O, Cekan P, Margraf D, Prisner TF, Sigurdsson ST (2009) Relative orientation of rigid nitroxides by PELDOR: beyond distance measurements in nucleic acids. *Angew Chem Int Ed Engl* 48(18):3292–3295. <https://doi.org/10.1002/anie.200805152> [Crossref][PubMed]
5. Schiemann O, Piton N, Plackmeyer J, Bode BE, Prisner TF, Engels JW (2007) Spin

- labeling of oligonucleotides with the nitroxide TPA and use of PELDOR, a pulse EPR method, to measure intramolecular distances. *Nat Protoc* 2(4):904–923. <https://doi.org/10.1038/nprot.2007.97>  
[Crossref][PubMed]
6. Piton N, Mu Y, Stock G, Prisner TF, Schiemann O, Engels JW (2007) Base-specific spin-labeling of RNA for structure determination. *Nucleic Acids Res* 35(9):3128–3143. <https://doi.org/10.1093/nar/gkm169>  
[Crossref][PubMed][PubMedCentral]
  7. Piton N, Schiemann O, Mu Y, Stock G, Prisner T, Engels JW (2005) Synthesis of spin-labeled RNAs for long range distance measurements by peldor. *Nucleosides Nucleotides Nucleic Acids* 24(5–7):771–775  
[Crossref]
  8. Schiemann O, Weber A, Edwards TE, Prisner TF, Sigurdsson ST (2003) Nanometer distance measurements on RNA using PELDOR. *J Am Chem Soc* 125(12):3434–3435. <https://doi.org/10.1021/ja0274610>  
[Crossref][PubMed]
  9. Milov A, Salikhov K, Shirov M (1981) Application of ELDOR in electron-spin echo for paramagnetic center space distribution in solids. *Fizika Tverdogo Tela* 23(4):975–982
  10. Martin RE, Pannier M, Diederich F, Gramlich V, Hubrich M, Spiess HW (1998) Determination of end-to-end distances in a series of TEMPO Diradicals of up to 2.8 nm length with a new four-pulse double electron electron resonance experiment. *Angew Chem Int Ed* 37(20):2833–2837. [https://doi.org/10.1002/\(SICI\)1521-3773\(19981102\)37:20<2833::AID-ANIE2833>3.0.CO;2-7](https://doi.org/10.1002/(SICI)1521-3773(19981102)37:20<2833::AID-ANIE2833>3.0.CO;2-7)  
[Crossref]
  11. Domnick C, Hagelueken G, Eggert F, Schiemann O, Kath-Schorr S (2019) Posttranscriptional spin labeling of RNA by tetrazine-based cycloaddition. *Org Biomol Chem* 17(7):1805–1808. <https://doi.org/10.1039/c8ob02597e>  
[Crossref][PubMed]
  12. Wuebben C, Blume S, Abdullin D, Brajtenbach D, Haege F, Kath-Schorr S, Schiemann O (2019) Site-directed spin Labeling of RNA with a gem-Diethylisoindoline spin label: PELDOR, relaxation, and reduction stability. *Molecules* 24(24):4482. <https://doi.org/10.3390/molecules24244482>  
[Crossref][PubMedCentral]
  13. Socoli G, Wachowius F, Bennati M, Hobartner C (2010) Probing secondary structures of spin-labeled RNA by pulsed EPR spectroscopy. *Angew Chem Int Ed Engl*

49(36):6443–6447. <https://doi.org/10.1002/anie.201000713>  
[Crossref][PubMed]

14. Babaylova ES, Ivanov AV, Malygin AA, Vorobjeva MA, Venyaminova AG, Polienko YF, Kirilyuk IA, Krumkacheva OA, Fedin MV, Karpova GG, Bagryanskaya EG (2014) A versatile approach for site-directed spin labeling and structural EPR studies of RNAs. *Org Biomol Chem* 12(19):3129–3136. <https://doi.org/10.1039/C3OB42154F>  
[Crossref][PubMed]
15. Kerzhner M, Abdullin D, Wiecek J, Matsuoka H, Hagelueken G, Schiemann O, Famulok M (2016) Post-synthetic spin-Labeling of RNA through click chemistry for PELDOR measurements. *Chemistry* 22(34):12113–12121. <https://doi.org/10.1002/chem.201601897>  
[Crossref][PubMed]
16. Weinrich T, Jaumann EA, Scheffer U, Prisner TF, Gobel MW (2018) A cytidine Phosphoramidite with protected Nitroxide spin label: synthesis of a full-length TAR RNA and investigation by in-line probing and EPR spectroscopy. *Chemistry* 24(23):6202–6207. <https://doi.org/10.1002/chem.201800167>  
[Crossref][PubMed]
17. Wachowius F, Höbartner C (2010) Chemical RNA modifications for studies of RNA structure and dynamics. *Chembiochem* 11(4):469–480. <https://doi.org/10.1002/cbic.200900697>  
[Crossref][PubMed]
18. Esquiaqui JM, Sherman EM, Ye JD, Fanucci GE (2014) Site-directed spin-labeling strategies and electron paramagnetic resonance spectroscopy for large riboswitches. *Methods Enzymol* 549:287–311. <https://doi.org/10.1016/B978-0-12-801122-5.00013-1>  
[Crossref][PubMed]
19. Esquiaqui JM, Sherman EM, Ye JD, Fanucci GE (2016) Conformational flexibility and dynamics of the internal loop and helical regions of the kink-turn motif in the glycine riboswitch by site-directed spin-Labeling. *Biochemistry* 55(31):4295–4305. <https://doi.org/10.1021/acs.biochem.6b00287>  
[Crossref][PubMed]
20. Duss O, Yulikov M, Jeschke G, Allain FH (2014) EPR-aided approach for solution structure determination of large RNAs or protein-RNA complexes. *Nat Commun* 5:3669. <https://doi.org/10.1038/ncomms4669>  
[Crossref][PubMed]
21. Zampetaki A, Albrecht A, Steinhofel K (2018) Long non-coding RNA structure and function: is there a link? *Front Physiol* 9:1201. <https://doi.org/10.3389/fphys.2018.01201>.

01201

[Crossref][PubMed][PubMedCentral]

22. Domnick C, Eggert F, Wuebben C, Bornewasser L, Hagelueken G, Schiemann O, Kath-Schorr S (2020) EPR distance measurements on long non-coding RNAs empowered by genetic alphabet expansion transcription. *Angew Chem Int Ed Engl* 59(20):7891–7896. <https://doi.org/10.1002/anie.201916447>  
[Crossref][PubMed][PubMedCentral]
23. Malyshев DA, Romesberg FE (2015) The expanded genetic alphabet. *Angew Chem Int Ed Engl* 54(41):11930–11944. <https://doi.org/10.1002/anie.201502890>  
[Crossref][PubMed][PubMedCentral]
24. Seo YJ, Matsuda S, Romesberg FE (2009) Transcription of an expanded genetic alphabet. *J Am Chem Soc* 131(14):5046–5047. <https://doi.org/10.1021/ja9006996>  
[Crossref][PubMed][PubMedCentral]
25. Malyshев DA, Seo YJ, Ordoukhalian P, Romesberg FE (2009) PCR with an expanded genetic alphabet. *J Am Chem Soc* 131(41):14620–14621. <https://doi.org/10.1021/ja906186f>  
[Crossref][PubMed][PubMedCentral]
26. Leconte AM, Romesberg FE (2006) Amplify this! DNA and RNA get a third base pair. *Nat Methods* 3(9):667–668. <https://doi.org/10.1038/nmeth0906-667>  
[Crossref][PubMed]
27. Hirao I, Mitsui T, Kimoto M, Yokoyama S (2007) An efficient unnatural base pair for PCR amplification. *J Am Chem Soc* 129(50):15549–15555. <https://doi.org/10.1021/ja073830m>  
[Crossref][PubMed]
28. Mitsui T, Kimoto M, Harada Y, Yokoyama S, Hirao I (2005) An efficient unnatural base pair for a base-pair-expanded transcription system. *J Am Chem Soc* 127(24):8652–8658. <https://doi.org/10.1021/ja0425280>  
[Crossref][PubMed]
29. Hirao I, Harada Y, Kimoto M, Mitsui T, Fujiwara T, Yokoyama S (2004) A two-unnatural-base-pair system toward the expansion of the genetic code. *J Am Chem Soc* 126(41):13298–13305. <https://doi.org/10.1021/ja047201d>  
[Crossref][PubMed]
30. Li L, Degardin M, Lavergne T, Malyshев DA, Dhami K, Ordoukhalian P, Romesberg FE (2014) Natural-like replication of an unnatural base pair for the expansion of the

- genetic alphabet and biotechnology applications. *J Am Chem Soc* 136(3):826–829. <https://doi.org/10.1021/ja408814g>  
[Crossref][PubMed]
31. Lavergne T, Lamichhane R, Malyshev DA, Li Z, Li L, Sperling E, Williamson JR, Millar DP, Romesberg FE (2016) FRET characterization of complex conformational changes in a large 16S ribosomal RNA fragment site-specifically Labeled using Unnatural Base pairs. *ACS Chem Biol* 11(5):1347–1353. <https://doi.org/10.1021/acscchembio.5b00952>  
[Crossref][PubMed][PubMedCentral]
  32. Seo YJ, Malyshev DA, Lavergne T, Ordoukhalian P, Romesberg FE (2011) Site-specific labeling of DNA and RNA using an efficiently replicated and transcribed class of unnatural base pairs. *J Am Chem Soc* 133(49):19878–19888. <https://doi.org/10.1021/ja207907d>  
[Crossref][PubMed][PubMedCentral]
  33. Someya T, Ando A, Kimoto M, Hirao I (2015) Site-specific labeling of RNA by combining genetic alphabet expansion transcription and copper-free click chemistry. *Nucleic Acids Res* 43(14):6665–6676. <https://doi.org/10.1093/nar/gkv638>  
[Crossref][PubMed][PubMedCentral]
  34. Ishizuka T, Kimoto M, Sato A, Hirao I (2012) Site-specific functionalization of RNA molecules by an unnatural base pair transcription system via click chemistry. *Chem Commun (Camb)* 48(88):10835–10837. <https://doi.org/10.1039/c2cc36293g>  
[Crossref]
  35. Kimoto M, Sato A, Kawai R, Yokoyama S, Hirao I (2009) Site-specific incorporation of functional components into RNA by transcription using unnatural base pair systems. *Nucleic Acids Symp Ser (Oxf)* 53:73–74. <https://doi.org/10.1093/nass/nrp037>  
[Crossref]
  36. Moriyama K, Kimoto M, Mitsui T, Yokoyama S, Hirao I (2005) Site-specific biotinylation of RNA molecules by transcription using unnatural base pairs. *Nucleic Acids Res* 33(15):e129. <https://doi.org/10.1093/nar/gni128>  
[Crossref][PubMed][PubMedCentral]
  37. Eggert F, Kulikov K, Domnick C, Leifels P, Kath-Schorr S (2017) Illuminated by foreign letters - strategies for site-specific cyclopropene modification of large functional RNAs via in vitro transcription. *Methods* 120:17–27. <https://doi.org/10.1016/j.ymeth.2017.04.021>  
[Crossref][PubMed]
  38. Eggert F, Kath-Schorr S (2016) A cyclopropene-modified nucleotide for site-specific

- RNA labeling using genetic alphabet expansion transcription. *Chem Commun (Camb)* 52(45):7284–7287. <https://doi.org/10.1039/c6cc02321e>  
[Crossref]
39. Domnick C, Eggert F, Kath-Schorr S (2015) Site-specific enzymatic introduction of a norbornene modified unnatural base into RNA and application in post-transcriptional labeling. *Chem Commun (Camb)* 51(39):8253–8256. <https://doi.org/10.1039/c5cc01765c>  
[Crossref]
40. Seo YJ, Hwang GT, Ordoukhalian P, Romesberg FE (2009) Optimization of an unnatural base pair toward natural-like replication. *J Am Chem Soc* 131(9):3246–3252. <https://doi.org/10.1021/ja807853m>  
[Crossref][PubMed][PubMedCentral]
41. Cochrane JC, Lipchock SV, Strobel SA (2007) Structural investigation of the GlmS ribozyme bound to its catalytic cofactor. *Chem Biol* 14(1):97–105. <https://doi.org/10.1016/j.chembiol.2006.12.005>  
[Crossref][PubMed]
42. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2005) GenBank. *Nucleic Acids Res* 33(Database issue):D34–D38. <https://doi.org/10.1093/nar/gki063>  
[Crossref][PubMed]
43. Oligonucleotide Calc.: <http://biotools.nubic.northwestern.edu/OligoCalc.html>

## 16. PELDOR Measurements on Nitroxide-Labeled Oligonucleotides

Tobias Hett<sup>1</sup> and Olav Schiemann<sup>1</sup> 

(1) Institute of Physical and Theoretical Chemistry, Rheinische Friedrich-Wilhelms-University, Bonn, Germany

 **Olav Schiemann**  
Email: [schiemann@pc.uni-bonn.de](mailto:schiemann@pc.uni-bonn.de)

### Abstract

In the past decades, pulsed dipolar electron paramagnetic resonance spectroscopy (PDS) has emerged as a powerful tool in biophysical chemistry to study the structure, dynamics, and function of biomolecules like oligonucleotides and proteins. Structural information is obtained from PDS methods in form of a distribution of distances between spin centers. Such spin centers can either be intrinsically present paramagnetic metal ions and organic radicals or may be attached to the biomolecule by means of site-directed spin labeling. The most common PDS experiment for probing interspin distances in the nanometer range is pulsed electron–electron double resonance (PELDOR or DEER). In the protocol presented here, we provide a step-by-step workflow on how to set up a PELDOR experiment on a commercially available pulsed EPR spectrometer, outline the data analysis, and highlight potential pitfalls. We suggest PELDOR measurements on nitroxide-labeled oligonucleotides to study the structure of either RNA-cleaving DNAzymes in complex with their RNA targets or modified DNAzymes with different functions and targets, in which deoxynucleotides are substituted by nitroxide-labeled nucleotides.

**Key words** Double electron–electron resonance (DEER) – Electron paramagnetic resonance (EPR) – Electron spin resonance (ESR) – Pulsed electron–electron double resonance (PELDOR) – Pulsed dipolar spectroscopy (PDS) – Spin label

---

## 1 Introduction

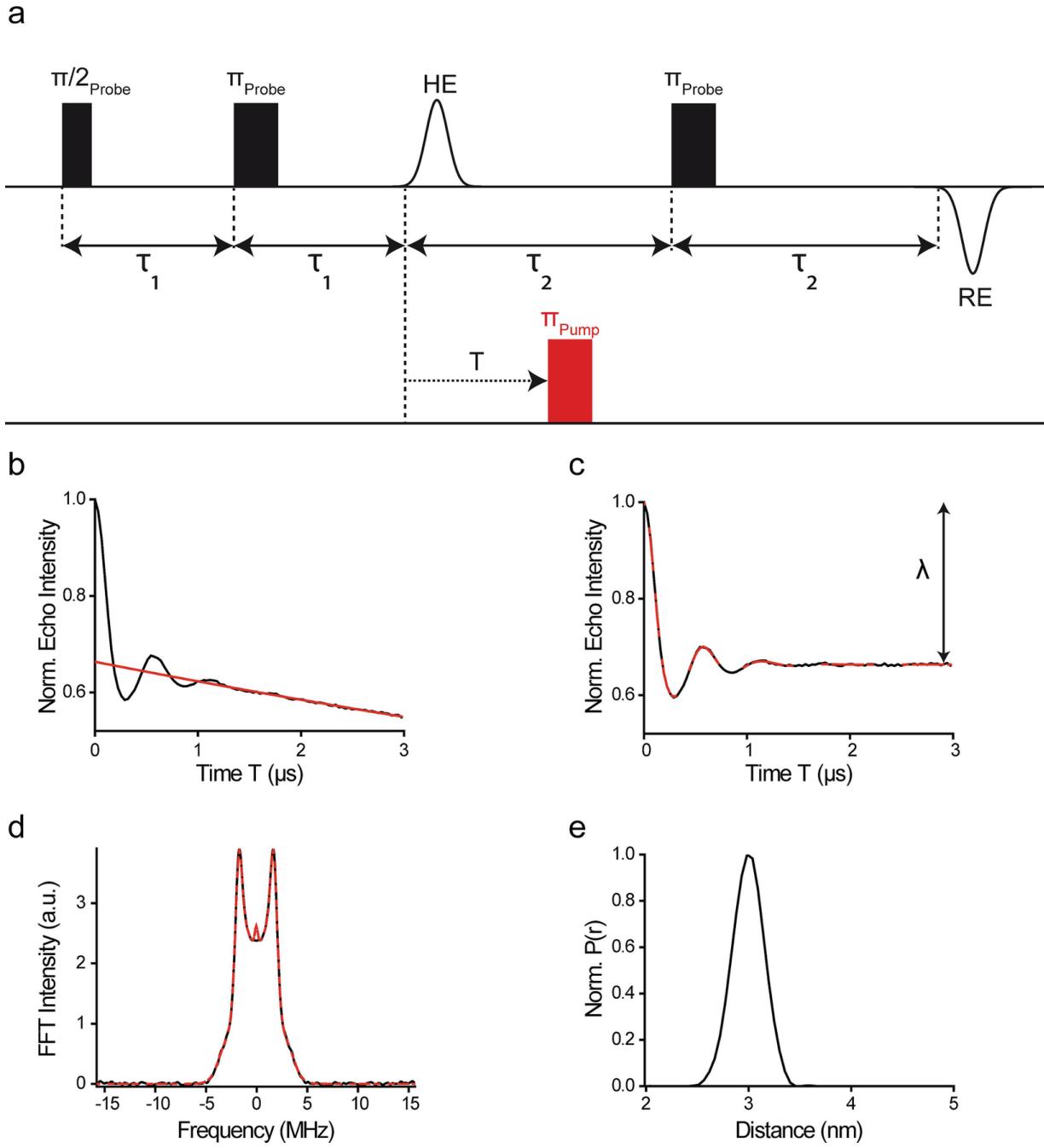
Electron paramagnetic resonance (EPR) spectroscopy provides various methods for studying the structure and dynamics of biomolecules [1, 2]. In order to be applicable, EPR requires the presence of at least one unpaired electron in the biomolecule, for example, in the form of paramagnetic metal ions or cofactor radicals. If the biomolecule does not contain unpaired electrons, these may be introduced via site-directed spin labeling (SDSL) with stable organic radicals [3] like nitroxides [4] and trityls [5–8], or with paramagnetic metal ions like gadolinium(III) [9, 10] or copper(II)-based complexes [11, 12]. EPR-based hyperfine spectroscopy methods allow then to probe the surrounding up to about 8 Å around the electron spin center with atomic resolution [13]. Pulsed dipolar EPR spectroscopy (PDS) provides access to structures on the nanometer range by measuring the distance-dependent dipolar coupling between electron spin centers. The most frequently used PDS method is Pulsed Electron–electron Double Resonance (PELDOR or DEER) [14–16], which has been applied with Angstrom precision in the distance range of 1.5–16 nm [17].

There are several reviews on the theoretical principles, advantages, and limits of PDS spectroscopy in general [18] and PELDOR [16, 19, 20] in particular. Briefly, PELDOR is a pump-probe EPR experiment: In a coupled two-spin system (A–B), a refocused Hahn echo (RE) is recorded on spin A via the probe sequence  $\pi/2-\tau_1-\pi-(\tau_1 + \tau_2)-\pi-\tau_2$ –RE, which is applied at the probe-frequency  $v_{\text{probe}}$  (Fig. 1a). Within the first interval  $\tau_2$ , spin B is flipped by a  $\pi$ -pulse at the pump-frequency  $v_{\text{pump}}$ . Moving the pump pulse within this interval  $\tau_2$  in incremental steps yields a modulated time trace, with the frequencies of this modulation encoding the dipolar electron–electron coupling between spins A and B. Usually, the recorded time trace is directly translated into a probability distribution of interspin distances by computational procedures as included in, for example, DeerAnalysis [21], GLADD/DD [22], or DeerLab [23] (Fig. 1b–e). The resulting distance

distribution can be analyzed with respect to the most probable and mean distance, the distribution width, and the modality of the distribution. With the help of computational methods like, for example, mtsslWizard [24], MMM [25], or the GFN-FF based CREST/MD [26], these distributions can be translated into structural and dynamical information. PELDOR is thus complementary to other biophysical methods like nuclear magnetic resonance (NMR) spectroscopy [27, 28], X-ray crystallography [29, 30], Electron Microscopy (EM) [31, 32], and Förster resonance energy transfer (FRET) spectroscopy [33, 34]. Importantly, PELDOR has the following advantages:

1. It has no restriction with respect to the size of the biomolecule, the largest complex studied by PELDOR is the ribosome [35].
2. In the case of nitroxides, it requires biomolecular concentrations of ~20 μM, but concentrations down to ~45 nM have been measured with PDS [8].
3. It can be applied to the biomolecule free in solution [35–39], in membranes [40, 41], or in whole cells [5, 10].
4. Usually, the measurements are conducted on frozen solutions of the biomolecule, but measurements on immobilized samples at room temperature have been performed [42, 43].
5. The two spin labels do not need to be different, thus orthogonal labeling is obsolete.
6. The labels are small and have short linkers, thus making distance-to-structure translation easier and minimizing the impact of the label on the structure of the biomolecule.
7. The distance distributions are obtained without the need for a reference and with Angstrom precision [44].
8. Beyond distance distributions, PELDOR also yields angular information [45], can be used to count the monomers in multimers [46, 47], and allows spatiotemporal resolution in the Angstrom and microsecond range when coupled with fast freeze-quench techniques

[48].



**Fig. 1** The PELDOR experiment. (a) Pulse sequence of the PELDOR experiment. The probe sequence is shown in black, the pump pulse in red. HE abbreviates the Hahn echo, RE the refocused echo. In the PELDOR experiment, the position of the pump pulse is incremented within the interval  $\tau_2$ , and the integrated intensity of the RE is plotted as a function of the dipolar evolution time  $T$ . (b) Original PELDOR time trace with a background fit indicated as a red line. (c) Background-corrected PELDOR time trace (form factor) with a fit indicated in red. The arrow denotes the modulation depth  $\lambda$ . (d) Fourier

transformation of (**b**) yields the so-called Pake pattern. (**e**) Probability distribution of distances as obtained from transformation of the time trace (**c**)

The protocol presented here guides the reader through the procedure of setting up a PELDOR experiment and highlights the key steps of data transformation from the primary time-domain signal into distance distributions. It focuses on Q-band PELDOR measurements on nitroxide-labeled RNA, for example to investigate RNA–DNAzyme complexes of RNA-cleaving DNAzymes or DNAzymes in which selected deoxynucleotides have been mutated to nitroxide-labeled ribonucleotides, as nitroxides are the most frequently used spin labels and Q-band (34 GHz) the most suitable frequency for a PELDOR experiment on nitroxide-bilabeled biomolecules. General guidelines on good practice regarding the set-up of the PELDOR experiment, the analysis of the time trace, and the interpretation of the distance distribution have also been defined in the white paper of the PELDOR/DEER community [49].

---

## 2 Materials

1. Biomolecule (1.6 nmol) spin-labeled with two nitroxide tags at a sufficiently high labeling efficiency (>80%), corresponding to 80 µL of a 20 µM solution of the biomolecule.
2. Deuterated ethylene glycol (EG-d<sub>6</sub>) or deuterated glycerol (glycerol-d<sub>8</sub>) as cryoprotectant, deuterated water (D<sub>2</sub>O) as solvent.
3. Eppendorf pipette and elongated pipette tips to transfer the sample into the EPR tube.
4. Dewar vessels of different sizes and volumes to shock-freeze the prepared oligonucleotide samples in EPR tubes (for example with 40 mm inner width, 90 mm inner height; 57 mm inner width, 210 mm inner height).
5. 100 L Liquid helium tank.
6. Liquid nitrogen for freezing the sample and liquid helium for cooling the sample in the resonator.

7. Safety goggles and cold protection gloves.
8. EPR sample tubes made from clear fused quartz (CFQ) (Q-band, 3 mm outer diameter (o. d.), 159 mm length).
9. Light duty tissue wipers.
10. Pulsed EPR Spectrometer (e.g., ELEXSYS E580, Bruker, Rheinstetten, Germany) equipped with a Q-band bridge and a 150 W travelling wave tube amplifier (TWT, model 187 Ka from Applied Systems Engineering, Fort Worth, USA), an ER 5106QT-2 resonator (Bruker), and the corresponding accessory devices required for low-temperature operation, that is, a CF935 liquid Helium cryostat (Oxford Instruments, Abingdon, UK), an iTc503S temperature controller (Oxford Instruments), a helium transfer line (Oxford Instruments NanoScience, Abingdon, UK; model LLT600 or LLT650), a turbomolecular pump for evacuating the cryostat (e.g., from Pfeiffer Vacuum, Alslar, Germany; HiCUBE 80 Eco), and a membrane pump (e.g., from KNF Neuberger GmbH, Freiburg, Germany; Type: PM26962-026.1.2) for maintaining a constant stream of cold helium gas.
11. Computer with MATLAB installed and the DeerAnalysis [21] toolbox imported into MATLAB. DeerAnalysis can be downloaded free of charge [50]. For DeerNet [51], installation of the Deep Learning Toolbox and the Signal Processing Toolbox for MATLAB are required.

The reader is advised to refer to the manual of the particular spectrometer, especially with regard to safety of the operator and the instrument. For clarity, references to control buttons and windows in the operating software are highlighted by quotation marks in the text.

---

### 3 Methods

### 3.1 Sample Preparation

Dissolve 1.6 nmol of spin-labeled RNA [52] in 64 µL D<sub>2</sub>O and mix with 16 µL of EG-d<sub>6</sub> or glycerol-d<sub>8</sub> (see **Note 1**). Transfer the 80 µL into a 3 mm o.d. Q-band EPR tube using an Eppendorf pipette with an elongated pipette tip. Freeze the sample by immersing the EPR tube carefully into liquid nitrogen; after freezing, the sample height should be ~1 cm. Safety precaution: Whenever handling EPR tubes with samples frozen in liquid nitrogen, wear safety goggles (see **Note 2**).

### 3.2 Switching on the Spectrometer

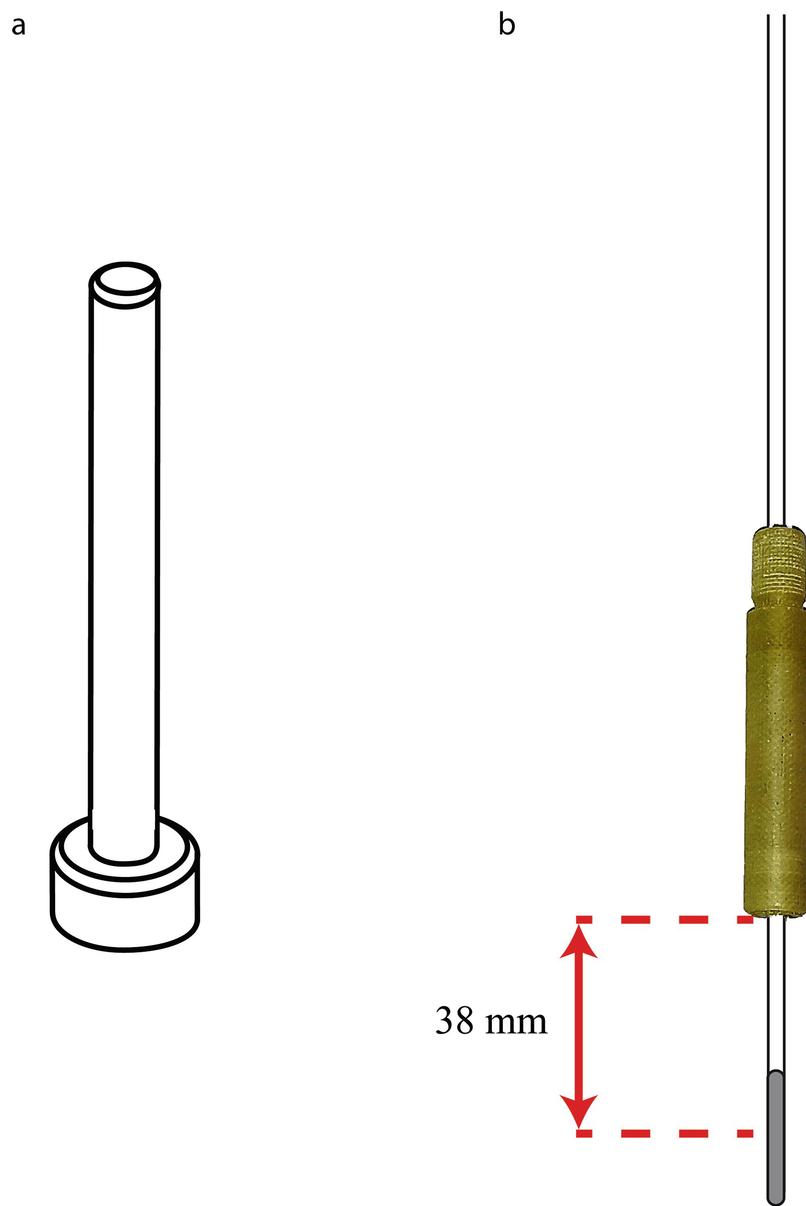
Switch on the heat exchanger of the spectrometer and check the inward and return flow temperature. The former should be around 10–15 °C and the latter should not exceed 25–30 °C. If your spectrometer is equipped with a closed cooling circuit, make sure that there is enough cooling water in the system. Next, switch on the spectrometer console, the magnet power supply, and the traveling-wave tube (TWT) amplifier. Keep the latter in **Standby mode (important!)**. When the console has started up as indicated by a permanent green light of the status LED on the microwave (MW) bridge, connect the operating Xepr software to the spectrometer by selecting the “Connect to Spectrometer” option in the “Acquisition” menu bar. Before starting measurements, the spectrometer, especially the MW bridge, should have warmed up for at least 1 h to ensure sufficient stability of the electric components.

### 3.3 Cooling down the Cryostat and Inserting the Sample

Switch on the temperature controller of the cryostat and set the target temperature to 50 K. Connect the turbomolecular pump system with the cryostat and switch it on. Make sure that the cryostat has reached a pressure of ~10<sup>-4</sup> mbar before proceeding with the following steps.

Connect the overflow valve of the liquid helium tank to the helium recovery system, if available, and make sure that the gas flow within the tubing is not blocked (e.g., due to bending). Ensure that the overflow valve is open so that no overpressure can build up in the tank. Open the needle valve on the helium transfer line and **slowly** (~ 2 min) insert the line into

the helium tank, then fix it with a spanner. Check for gas flow by immersing the outlet of the transfer line into ethanol, wipe it, and connect the transfer line to the cryostat of the spectrometer. Connect the membrane pump with the transfer line. Close the cryostat with a stopper (Fig. 2a) or a sample rod to prevent condensation of air when cooling down. Open the needle valve at the transfer line approximately  $\frac{1}{4}$  turn and switch on the membrane pump. It usually takes up to 20–30 min to cool the cryostat from room temperature to 50 K (see Note 3).



**Fig. 2** Sample holder. (a) Stopper to close the cryostat and thus prevent condensation of air when cooling down. (b) Sample holder with an EPR tube mounted. The red arrows

indicate the distance from the lower end of the holder to the center of the sample (38 mm)

In the meantime, mount the EPR tube into the sample holder and screw the holder into the sample rod. The position of the sample holder is to be adjusted in such a way that the center of the frozen sample is 38 mm [53] below the lower end of the tube holder to ensure optimal positioning of the sample in the resonator (Fig. 2b). Quickly wipe the tube with a light duty tissue wiper to remove potential contaminants from the exterior of the tube. Before inserting the tube into the cryostat, ensure that the high power “Attenuation” is set to 60 dB and that the spectrometer is in the “Standby” mode.

When a temperature of approximately 50 K has been reached, stop the membrane pump and wait until the needle of the gas flow controller (Fig. 3) has dropped to zero. Remove the stopper from the cryostat, gently shake the sample rod to brush off liquid nitrogen, and quickly insert it into the cryostat. Do not keep the cryostat open unnecessarily long to prevent condensation of air in it (see **Note 4**). Give the sample and the cryostat at least 15 min to reach a thermal equilibrium before proceeding. If sample and cryostat are not in equilibrium, frequency and phase drifts might occur during the measurement.



**Fig. 3** Gas flow controller

### 3.4 General Remarks on the Xepr Software

The following section describes the basic functions and windows of the Xepr software used for setting up pulsed EPR experiments. Screenshots of the respective program windows are shown in Figs. 4 and 5.

1. Viewport and main window (Fig. 4a, b): The toolbars in the main window permit calling various functions of the software (e.g., for saving and processing data, “Spectrometer Configuration”) and opening further subwindows (e.g., “MW Bridge Tuning” panel, “FT EPR Parameters”). Figure 4b shows a close-up view of the most important icons used throughout the experiments. In the viewport, the recorded data and the current number of averages are displayed.
2. “FT Bridge” (Fig. 4c): Permits controlling parameters related to the MW Bridge (e.g., “CW Mode” and “Pulse Mode” in the “Bridge Configuration” tab, Fig. 4c-1), parameters regarding the detection system (e.g., “Video Gain,” “Video Bandwidth,” “Attenuation,” “Signal Phase” in the “Receiver Unit” tab, Fig. 4c-2), and phase and amplitude settings of the Microwave Pulse Forming Unit (MPFU) channels (“MPFU Control” tab, Fig. 4c-3).
3. “SpecJet” (Fig. 4d) is a digital oscilloscope displaying the echo(es) generated by the pulse sequence currently executed from the Pulse Tables (see below). The real channel of the quadrature detection system is displayed in green, the imaginary channel in yellow.
4. “FT EPR Parameters” (Fig. 4e): Permits adjusting various spectroscopic variables, for example, the pulse sequence via the Pulse Tables (“Patterns” tab, Fig. 4e-1), the magnetic field (e.g., “Center Field” and “Sweep Width” in the “Field” tab, Fig. 4e-2), parameters regarding the MW bridge (e.g., “Attenuation,” “Video Gain,” “Current ELDOR Frequency,” “ELDOR Attenuation” in the “Microwave” tab, Fig. 4e-3), and interfaces for data acquisition control (e.g., axis quantities and sizes, acquisition from the Pulse Tables and PulseSPEL in the “Acquisition” tab, Fig. 4e-4). It also includes a button for calling PulseSPEL (see below).
5. “Pulse Tables” (Fig. 4e-1): Option for manually programming pulse

sequences by directly providing the pulse lengths, interpulse delays, and incrementation schemes. It is recommended for quickly setting up preliminary experiments, for example, the echo-detected field-swept

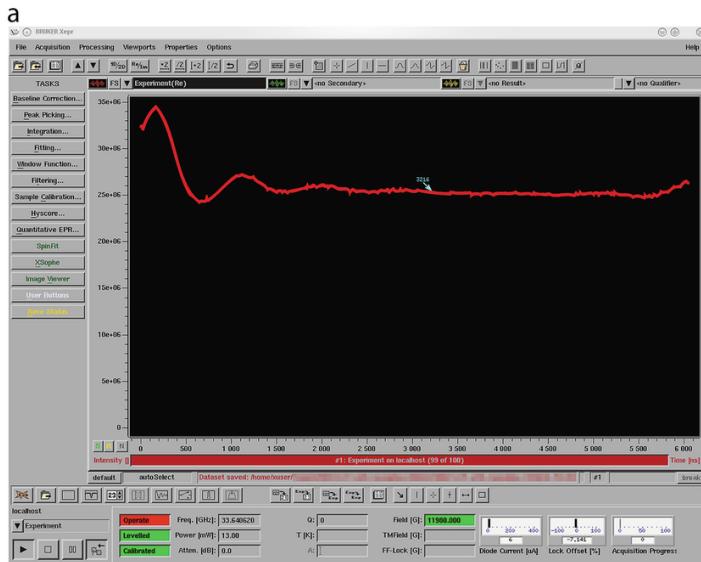
EPR spectrum. The “Pulse Tables” are located in the “Patterns” tab in the “FT EPR Parameters” window (Fig. 4e-1). For experiments conducted using the Pulse Tables, *absolute* timing with respect to the beginning of the whole pulse sequence is used (see **Note 5**).

6. “PulseSPEL” (Bruker **Pulse SPE**ctroscopy Language) is a programming language for interacting with the spectrometer, which permits to easily set up also more sophisticated sequences with numerous pulses. By contrast to the pulse tables, it allows to implement phase cycling and modulation averaging procedures. PulseSPEL variables and programs have to be compiled and validated, respectively, by clicking the “Compile” and “Validate” buttons before execution. Afterward, the experiment can be started by clicking the “Run” button in the main window (Fig. 4b). Changes to PulseSPEL variables can be conveniently made using the variable box (“FT EPR Parameters”/“Acquisition,” Fig. 4e-4) without explicitly compiling them, partially even during the running experiment (e. g. number of scans n).

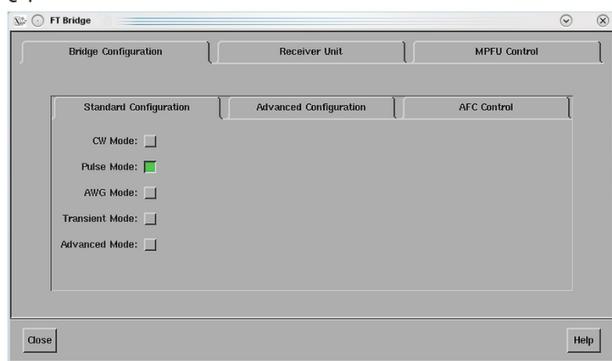
By contrast to the Pulse Tables, PulseSPEL uses *relative* timing of the pulse sequence and refers to the previous interpulse delay in the sequence. For a comparison of the timing in the Pulse Tables and PulseSPEL, refer to **Note 5**.

7. “MW Bridge Tuning” panel (Fig. 5a): Permits setting parameters related to the MW Bridge (e.g., MW frequency and phase) and to set the MW Bridge into “Standby,” “Tune,” and “Operate” mode.

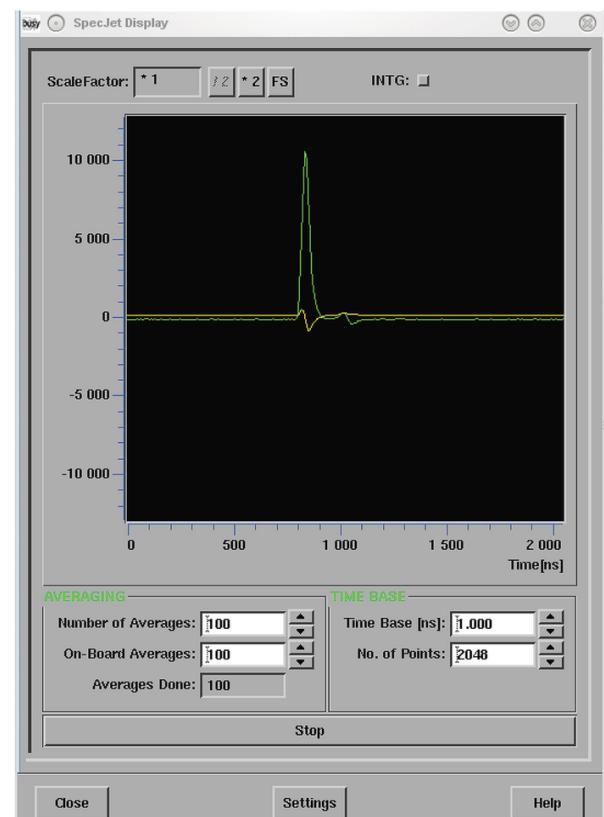




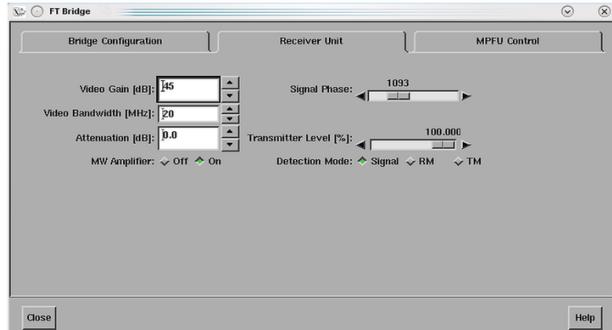
C-1



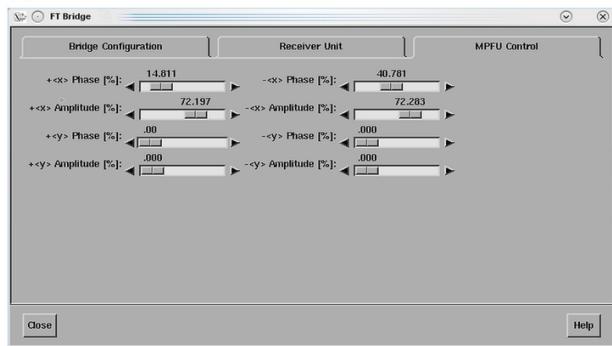
d

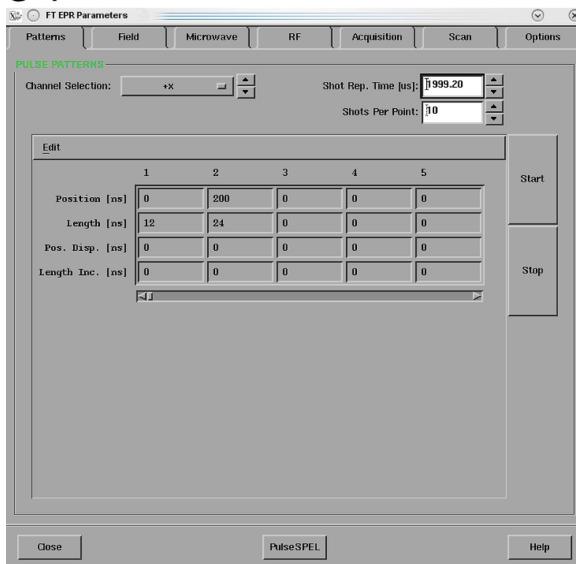
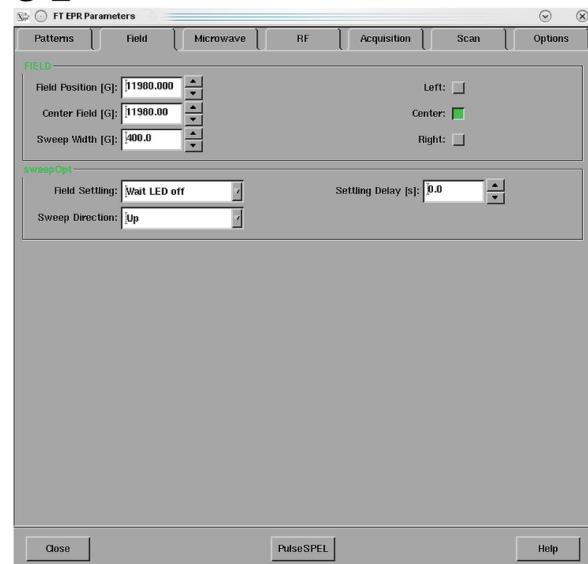
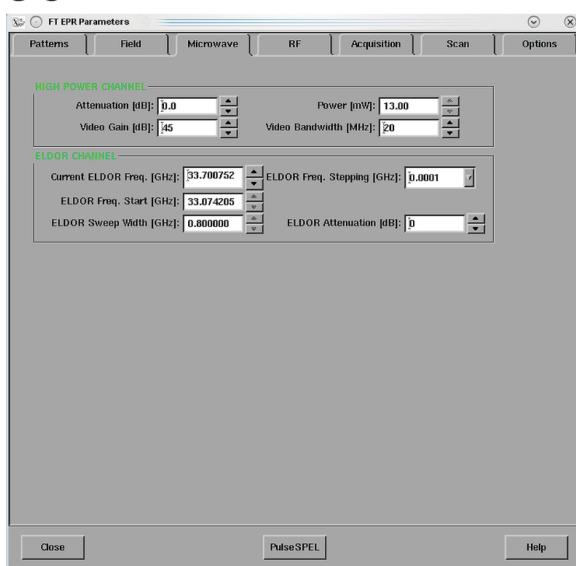
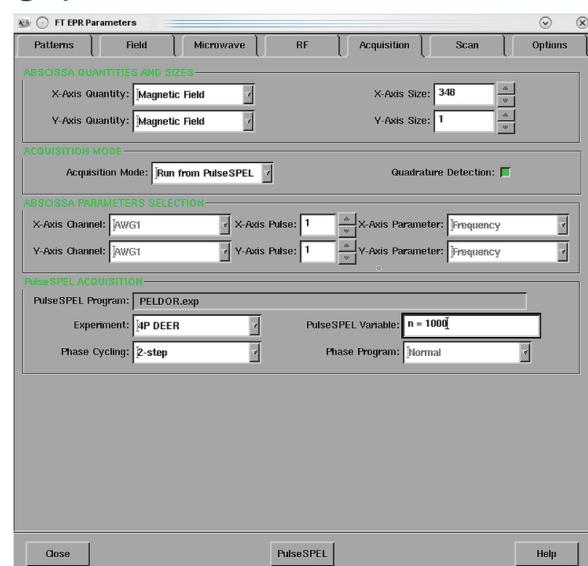


C-2

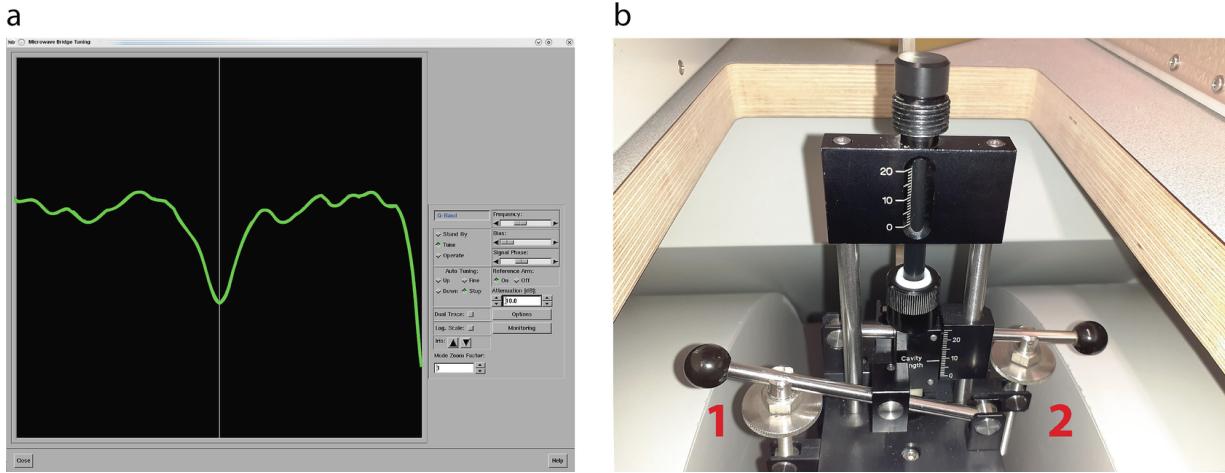


C-3



**e-1****e-2****e-3****e-4**

**Fig. 4** Program windows of the Xepr software. **(a)** Main window with viewport. **(b)** Close-up view of relevant control buttons of the software (top to bottom): “Create new experiment,” “Activate button,” “MW Tuning Panel,” “FT EPR Parameters,” “FT Bridge,” “SpecJet,” “Run button,” “Save Data to Disk.” **(c)** FT Bridge and its subpanels **(c-1)** “Bridge Configuration,” **(c-2)** “Receiver Unit,” **(c-3)** “MPFU Control.” **(d)** SpecJet. **(e)** FT EPR Parameters with its subpanels **(e-1)** “Patterns,” **(e-2)** “Field,” **(e-3)** “Microwave,” **(e-4)** “Acquisition”



**Fig. 5** (a) Screenshot of the MW bridge tuning panel. The dip in the center of the window corresponds to the “real” resonator dip, and the dip on the right is the “fake” dip. (b) Screws to adjust the resonator coupling (1) and the cavity length (2)

### 3.5 Tuning the Spectrometer and Safety Check

Before running PELDOR experiments, the resonator has to be overcoupled such that its bandwidth is large enough to accommodate the pulses of both the pump and the probe frequency. The ER 5106QT-2 resonator used herein has two adjustment arms, one for the coupling and another one for the length of the cavity. By rotating the respective screw on the resonator (Fig. 5b), the cavity coupling (left screw, 1 in Fig. 5b) and the cavity length (right screw, 2 in Fig. 5b) can be altered. The resonance frequency of the ER 5106-QT2 is 34 GHz [53], and it can be fine-adjusted by the cavity length.

Open the “MW Bridge Tuning” panel. Go into the “Tune” mode, set the attenuation to 10 dB, and adjust the MW frequency to 33.7 GHz with the frequency slider. Next, adjust the cavity length (right screw, 2 in Fig. 5b) so that the resonator tuning dip becomes visible in the “MW Bridge Tuning” window (Fig. 5a). Note that there is a “real” and a “fake” dip for the ER 5106QT-2: The “fake” dip can be identified as it moves markedly when the cavity coupling (left screw, 1 in Fig. 5b) is changed, whereas the “real” dip stays in place. Adjust the cavity length so that the “real” dip is centered and the “fake” dip on the right of the tuning window. This corresponds to the overcoupling condition (Fig. 5a). Usually, it is achieved at a cavity length of around 9 mm (*see Note 6*).

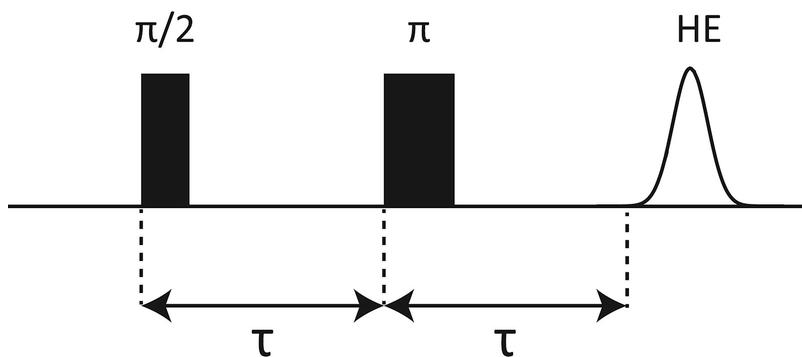
Create a new experiment by clicking the “Exp” button and choose the option “Pulse/Advanced.” Click the “Activate” button and perform the

safety check of the detection system as described in the instrument manual before switching the TWT into the Operate mode. Never (!) attempt to switch the TWT amplifier into the Operate mode without having observed the defense pulses, otherwise the detection system may get damaged. Never return into the “CW” mode without having switched the TWT to Standby. Make sure that the dead time set in the spectrometer configuration files is long enough so that no ringdown of the MW pulses occurs during signal detection. If in doubt, consult the instrument manual or your spectrometer administrator (see Note 7).

## 3.6 Preliminary Experiments

### 3.6.1 Standing Hahn Echo and Echo-Detected Field-Swept EPR Spectrum

Set the “Center Field” to a value corresponding to  $g \approx 2.00$  at the given MW frequency (e.g.,  $\approx 11,980$  G at 33.7 GHz). Program the Hahn echo sequence ( $\pi/2 - \tau - \pi - \tau - \text{echo}$ , Fig. 6) using the Pulse Tables: Choose the “+x” Channel in the “Channel Selection” menu, set a  $\pi/2$ -pulse of 12 ns length at the position 0 ns and a  $\pi$ -pulse of 24 ns length at the position 200 ns. (Fig. 4e-1). Confirm each entry in the Pulse Tables by pressing “enter” on the keyboard.



**Fig. 6** Hahn echo (HE) pulse sequence. For the two-pulse ESEEM experiment, the interpulse delays  $\tau$  are incremented and the HE amplitude is monitored

Next, select the “Acquisition Trigger” channel and adjust the acquisition trigger length to 4 ns at the position 0 ns. Set the “Integrator Time Base” to 1.0 ns, the number of “Shots Per Point” to 10 and the shot repetition time (“Shot Rep. Time”) to 3000  $\mu$ s, which is usually a good starting point for

nitroxides at 50 K. Set the “Video Bandwidth” (“FT Bridge”\ “Receiver Unit”) to 20 MHz. Next, click the “Start” buttons in the Pulse Tables window and in “SpecJet,” and decrease the MW attenuation to observe the Hahn echo. The exact temporal position of the echo depends on the dead time of the spectrometer, refer to the manual of the particular instrument for details (*see Note 5*).

Find the attenuation at which the echo amplitude is maximized and adjust the video gain so that no clipping occurs; for this purpose, set “Number of Averages” = 1 in “SpecJet.” At the microwave power level, which maximizes the echo amplitude, the applied pulses correspond to  $\pi/2$  and  $\pi$ -pulses, which tilt the magnetization vector by  $90^\circ$  and  $180^\circ$ , respectively. If the echo amplitude does not go through a maximum when lowering the attenuation to 0 dB, increase the pulse lengths (e.g.,  $\pi/2 = 16$  ns,  $\pi = 32$  ns; *see Note 8*). Adjust the “Signal Phase” to maximize the amplitude in the real signal channel and to bring it close to zero in the imaginary channel. Slight adjustments of the “Center Field” may be helpful in this step to fully bring the sample on resonance.

Next, record the echo-detected field-swept EPR spectrum, which is the integral of the Hahn echo plotted as a function of the magnetic field strength. In the Pulse Tables window, adjust the “Acquisition Trigger” position so that the echo starts on the left side of the “SpecJet” window and set the “Acquisition Trigger” length such that the whole echo is integrated ( $\sim 120$  ns). For the field-swept EPR experiment, integrating the whole echo is crucial to obtain a distortion-free spectrum and to maximize the spectral resolution [34]. Set the “Sweep Width” to 400 G and the number of points (“FT EPR Parameters”/ “Acquisition”/ “X-Axis Size”) to 800. Press the “Run” button in the main window to record the echo-detected field-swept EPR spectrum and save it to disk. Usually, a small number of scans (e.g.,  $n = 3$ ) is sufficient to obtain a good signal-to-noise ratio (SNR).

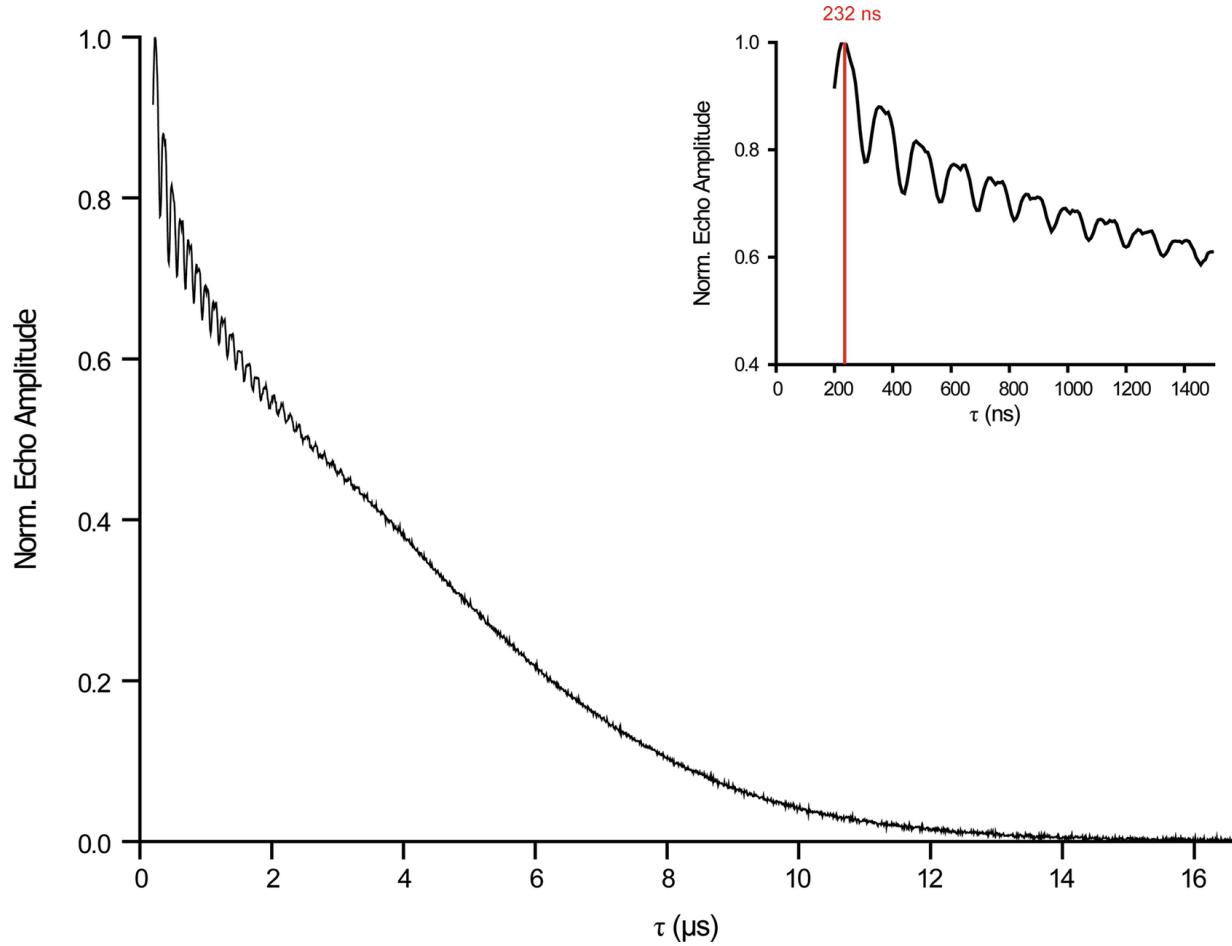
### **3.6.2 Two-Pulse Electron Spin Echo Envelope Modulation (ESEEM)**

The Two-Pulse ESEEM experiment monitors the Hahn echo (Fig. 6) amplitude (y-quantity) as a function of the interpulse delay  $\tau$  (x-quantity) and thus provides information on the transverse electron spin relaxation. Knowledge of this is relevant for choosing a proper value for the dipolar

evolution time window in the PELDOR experiment. Note that the default PulseSPEL program for Two-Pulse ESEEM plots the Hahn echo amplitude as a function of  $\tau$ ; in the literature, however, the echo decay curves are occasionally shown as a function of  $2\tau$ .

Read off the magnetic field position at which the maximal signal intensity has been obtained in the field-swept spectrum and set the “Center Field” to this value. Open PulseSPEL (“FT EPR Parameters”)\“Acquisition”\“PulseSPEL”). Load the PulseSPEL program “2p\_ESEEM.exp” and the corresponding variable definitions “descrESEEM.def” located in the folder xeprFiles/PulseSPEL/sharedPulseSPEL/Standard/Spel2009/ESEEM (see **Note 9**).

After loading the experiment and the variable definitions file, click the buttons “Compile,” “Show Program,” “Validate,” in the “PulseSPEL” window in the given order. Select “Run from PulseSPEL” (“FT EPR Parameters”/“Acquisition”) and the experiment “2P ESE Setup.” Click the “Run” button in the main window to record the Hahn echo. In the viewport, read off the time at which the maximum of the echo occurs and set this value as the instrument-related acquisition delay parameter  $d0$  in PulseSPEL (e.g.,  $d0 = 432$  ns, see **Notes 5** and **9**). Set the number of points for the Two-Pulse ESEEM program (dim2 in the PulseSPEL program) to 1024. Then, select the “2P ESEEM” experiment from the dropdown list with a two-step phase cycle and execute it by clicking “Run” in the main window. If the signal amplitude of the acquired Hahn echo decay curve does not reach the zero level, increase the number of points and/or the time increment  $d30$  in the PulseSPEL program. In order to obtain a sufficient SNR, average multiple scans (e.g.,  $n = 3$ ) and save the trace to disk afterward (Fig. 7).



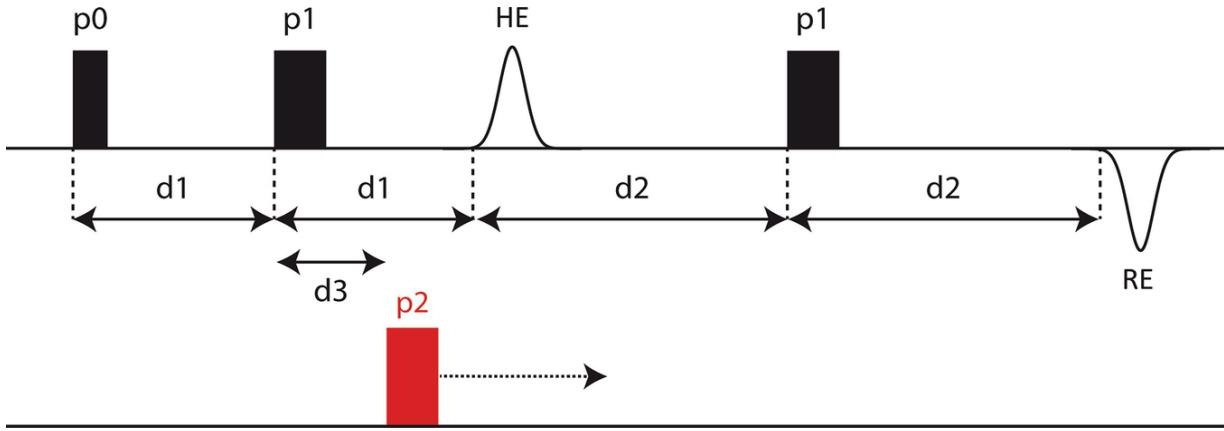
**Fig. 7** Two-pulse electron spin echo envelope modulation (ESEEM) experiment. The Hahn echo amplitude is plotted as a function of the interpulse delay  $\tau$  (Fig. 6). The inset shows a zoom-in on the first 1400 ns of the trace with the maximum at 232 ns indicated by a red line

### 3.7 Setting up the PELDOR Experiment

As the PELDOR experiment requires phase cycling and nuclear modulation averaging to remove unwanted echoes and ESEEM, it is commonly run from a PulseSPEL program. Pulsed EPR spectrometers distributed by Bruker contain a default PulseSPEL program for the PELDOR experiment; however, many research groups modified this program, developing their own syntax and conventions. Herein, we present the PELDOR setup using our modified version of the standard program; for reference, our PELDOR program can be found on GitHub [54].

#### 3.7.1 *Choice of Variables for the Hahn Echo*

Load the PELDOR program and the corresponding variable definitions into PulseSPEL and click the “Compile,” “Show Program,” “Validate” buttons in the given order. For clarity, the variable definitions used in the PulseSPEL program are shown in Fig. 8.



**Fig. 8** Pulse sequence of the PELDOR experiment with the pulses and interpulse delays labeled according to the nomenclature in the “4P DEER” PulseSPEL program. HE abbreviates the Hahn echo, RE the refocused echo. The detection sequence is shown in black, the pump pulse in red. Note that the initial position of the pump pulse is set before the Hahn echo, which permits proper determination of the zero-time in the PELDOR trace. The following correlation between the PulseSPEL variable definitions shown here and the conventional variables applies:  $d_1 = \tau_1$ ;  $d_2 = \tau_2$ ;  $d_3$  = spectrometer dead time delay;  $p_0 = \pi/2_{\text{probe}}$ ;  $p_1 = \pi_{\text{probe}}$ ;  $p_2 = \pi_{\text{pump}}$

For the  $\pi/2$  and  $\pi$ -pulses of the probe sequence (variables  $p_0$  and  $p_1$  in the PulseSPEL program), set the pulse lengths which you identified as optimal for the Hahn echo.

The proper choice of the interpulse delay  $\tau_1$  ( $d_1$  in PulseSPEL) is governed by two aspects: (a) The loss of echo intensity due to transverse relaxation with increasing  $\tau_1$  and (b) the time interval the PELDOR trace is recorded prior to its maximum. In order to fulfil aspect (a), it would be desirable to set  $\tau_1$  as short as possible. However, for clear identification of the maximum in the PELDOR trace, which arises when the positions of the pump pulse and the Hahn echo coincide, the initial position of  $\pi_{\text{pump}}$  is set before the Hahn echo (Fig. 8). As the spectrometer has an instrument-based dead time delay ( $d_3$ , 100 ns in this case) between the first  $\pi_{\text{probe}}$ -pulse and the pump pulse,  $\tau_1$  has a lower limit. In the literature, values of  $\tau_1$  ranging

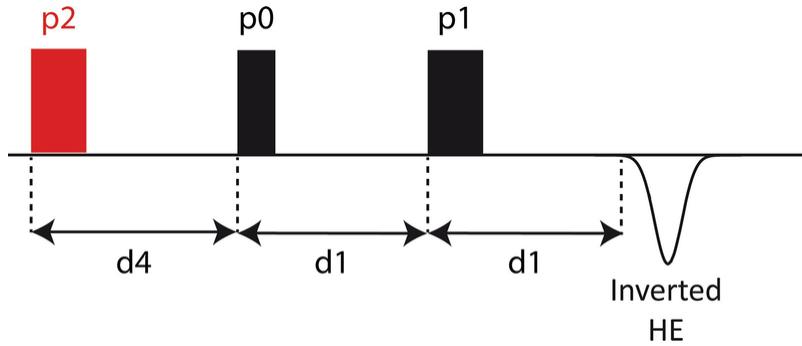
between 200 ns [55] and 400 ns [56] have been reported. As a guideline,  $\tau_1$  may be set to the value which yields the global maximum in the Two-Pulse ESEEM experiment, in this case 232 ns (Fig. 7).

### 3.7.2 Optimization of the Hahn Echo at the Pump Frequency

Set the MW attenuation to 0 dB, select the “2P ESE Setup” program (a PulseSPEL program for the Hahn echo), and press the “Run” button in the main window. Press “Start” in the Pulse Tables window, “Run” in “SpecJet,” and open the “MPFU Control” tab in the “FT Bridge” window. Adjust the “+<x> Amplitude” such that the pulses correspond to  $\pi/2$  and  $\pi$ , thus maximizing the Hahn echo observed in “SpecJet”; with a pulse length of  $\pi/2 = 12$  ns, the channel amplitude slider bar should show a value of around 68%. Use the *global* “Signal Phase” shifter (e.g., in “FT Bridge”/“Receiver Unit”) to adjust the MW phase such that the entire signal is detected in the real channel. The imaginary channel should be zero on average (Fig. 4d).

### 3.7.3 Inversion of the Hahn Echo

Select the “3P ELDOR Setup” experiment (Fig. 9). In this experiment, the inversion pulse  $\pi_{\text{pump}}$  ( $p2$  in the PulseSPEL program) is applied before the Hahn echo sequence, thus inverting the echo amplitude. In “FT EPR Parameters”/“Microwave,” set the “Current ELDOR Frequency,” that is, the frequency of the pump pulse  $v_{\text{pump}}$ , to the current MW frequency (~33.7 GHz). The pump frequency is later needed for computing the probe frequency, thus write it down. Press the “Run” button in the main window. Read off the time position of the echo maximum in the viewport, add it to the current value of  $d0$  in PulseSPEL, and update  $d0$  to this new value, for example, via the “PulseSPEL Variable” box in the “FT EPR Parameters”/“Acquisition” tab. Press “Run” again and check that the echo maximum is located on the left of the viewport. Start “SpecJet” and stepwise decrease the “ELDOR Attenuation” to 0 dB, which should lead to the inversion of the echo amplitude (see **Note 10**).

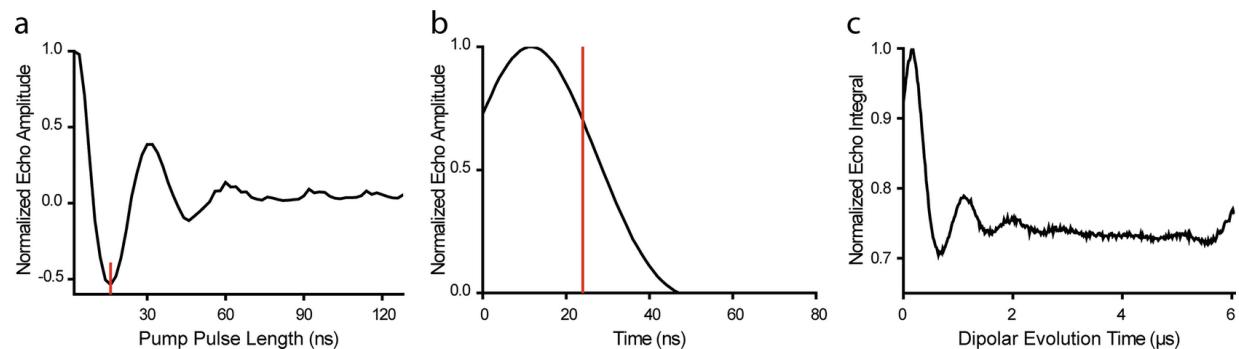


**Fig. 9** Pulse sequence of the Three-Pulse ELDOR Setup experiment. In this experiment, all three pulses are applied at the pump frequency (33.7 GHz). For the Three-Pulse ELDOR Nutation experiment, the length of the pump pulse  $p_2$  is incremented and the amplitude of the inverted HE is monitored. The following correlation of the PulseSPEL variable definitions as shown in the figure and the conventional variables applies:  $d_1 = \tau$ ;  $d_4 = T$ ;  $p_0 = \pi/2_{\text{probe}}$ ;  $p_1 = \pi_{\text{probe}}$ ;  $p_2 = \pi_{\text{pump}}$

### 3.7.4 Optimization of the Pump Pulse Length

Press the “Stop” buttons in the Pulse Tables and in “SpecJet”, select the “3P ELDOR Nutation” experiment from the dropdown menu, and press the “Run” button in the main window. In this experiment, the Hahn echo amplitude is recorded as a function of the pump pulse length  $\pi_{\text{pump}}$  (Fig. 9).

Read off the optimal pump pulse length which permits full inversion of the echo, that is, which yields the global minimum in the nutation trace (Fig. 10a), and set it as variable  $p_2$ . Ideally, this value should be between  $\sim 12$  ns and  $\sim 18$  ns (see Note 11). If two pump pulse lengths lead to a comparable inversion of the echo (e.g., 16 ns and 14 ns), choose the shorter one, which is associated with a larger excitation bandwidth.

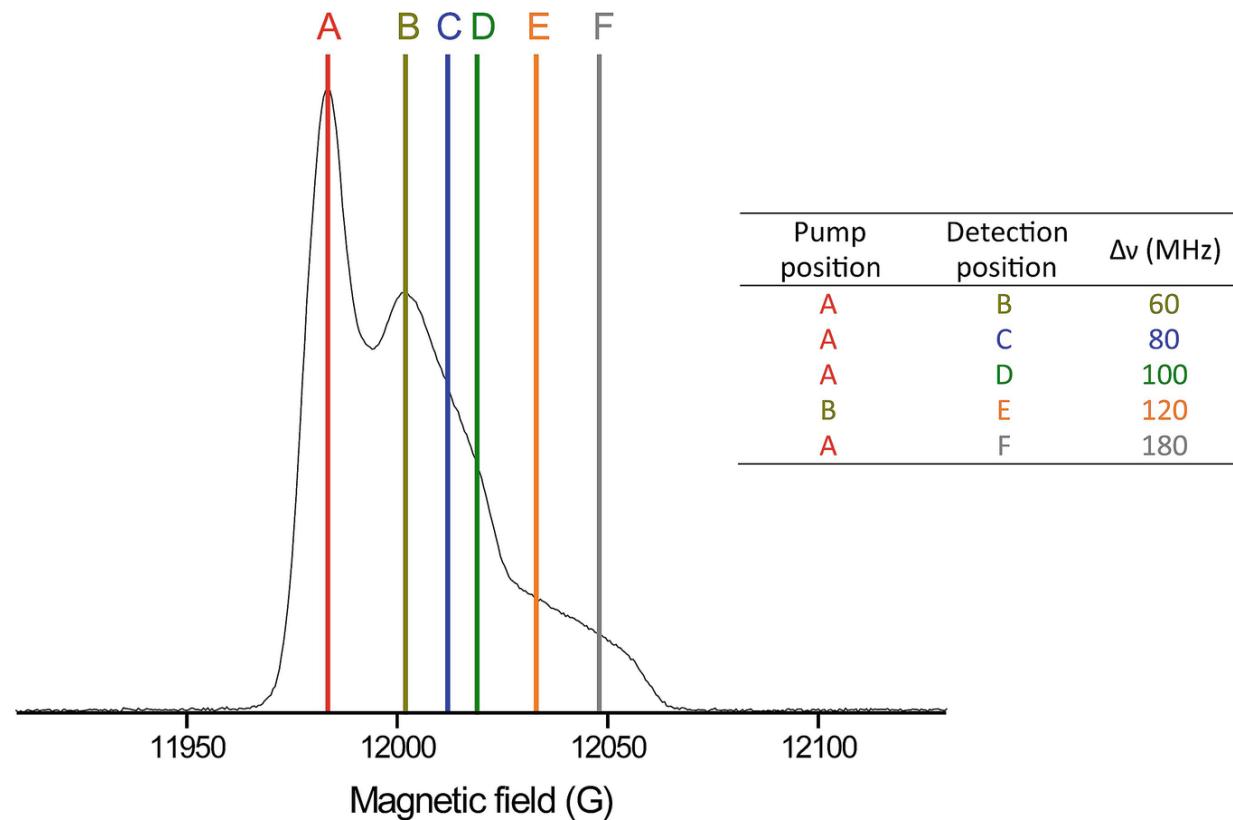


**Fig. 10** Key steps in setting up the PELDOR experiment. (a) Transient nutation experiment (“3P ELDOR Nutation”) to determine the optimal pump pulse length, 16 ns in this case as marked by the red bar. (b) Plot of the refocused echo (“4 P DEER Setup”); the

acquisition delay offset  $d_0$  has been adjusted so that the maximal echo amplitude is located at 12 ns. The red bar indicates the integration gate with a length of 24 ns. (c) PELDOR time trace as obtained from the spectrometer prior to shifting of the zero-time and cutting of the pulse overlap artefact at the end of the trace

### 3.7.5 *Changing the MW Frequency to the Probe Position and Safety Check*

In the next step, change the MW frequency to the probe position: Set the “Attenuation” to 60 dB, the “ELDOR Attenuation” to 30 dB, and switch the TWT amplifier into the Standby mode. As the cathode voltage of the TWT has dropped to 0 V, switch the spectrometer into “CW” mode (“FT Bridge”). Change the MW frequency in the “MW Bridge Tuning” panel to the probe frequency, for example, 80 MHz lower than  $v_{\text{pump}}$ , depending on the desired frequency offset  $\Delta v$  (see Fig. 11 and Note 12).



**Fig. 11** Echo-detected field-swept EPR spectrum with the positions of the pump pulse and the observer pulses indicated for different frequency offsets  $\Delta v$

Perform again the safety check of the detection system. If the defense pulses are present, switch the TWT into the Operate mode.

### ***3.7.6 Optimization of the Hahn Echo at the Probe Frequency***

In the next step, the Hahn echo will be optimized at the probe frequency. Set the “PulseSPEL” variable  $d0 = 360$  ns, select the “2P ESE Setup” experiment with the phase cycle option “+<x> none”, and press the “Run” button in the main window. Press “Start” in the Pulse Tables window, “Run” in “SpecJet,” and decrease the attenuation stepwise to 0 dB. Adjust the video gain to prevent clipping. Then, adjust the “+<x> Amplitude” slider bar in the “MPFU Control” panel such that the echo in the real signal channel is maximized (~72% in the slider bar) and set the “+<x> Phase” such that the real part of the signal becomes maximally negative. Next, select the “-<x> none” phase cycle option and press “Run” again. Adjust the “-<x> Amplitude” slider bar such that the echo in the real signal channel is maximized (~72% in the slider bar), and adjust the “-<x> Phase” slider bar such that the echo amplitude is maximally positive. For both phase cycling options, the imaginary signal channel should be zero on average. Finally, check that the echoes of the phase cycling options “+<x> none” and “-<x> none” have equal absolute amplitudes (see **Note 13**).

### ***3.7.7 Choice of the Dipolar Evolution Window $\tau_2$***

Set the PulseSPEL variable  $d2$ , that is, the length of the dipolar evolution window  $\tau_2$  (Fig. 1a) of the PELDOR time trace, to a value appropriate for the specific case. Generally, longer dipolar evolution windows result in more reliable data and ease background fitting, but come at the expense of longer acquisition times. Thus, the dipolar evolution window should on the one hand not be shorter than at least 1.5 periods of the oscillation [26]. On the other hand, its upper limit is determined by the transverse electron spin relaxation, that is, the maximal length of  $d2$  is given by the time value at which the intensity has almost completely vanished in the Two-Pulse ESEEM experiment (e.g., ~11  $\mu$ s in Fig. 7). If there is no prior knowledge on the expected interspin distance, start a PELDOR experiment with an intermediate length of the dipolar evolution window; a suitable value for  $d2$  may be determined in this case from the echo decay in the Two-Pulse

ESEEM experiment. If you realize that a longer time window is needed for resolving at least 1.5 oscillation periods, abort the experiment and change the PulseSPEL variable  $d2$  accordingly.

### ***3.7.8 Optimization of the Refocused Echo***

Select the PulseSPEL program “4P DEER Setup,” the phase cycling option “2-step”, and press the “Run” button in the main window to execute the standing PELDOR experiment.

Press “Start” in the Pulse Tables tab, “Run” in “SpecJet”, and set the “Number of Averages” to 1 in “SpecJet”. Adjust the video gain such that no clipping of the refocused echo occurs; usually, a video gain between 36 dB and 48 dB is a good choice. Set the number of averages in “SpecJet” to a sufficiently high value to obtain a clear image of the refocused echo (100–1000 averages) and press “Run” in the main window. Read off the time value at which the maximum of the refocused echo occurs. Set the integration gate width (PulseSPEL variable  $pg$ ) to the length of the longest pulse in the PELDOR sequence; this is most likely the  $\pi$ -pulse of the probe sequence (e.g., 24 ns). Adjust the acquisition trigger offset  $d0$  such that the echo maximum is located at the center of the integration gate (e.g., at 12 ns, Fig. 10b). Adjusting the integration gate width to the length of the longest pulse in the sequence and centering the echo within the gate maximizes the SNR for PDS experiments [57].

### ***3.7.9 Setting Parameters for Nuclear Modulation Averaging***

Adjust the parameters for modulation averaging to suppress deuterium ESEEM in the PELDOR time trace. Averaging 8 steps (variable  $m$  in the PulseSPEL program) over one ESEEM oscillation period usually leads to an almost complete suppression of ESEEM in the PELDOR experiment. The time increment for modulation averaging ( $d31$  in the PulseSPEL program) can be computed by  $(1/v_{\text{Larmor}})/m$ , where  $v_{\text{Larmor}}$  is the Larmor frequency of  $^2\text{H}$  at the given field. At Q-band, this results in  $d31 = 16$  ns for  $m = 8$  [35].

### **3.7.10 Setting the Time Resolution for the PELDOR Experiment**

Decide on the temporal resolution of the PELDOR trace, that is, the time increment between two data points on the trace (PulseSPEL variable  $d30$ ). As a rule of thumb, the shorter the period of the dipolar oscillation is, the shorter  $d30$  should be to obtain a sufficiently high resolution. However, a too high resolution in combination with a long dipolar evolution time window (PulseSPEL variable  $d2$ ) will result in excessively long acquisition times. For instance, the interspin distance of 1.5 nm corresponds to a dipolar frequency of  $\sim$ 15 MHz, that is, the oscillation period is  $\sim$ 66 ns; in this case, a time step of 4 ns is recommended, especially to resolve the fast initial decay in the time trace. A trace length of 1  $\mu$ s would be enough to cover several oscillations.

At a distance of 3.5 nm, the dipolar coupling frequency is  $\sim$ 1.2 MHz and the oscillation period  $\sim$ 800 ns. Thus, a time step of 8 ns or 16 ns would be appropriate and the length of the time trace should be  $\sim$ 3–4  $\mu$ s, so that the intermolecular background can be fitted reasonably.

Next, compute the number of points (“dim5” in the PulseSPEL program) to be recorded on the PELDOR time trace, which is given by the equation

$$\text{dim5} = \frac{d1 + d2 - 2 \times d3}{d30}$$

wherein  $d1$ ,  $d2$  = lengths of the interpulse delays  $\tau_1$  and  $\tau_2$  in the PELDOR sequence;  $d3$  = dead time delay of the spectrometer (Fig. 8);  $d30$  = time step of the PELDOR trace; all of these values can be found in the PulseSPEL program. Set the calculated value for  $\text{dim5}$  in the PulseSPEL program. Select the “4P DEER” (Fig. 8) experiment and press the “Run” button. Click the “Re/Im” button in the main window to display the imaginary channel in the viewport, which should fluctuate around zero if the MW phase has been adjusted properly. If there is an appreciable amount of signal in the imaginary channel, adjust the “Signal Phase” in the “FT Bridge”/ “Receiver Unit” window, abort, and restart the “4P DEER” program.

Set the number of scans to a sufficiently high value (e.g.,  $n = 1000$ , depending on the sample and your parameter settings) and let the

measurement run. The SNR of the time trace will improve with the square root of the acquisition time; depending on the spin concentration and the time trace length, acquisition times of 4–24 h are common (see **Note 14**). After stopping the measurement, save the recorded trace to disk (see Fig. 10c, **Note 15**).

For remote monitoring of the current measurement status and for optionally saving intermediate states of the measurement, the MATLAB-application “ScanState.m” may be used, which is freely available [58]. The option to continuously save the traces is especially advantageous if software or hardware errors occur during the measurement, which may lead to deterioration of the data quality (temperature increase due to liquid helium shortage, frequency and phase drifts, TWT faults, software crash, etc.).

## 3.8 Removing the Sample

Drag all MPFU channel slider bars to zero (“MPFU Control”). Set the high power “Attenuation” to 60 dB, the “ELDOR Attenuation” to 30 dB, and switch the TWT into the Standby mode. Go into the “CW” mode (“FT Bridge”/“Bridge Configuration”), readjust the MW frequency to 33.7 GHz with the slider bar, and set the MW Bridge to “Standby” (“MW Bridge Tuning” panel). Switch off the membrane pump and wait until the needle of the gas flow controller has dropped to zero before removing the sample rod from the cryostat (see **Note 16**).

At this point, another measurement on a different sample can be started by following this protocol from Subheading 3.3. Otherwise, the spectrometer should be switched off as described in Subheading 3.9.

## 3.9 Switching off the Spectrometer

First, switch off the TWT by pressing the On/Off-button. It will take a few minutes to cool down and turn off. Close all subwindows of the Xepr-software, but leave the main window open. Disconnect from the EPR spectrometer by clicking “Acquisition”\ “Disconnect from Spectrometer” and then close the main window. Switch off the magnet power supply, the console, and the heat exchanger in the given order. Close the needle valve of the helium transfer line.

## 3.10 Data Analysis

### **3.10.1 Data Formats and Data Conversion**

The Xepr software saves the data in the Bruker BES3T® format generating .DSC and .DTA-files, with the former containing experiment-related acquisition parameters such as the code of a PulseSPEL program and the latter the spectral data in binary format. DTA-files can be conveniently converted into ASCII files (.dat) using the MATLAB-application “DTA2dat” [59], which is based on the eprload-function provided with the EasySpin package [60]. Optionally, exporting the data directly in the ASCII format is also possible in the Xepr software.

### **3.10.2 Analysis of PELDOR Data**

For analyzing PELDOR time traces, various programs have been established, for example, LongDistances [61], DIPFIT [62], GLADD/DD [22], DeerAnalysis [21], and DeerLab [23]. The following section is dedicated to the analysis of PELDOR time traces by means of the MATLAB-based DeerAnalysis program, which can be downloaded free of charge [50].

Time traces saved in the BES3T® format can be directly imported into DeerAnalysis. Analysis of PELDOR traces is commonly subdivided into two steps: (a) the separation of intermolecular and intramolecular contributions to the time trace by fitting the background, and (b) the translation of the background-corrected time trace into a distance distribution. DeerAnalysis provides different computational approaches to fulfil this task, a selection of these is briefly discussed here.

#### **1. Data preprocessing and background removal**

The first step in data analysis is to shift the time axis so that the maximum of the dipolar trace is at the zero time. Additionally, artefacts possibly occurring at the end of the trace due to overlapping of the refocusing pulse and the pump pulse should be cut off (Fig. 10c). Next, the intramolecular dipolar coupling has to be separated from the background, the latter resulting from intermolecular dipolar interactions. Usually, a good starting point is to assume a 3D-homogeneous background and to use the autooptimize function (!-button in the software) for the background start. The quality of the background correction can be judged by inspecting the Fourier

transform of the time domain data. The so-called Pake pattern should neither show a deep hole, nor a spike at the zero frequency [41].

## 2.

### Transformation of the time trace into the distance distribution

DeerAnalysis provides different methods for solving the ill-posed problem of translating the background-corrected time trace, also called form factor, into a distance distribution. Approximate Pake Transformation (APT) is a comparatively fast approach and can provide an initial guess of the distance distribution [21]. Regularization techniques such as model-free Tikhonov regularization can be used to find a more stable solution to the ill-posed problem [21]. After computing the distance distribution, the influence of background correction should be examined by the validation tool in DeerAnalysis, which systematically varies the background correction parameters within given ranges, and statistically evaluates the resulting distance distributions. Uncertainty estimates will be shown as grey shaded areas. Apart from APT and Tikhonov regularization, parametrized models such as Gaussian distance distributions can be used to fit the time traces [21, 63]. This also permits to generate user-defined models.

All of the approaches mentioned above require crucial input from the user, for example, regarding background correction or the choice of the regularization parameter  $\alpha$ . In order to circumvent this user interference and potential biasing of distance distributions, data analysis can also be done by trained neural networks with the DeerNet [51] feature in DeerAnalysis. Within this option, user input is limited to setting the zero-time and the time trace cutoff; background correction and transformation of the form factor into a distance distribution is done by the program automatically. Note that DeerNet requires the Deep Learning Toolbox and the Signal Processing Toolbox to be installed and licensed in MATLAB.

## 3. Data interpretation

The distance information extracted from the PELDOR data can be translated into structural models by means of in silico spin labeling software, for example, the mtsslWizard toolbox [24] for PyMOL, the MMM [25] program for MATLAB, and the GFN-FF based CREST/MD approach [26]. These programs calculate distance distributions based on the structure of the spin label and the structure of the biomolecule, where the latter can often be obtained from the protein

---

data bank [64]. The error of these *in silico* methods is up to 3.5 Å for mtsslWizard [65], down to 2 Å for CREST/MD [26].

As said above, distance distributions can be interpreted in terms of the most probable distance, the mean distance, and the modality. However, especially when it comes to the modality and shoulders of the distribution, one should beware of overinterpretation. These features should only be interpreted if the underlying time trace is of high quality; if these features are interpreted, they must not vanish in the validation of the PELDOR data and further evidence should be gained from other biophysical methods that supports the interpretation [49]. When analyzing distance distributions, one should also keep in mind that shoulders or peaks may be associated either with different conformations of biomolecules or of the label. One should also consider possible changes of the structure and function of the biomolecule upon spin labeling. If there are biochemical assays which can be used to prove the proper function of the labeled biomolecule, these should be performed [49].

### **3.10.3 Data Quality and Reproducibility**

The quality of PELDOR data can be judged in terms of the modulation depth  $\lambda$  (Fig. 1) and the SNR of the time trace. It thus serves as a control of the sample quality and the measurement set-up.

For a two-spin system, the modulation depth achievable in the Q-band with rectangular pulses and a 150 W TWT amplifier amounts to 30–35%. The modulation depth decreases if the sample contains simply labeled biomolecules, either due to a low labeling efficiency (<80%) or due to partial degradation of the spin label. Free label remaining in the sample can also affect the modulation depth adversely. In terms of setting up the PELDOR measurement, double-checking the chosen acquisition parameters can help to exclude technical pitfalls (see **Note 15**).

The SNR of a PELDOR time trace can be given as the modulation depth divided by the standard deviation of the noise in the trace. Separation of the dipolar signal and the noise is achieved by polynomial fitting. The SNR can be determined using, for example, the program SnrCalculator, which is freely available [66]. As a rule of thumb, meaningful data analysis is

possible if the absolute SNR is  $>20$ , with good data having an SNR larger than 100.

In general, the SNR depends linearly on the spin concentration of the sample and on the square root of the measurement time. However, it should also be considered that the intermolecular background decay in the PELDOR trace correlates with the spin concentration, and background fitting can become difficult at high spin concentrations. Finally, it should be noted that it is good practice to repeat a PELDOR measurement on independent samples prepared in the same manner, for example, in triplicate, to verify the results.

---

## 4 Notes

1.

### Cryoprotectants and their influence on the biomolecule

The cryoprotectant acts as a glass-forming agent and prevents clustering of the biomolecules [67]. Choose the cryoprotectant which has the least impact on the structure and function of the biomolecule, and keep in mind that some biomolecules can bind and interact with EG-d<sub>6</sub> or glycerol-d<sub>8</sub>. In the literature, different amounts of cryoprotectant have been reported, ranging from ~20% v/v [68] up to 50% v/v [69]. The optimal ratio of the cryoprotectant and the biomolecule has to be determined experimentally for each case.

2.

### Freezing samples in the EPR tube

Tubes may crack during shock-freezing or upon rapid warming of the frozen tube, and possibly condensed liquid nitrogen may evaporate. This can result in uncontrolled ejection of the sample and in rupture of the tube. When working with larger amounts of liquid nitrogen, wear cold protection gloves in addition to the safety goggles.

3. Problems in cooling down and reaching a stable temperature of 50 K.

If the cryostat cannot be cooled down at all, check the following points in the given order: Is the filling level of the helium tank sufficient? Has all tubing been connected tightly and to the correct ports of the gas flow controller and the helium pump? Double-check the inscription of the respective ports. If existing, is the helium recovery port open? Is the transfer line permeable and not blocked, and is it tightly screwed into the cryostat?

If the temperature does not stabilize at 50 K, make sure that the flow of cold helium gas amounts to ca. 1 L/h. If the flow is far off, it is difficult to reach a stable temperature. Also, ensure that the vacuum within the transfer line insulation is properly maintained and, if necessary, evacuate the transfer line overnight.

#### 4.

##### Sample insertion

Avoid transferring drops of liquid nitrogen sticking to the sample rod by gently shaking it before insertion into the cryostat. In this context, it is convenient to use a *small* Dewar vessel (40 mm inner width, 90 mm inner height) so that the EPR tube and the lower part of the sample holder can be immersed into liquid nitrogen, but not the rod itself. Use a ring stand to prevent the sample rod from falling over.

#### 5. Comparison of the Pulse Tables and PulseSPEL

Timing of the pulse sequence differs between the Pulse Tables and PulseSPEL. Whereas the Pulse Tables use absolute timing with respect to the beginning of the pulse sequence to define pulses, interpulse delays, and echo positions, PulseSPEL refers to the previous event in the sequence. This difference is exemplified by means of the Hahn echo sequence (12 ns–200 ns–24 ns–200 ns), and the timing is given for the Pulse Tables and PulseSPEL (Fig. 12).

Running the Hahn echo from the Pulse Tables (Fig. 12a, b), detection starts at the beginning of the sequence and the echo will occur at ~800 ns, reaching its maximum at 832 ns. As shown in Fig. 12a, the exact echo position depends on the instrument delay, which is ~400 ns in this case. For recording the echo-detected field-swept EPR spectrum, the acquisition trigger should thus be set to the position ~790 ns with a length of ~120 ns. Note: Even though the time axis starts with the first pulse, the pulses themselves are not recorded as the detector is “closed” to protect it from the high-power MW radiation.

If the Hahn echo sequence is run from PulseSPEL (Fig. 12c,d), detection starts at the end of the second interpulse delay  $\tau$ . Thus, the offset between the origin of the time axis and the echo start is ~400 ns, corresponding to the instrument delay.

Pulse Tables and PulseSPEL have in common that the interpulse

delays always start at the leading edge of a pulse [70].

## 6.

It is difficult to achieve a stable tuning condition

Unstable tuning can have several reasons. Check and try to address the following points:

- (a) The MW frequency is unstable. Possible reasons are as follows:
  - (1) The MW Bridge was not warmed up long enough. Give the instrument a longer period of time to warm up before starting experiments, at least 1 h (2). The cooling water entering the MW Bridge is too warm, possibly due to chiller faults or insufficient cooling water in the system. Check the cooling water level and check if the cooling aggregate has turned off.
- (b) Sample and cryostat are not in the thermal equilibrium. Give the sample at least 15 min to reach and stabilize at the temperature of 50 K.
- (c) If the tuning dips start moving even without any changes being made to the spectrometer, this might indicate condensation of liquid nitrogen in the cryostat. That often comes along with an unusual tuning picture different from the one shown in Fig. 5a. Remove the sample as described in Subheading 3.8, warm up the spectrometer to at least 120 K for at least 20 min, and start from the beginning. If this does not help in resolving the problem, warm up the spectrometer to room temperature and blow a stream of nitrogen gas through the resonator overnight.

## 7. Safety check

As a warning reminder, never switch the TWT into the Operate mode if the defense pulses have not been displayed in “SpecJet” as this would damage the detection system. Try the following steps if you cannot see the defense pulses: Make sure that a valid pulse sequence is being executed from the Pulse Tables and that “SpecJet” is running with enough points being displayed. Make sure that the “Video Gain” has been set to 45 dB and that the “MW Amplifier”

button has been switched to “On” (“FT Bridge”/“Receiver Unit”). Occasionally, even though the defense pulses actually are present, they cannot be seen due to the phase settings. Drag the shifter of the “Signal Phase” (e.g., in “FT Bridge”/“Receiver Unit”) to change the phase settings a bit and see whether the defense pulses become visible. If you are still unsuccessful in checking the defense pulses, contact your spectrometer administrator as some basic settings in the software or hardware might be faulty.

## 8. Finding and optimizing the Hahn echo

If no echo can be observed at all, firstly ensure that the pulse sequence has been programmed properly and check that “SpecJet” displays enough points to detect a signal up to the time where the echo is expected. This can vary depending on the individual dead time of the spectrometer (see **Note 5**). If necessary, increase the “Number of Points” in “SpecJet.” Make sure that the TWT has been switched into the Operate mode. Double-check the combination of magnetic field position and MW frequency.

Next, check proper tuning (Subheading 3.5) and proper sample positioning in the resonator. Refer to Subheading 3.8 for safe sample removal from the spectrometer. Make sure that the EPR tube has not been pushed into the sample rod during insertion. This may occur if the cavity is blocked by ice (frozen nitrogen or air moisture). In that case, warm up and completely defrost the spectrometer (see **Notes 4** and **6c**).

If the resonator has been properly overcoupled, the maximal echo amplitude should be obtained at an attenuation of approximately 3–5 dB when using pulse lengths of  $\pi/2 = 12$  ns and  $\pi = 24$  ns. If the echo amplitude keeps growing when decreasing the attenuation <3 dB, go back to Subheading 3.5 and check the tuning. If the amplitude of the Hahn echo still cannot be maximized, the transmitted power is insufficient to achieve  $\pi/2$  and  $\pi$ -pulses at the given pulse lengths. In this case, increase the pulse lengths (e.g., to  $\pi/2 = 14$  ns and  $\pi = 28$  ns or  $\pi/2 = 16$  ns and  $\pi = 32$  ns) and try again.

If the echo shows a low signal-to-noise ratio and requires a rather high “Video Gain” setting, the spin concentration in the sample may not be sufficient. Perform a *cw*-EPR spin counting experiment; the spin concentration in the sample should amount to at least ~20  $\mu\text{M}$  to obtain a decent signal.

9.

## Two-Pulse ESEEM

Use the pulse lengths for  $\pi/2$  and  $\pi$  which you identified as optimal for the standing Hahn echo. The following variable settings for the Two-Pulse ESEEM experiment should work if the Hahn echo has been optimized with pulse lengths of 12 ns and 24 ns for  $\pi/2$  and  $\pi$ , respectively:  $p0 = 12$  ns;  $p1 = 24$  ns;  $d0 = 0$  ns (for setup) and  $d0 = 432$  ns (for experiment);  $d1 = 200$  ns;  $d30 = 8$  ns;  $h = 10$ ;  $n = 1$ .

Herein,  $p0$  and  $p1$  is the length of the  $\pi/2$  and  $\pi$ -pulses, respectively;  $d0$  is a constant instrument-related acquisition trigger delay;  $d1$  is the interpulse delay denoted  $\tau$  in the Hahn echo sequence (Fig. 6);  $d30$  is the time step by which  $d1$  is incremented;  $h$  is the number of shots per point;  $SRT$  is the shot repetition time, and  $n$  is the number of scans to accumulate.

10.

## Inversion of the Hahn echo

If no inversion of the echo is observed, make sure that the “Current ELDOR Frequency” has been set to the correct value, that is, the current spectrometer frequency at which the pump pulse should be applied. Ensure that the pump pulse length (variable  $p2$  in PulseSPEL) is between 12 and 18 ns. If the echo still does not invert upon reducing the “ELDOR Attenuation,” a power loss in the ELDOR channel may have occurred. Contact your spectrometer administrator.

11.

The nutation experiment reports a full inversion of the echo at pump pulse lengths  $>18$  ns:

(a)

Make sure that the “ELDOR Attenuation” has been set to 0 dB.

(b)

Check proper tuning (see Subheading 3.5 and Note 6). Make sure that the resonator dip has been centered in the in the “MW Bridge Tuning” panel.

(c)

Unusually long pump pulse lengths may be indicative of liquid nitrogen condensation in the cryostat (see Note 6c).

12. Choice of the frequency offset  $\Delta\nu$

Generally a smaller offset  $\Delta\nu$  between the numn and observer

frequency leads to a higher signal intensity and thus an improved SNR. However, choosing a too small  $\Delta v$  ( $\Delta v < 50$  MHz) will lead to overlapping excitation bandwidths of the probe and pump pulses and thus deteriorate the modulation depth. As a first attempt, use an offset of  $\Delta v = 80$  MHz or 100 MHz.

Regarding the choice of the frequency offset, it should be kept in mind that orientation selection can occur, especially when rigid spin labels are used [15, 71]. Orientation selection means that spins on resonance at a given combination of magnetic field and MW frequency correspond to a certain selected orientation. It implies that not all orientations of the distance vector with respect to the magnetic field direction are excited [72] and manifests itself in the oscillation periods of the time traces depending on the used frequency offset [73–75]. Orientation selection can be recognized best from the Fourier transform (FFT) of the time traces which, in the presence of orientation selection, often lacks the parallel component of the Pake pattern [15]. It is good practice to check the presence of orientation selection by recording PELDOR traces at different frequency offsets (e.g., 60 MHz; 80 MHz; 100 MHz; 120 MHz; 180 MHz, Fig. 11). If orientation selection occurs, the PELDOR traces as such can no longer be analyzed individually. In order to minimize orientational selectivity, divide the signal intensity values of the traces by the respective number of scans and sum up all individual traces. The resulting time trace can then be processed by, for example, DeerAnalysis to obtain the distance distribution.

Apart from this method of analyzing orientation-selective PELDOR traces, there are programs such as PeldorFit [45] which can also extract angular information from the time trace in addition to the distance.

### 13. No echo is seen

If no echo is seen at this stage of the experiment setup, check that the parameters set in PulseSPEL (interpulse delays  $d1$  and  $d2$ ; pulse lengths  $p0$ ,  $p1$ ,  $p2$ ; acquisition trigger delay  $d0$ ) are correct and ensure that enough points are shown in “SpecJet” to monitor the echo. Make sure not to mix up the units of the timing (e.g.,  $SRT$  is given in  $\mu$ s;

pulse lengths and delays are mostly given in ns in PulseSPEL;  $d2$  in the PulseSPEL variable box is given in  $\mu$ s, however).

14.

Acquisition time:

In general, the approximate acquisition time for a given number of scans  $n$  can be determined by

$$\text{Acquisition time} = \text{dim5} \times \text{PC} \times m \times \text{SRT} \times h \times n$$

Herein, dim5 = number of points on the trace as set in the PulseSPEL program; PC = number of phase cycling steps (usually 2 for PELDOR);  $m$  = number of modulation averaging steps (usually 8 for PELDOR); SRT = shot repetition time,  $h$  = number of shots per point,  $n$  = number of scans. Note, however, that the herein computed duration should be understood as a lower limit since overhead occurs due to reprogramming of the pulse programmer (called “PatternJet” in the Bruker instruments) during the pulse sequence [76]. Depending on the version of the PatternJet and especially for long experiments, deviations of up to several minutes between the predicted and the true acquisition time can occur.

15. The modulation depth of the PELDOR time trace and the SNR are lower than expected:

(a)

Double-check the attenuation (high power “Attenuation” and “ELDOR Attenuation” should have been set to 0 dB), PulseSPEL variable settings (especially pump pulse length  $p2$ ), amplitude and phase settings (“MPFU Control”), frequencies (probe and “Current ELDOR Frequency”), and the magnetic field position (“Center Field”).

(b)

It is possible that the MW Bridge was not sufficiently warmed up and/or the sample and cryostat were not in the thermal equilibrium before starting the experiment. Concomitant phase and frequency drifts might have led to deterioration of the data quality.

(c)

The sample was not optimally prepared and contains, for example, simply labeled biomolecules or free spin label. Verify proper spin labeling by a <sup>1</sup>H EDD spin counting experiment

## proper spin trapping by a low-EPR spin counting experiment.

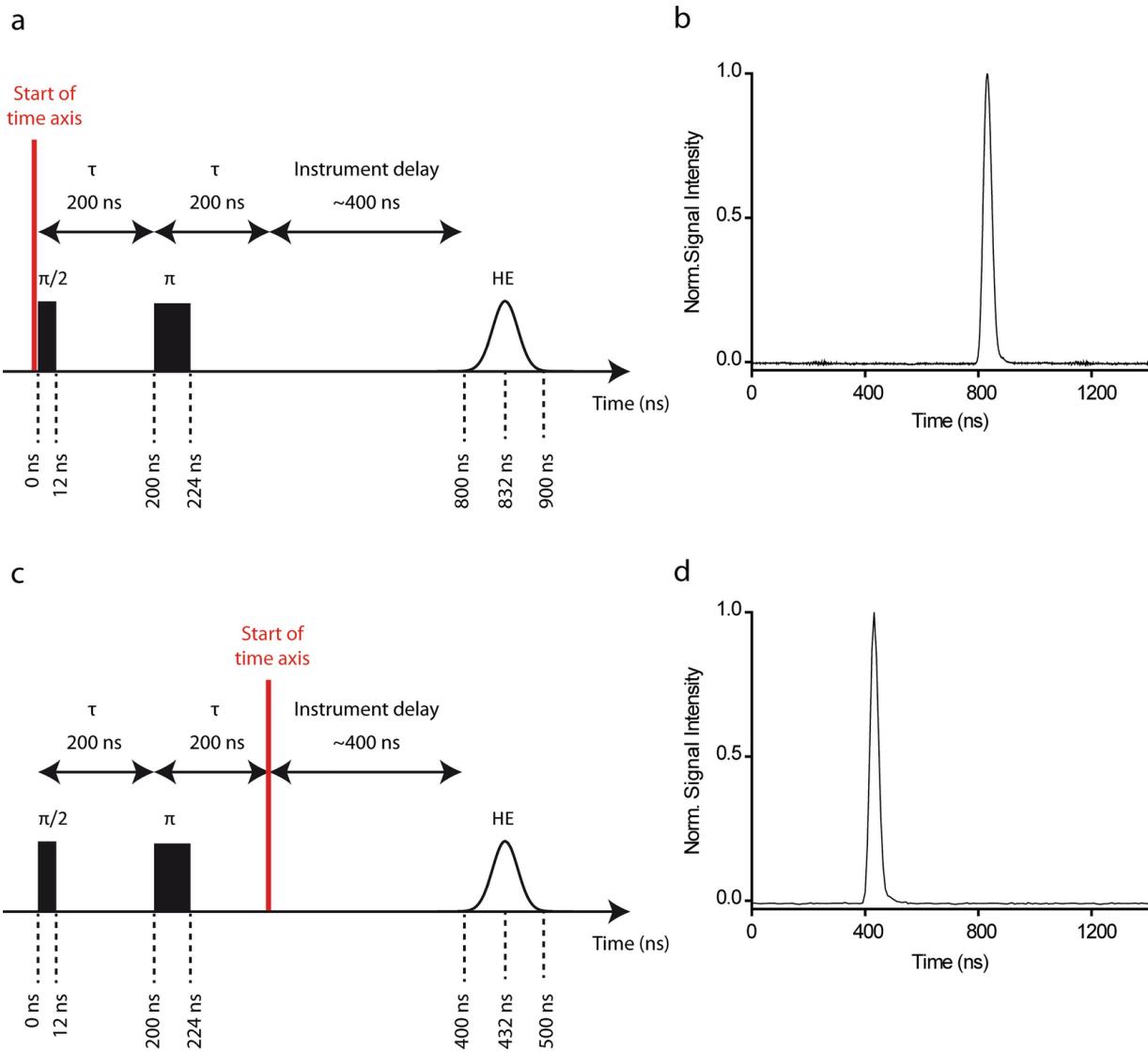
16.

The sample cannot be removed from the cryostat

It may occur that the sample tube freezes and gets stuck in the cryostat, especially after opening and closing the cryostat many times when changing the sample, as this will inevitably lead to condensation of air. If the sample rod cannot be removed at all from the spectrometer, warm up the cryostat to a temperature above 77 K and try again to remove the rod. If only the EPR tube has got stuck in the cryostat, carefully reinsert the sample rod and warm up the cryostat to a temperature above 77 K. Then, try again to remove the tube. If this does not help, warm up the spectrometer to room temperature.

However, keep in mind that the tube may burst when warming up and would thus contaminate the resonator. In this case, careful cleaning as described in the spectrometer manual is mandatory.

In order to prevent sample tubes from getting stuck in the cryostat, try to avoid transferring drops of liquid nitrogen into it by using a small Dewar vessel as described in **Note 4**. Also, minimize the time the resonator is opened to prevent condensation of air.



**Fig. 12** Timing of pulse sequences exemplified by the Hahn echo. **(a)** *Absolute* timing with respect to the beginning of the pulse sequence as used by the pulse tables. The time axis starts at the beginning of the pulse sequence; in consideration of the instrument delay, the echo arises at ~800 ns. **(b)** Hahn echo recorded from the pulse tables. **(c)** *Relative* timing with respect to the end of the last interpulse delay of the pulse sequence as used by PulseSPEL. The time axis starts at the end of the last interpulse delay; in consideration of the instrument delay, the echo arises at ~400 ns. **(d)** Hahn echo recorded from PulseSPEL

## References

1. Goldfarb D, Stoll S (eds) (2018) EPR spectroscopy: fundamentals and methods. eMagRes books. Wiley, Chichester, West Sussex
2. Schweiger A, Jeschke G (2001) Principles of pulse electron paramagnetic resonance.

Oxford University Press, Oxford

3. Timmel CR, Harmer JR (2013) Structural information from spin-labels and intrinsic paramagnetic centres in the biosciences. In: Structure and bonding, vol 152. Springer, Berlin
4. Marsh D (2020) Spin-label electron paramagnetic resonance spectroscopy. CRC Press, Taylor & Francis Group, Boca Raton
5. Fleck N, Heubach CA, Hett T et al (2020) SLIM: a short-linked, highly redox-stable trityl label for high-sensitivity in-cell EPR distance measurements. *Angew Chem Int Ed Engl* 59(24):9767–9772. <https://doi.org/10.1002/anie.202004452>  
[Crossref][PubMed][PubMedCentral]
6. Jassoy JJ, Heubach CA, Hett T et al (2019) Site selective and efficient spin labeling of proteins with a maleimide-functionalized trityl radical for pulsed dipolar EPR spectroscopy. *Molecules* 24(15):2735. <https://doi.org/10.3390/molecules24152735>  
[Crossref][PubMedCentral]
7. Yang Y, Pan B-B, Tan X et al (2020) In-cell trityl-trityl distance measurements on proteins. *J Phys Chem Lett* 11(3):1141–1147. <https://doi.org/10.1021/acs.jpclett.9b03208>  
[Crossref][PubMed][PubMedCentral]
8. Fleck N, Heubach C, Hett T, Spicher S, Grimme S, Schiemann O (2021) Ox-SLIM: synthesis of and site-specific labelling with a highly hydrophilic trityl spin label. *Chem Eur J* 27(16):5292–5297. <https://doi.org/10.1002/chem.202100013>
9. Qi M, Gross A, Jeschke G et al (2014) Gd(III)-PyMTA label is suitable for in-cell EPR. *J Am Chem Soc* 136(43):15366–15378. <https://doi.org/10.1021/ja508274d>  
[Crossref][PubMed]
10. Yang Y, Yang F, Gong Y-J et al (2018) High sensitivity in-cell EPR distance measurements on proteins using an optimized Gd(III) spin label. *J Phys Chem Lett* 9(20):6119–6123. <https://doi.org/10.1021/acs.jpclett.8b02663>  
[Crossref][PubMed]
11. Singewald K, Bogetti X, Sinha K et al (2020) Double histidine based EPR measurements at physiological temperatures permit site-specific elucidation of hidden dynamics in enzymes. *Angew Chem Int Ed Engl* 59(51):23040–23044. <https://doi.org/10.1002/anie.202009982>  
[Crossref][PubMed]
12. Cunningham TF, Puttermann MR, Desai A et al (2015) The double-histidine Cu<sup>2+</sup>-binding motif: a highly rigid, site-specific spin probe for electron spin resonance

- distance measurements. *Angew Chem Int Ed Engl* 54(21):6330–6334. <https://doi.org/10.1002/anie.201501968>  
[Crossref][PubMed][PubMedCentral]
13. Sahu ID, McCarrick RM, Lorigan GA (2013) Use of electron paramagnetic resonance to solve biochemical problems. *Biochemistry* 52(35):5967–5984. <https://doi.org/10.1021/bi400834a>  
[Crossref][PubMed]
  14. Milov AD, Salikhov KM, Schirov MD (1981) Application of ELDOR in electron-spin echo for paramagnetic center space distribution in solids. *Fiz Tverd Tela* 23:975–982
  15. Pannier M, Veit S, Godt A et al (2000) Dead-time free measurement of dipole-dipole interactions between electron spins. *J Magn Reson* 142(2):331–340. <https://doi.org/10.1006/jmre.1999.1944>  
[Crossref][PubMed]
  16. Jeschke G (2012) DEER distance measurements on proteins. *Annu Rev Phys Chem* 63:419–446. <https://doi.org/10.1146/annurev-physchem-032511-143716>  
[Crossref][PubMed]
  17. Schmidt T, Wälti MA, Baber JL et al (2016) Long distance measurements up to 160 Å in the GroEL Tetradecamer using Q-band DEER EPR spectroscopy. *Angew Chem Int Ed Engl* 55(51):15905–15909. <https://doi.org/10.1002/anie.201609617>  
[Crossref][PubMed][PubMedCentral]
  18. Schiemann O, Prisner TF (2007) Long-range distance determinations in biomacromolecules by EPR spectroscopy. *Q Rev Biophys* 40(1):1–53. <https://doi.org/10.1017/S003358350700460X>  
[Crossref][PubMed]
  19. Reginsson GW, Schiemann O (2011) Pulsed electron-electron double resonance: beyond nanometre distance measurements on biomacromolecules. *Biochem J* 434(3):353–363. <https://doi.org/10.1042/BJ20101871>  
[Crossref][PubMed]
  20. Borbat PP, Freed JH (2013) Pulse dipolar electron spin resonance: distance measurements. In: Timmel CR, Harmer JR (eds) Structural information from spin-labels and intrinsic paramagnetic centres in the biosciences, vol 152. Springer, Berlin, pp 1–82  
[Crossref]
  21. Jeschke G, Chechik V, Ionita P et al (2006) DeerAnalysis2006—a comprehensive software package for analyzing pulsed ELDOR data. *Appl Magn Reson* 30(3–4):473–498. <https://doi.org/10.1007/BF03166213>

[Crossref]

22. Brandon S, Beth AH, Hustedt EJ (2012) The global analysis of DEER data. *J Magn Reson* 218:93–104. <https://doi.org/10.1016/j.jmr.2012.03.006>  
[Crossref][PubMed][PubMedCentral]
23. Fábregas Ibáñez L, Jeschke G, Stoll S (2020) DeerLab: a comprehensive software package for analyzing dipolar electron paramagnetic resonance spectroscopy data. *Magn Reson* 1(2):209–224. <https://doi.org/10.5194/mr-1-209-2020>  
[Crossref]
24. Hagelueken G, Abdullin D, Schiemann O (2015) mtsslSuite: probing biomolecular conformation by spin-labeling studies. *Methods Enzymol* 563:595–622. <https://doi.org/10.1016/bs.mie.2015.06.006>  
[Crossref][PubMed]
25. Polyhach Y, Bordignon E, Jeschke G (2011) Rotamer libraries of spin labelled cysteines for protein studies. *Phys Chem Chem Phys* 13(6):2356–2366. <https://doi.org/10.1039/C0CP01865A>
26. Spicher S, Abdullin D, Grimme S et al (2020) Modeling of spin-spin distance distributions for nitroxide labeled biomacromolecules. *Phys Chem Chem Phys* 22(42):24282–24290. <https://doi.org/10.1039/D0CP04920D>  
[Crossref][PubMed]
27. Reckel S, Gottstein D, Stehle J et al (2011) Solution NMR structure of Proteorhodopsin. *Angew Chem* 123(50):12148–12152. <https://doi.org/10.1002/ange.201105648>  
[Crossref]
28. Fürtig B, Richter C, Wöhnert J et al (2003) NMR spectroscopy of RNA. *Chembiochem* 4(10):936–962. <https://doi.org/10.1002/cbic.200300700>  
[Crossref][PubMed]
29. Westhof E (2015) Twenty years of RNA crystallography. *RNA* 21(4):486–487. <https://doi.org/10.1261/rna.049726.115>  
[Crossref][PubMed][PubMedCentral]
30. Cate JH, Doudna JA (2000) Solving large RNA structures by X-ray crystallography. *Methods Enzymol* 317:169–180. [https://doi.org/10.1016/s0076-6879\(00\)17014-4](https://doi.org/10.1016/s0076-6879(00)17014-4)  
[Crossref][PubMed]
31. Khatter H, Myasnikov AG, Natchiar SK et al (2015) Structure of the human 80S ribosome. *Nature* 520(7549):640–645. <https://doi.org/10.1038/nature14427>  
[Crossref][PubMed]

32. Wrapp D, Wang N, Corbett KS et al (2020) Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* 367(6483):1260–1263. <https://doi.org/10.1126/science.abb2507>  
[Crossref][PubMed][PubMedCentral]
33. Stephenson JD, Kenyon JC, Symmons MF et al (2016) Characterizing 3D RNA structure by single molecule FRET. *Methods* 103:57–67. <https://doi.org/10.1016/jymeth.2016.02.004>  
[Crossref][PubMed]
34. Lakowicz JR (2006) Principles of fluorescence spectroscopy, 3rd edn. Springer, New York, NY  
[Crossref]
35. Malygin AA, Krumkacheva OA, Graifer DM et al (2019) Exploring the interactions of short RNAs with the human 40S ribosomal subunit near the mRNA entry site by EPR spectroscopy. *Nucleic Acids Res* 47(22):11850–11860. <https://doi.org/10.1093/nar/gkz1039>  
[Crossref][PubMed][PubMedCentral]
36. Pliotas C, Ward R, Branigan E et al (2012) Conformational state of the MscS mechanosensitive channel in solution revealed by pulsed electron-electron double resonance (PELDOR) spectroscopy. *Proc Natl Acad Sci U S A* 109(40):E2675–E2682. <https://doi.org/10.1073/pnas.1202286109>  
[Crossref][PubMed][PubMedCentral]
37. Constantinescu-Aruxandei D, Petrovic-Stojanovska B, Schiemann O et al (2016) Taking a molecular motor for a spin: helicase mechanism studied by spin labeling and PELDOR. *Nucleic Acids Res* 44(2):954–968. <https://doi.org/10.1093/nar/gkv1373>  
[Crossref][PubMed]
38. Duss O, Michel E, Yulikov M et al (2014) Structural basis of the non-coding RNA RsmZ acting as a protein sponge. *Nature* 509(7502):588–592. <https://doi.org/10.1038/nature13271>  
[Crossref][PubMed]
39. Duss O, Yulikov M, Jeschke G et al (2014) EPR-aided approach for solution structure determination of large RNAs or protein-RNA complexes. *Nat Commun* 5:3669. <https://doi.org/10.1038/ncomms4669>  
[Crossref][PubMed]
40. Ward R, Pliotas C, Branigan E et al (2014) Probing the structure of the mechanosensitive channel of small conductance in lipid bilayers with pulsed electron-

- electron double resonance. *Biophys J* 106(4):834–842. <https://doi.org/10.1016/j.bpj.2014.01.008>  
[Crossref][PubMed][PubMedCentral]
41. Joseph B, Tormyshev VM, Rogozhnikova OY et al (2016) Selective high-resolution detection of membrane protein-ligand interaction in native membranes using trityl-nitroxide PELDOR. *Angew Chem Int Ed Engl* 55(38):11538–11542. <https://doi.org/10.1002/anie.201606335>  
[Crossref][PubMed]
  42. Meyer V, Swanson MA, Clouston LJ et al (2015) Room-temperature distance measurements of immobilized spin-labeled protein by DEER/PELDOR. *Biophys J* 108(5):1213–1219. <https://doi.org/10.1016/j.bpj.2015.01.015>  
[Crossref][PubMed][PubMedCentral]
  43. Gränz M, Erlenbach N, Spindler P et al (2018) Dynamics of nucleic acids at room temperature revealed by pulsed EPR spectroscopy. *Angew Chem Int Ed Engl* 57(33):10540–10543. <https://doi.org/10.1002/anie.201803682>  
[Crossref][PubMed]
  44. Stelzl LS, Erlenbach N, Heinz M et al (2017) Resolving the conformational dynamics of DNA with Ångstrom resolution by pulsed electron-electron double resonance and molecular dynamics. *J Am Chem Soc* 139(34):11674–11677. <https://doi.org/10.1021/jacs.7b05363>  
[Crossref][PubMed]
  45. Abdullin D, Hagelueken G, Hunter RI et al (2015) Geometric model-based fitting algorithm for orientation-selective PELDOR data. *Mol Phys* 113(6):544–560. <https://doi.org/10.1080/00268976.2014.960494>  
[Crossref]
  46. Hagelueken G, Ingledew WJ, Huang H et al (2009) PELDOR spectroscopy distance fingerprinting of the octameric outer-membrane protein Wza from *Escherichia coli*. *Angew Chem Int Ed Engl* 48(16):2904–2906. <https://doi.org/10.1002/anie.200805758>  
[Crossref][PubMed][PubMedCentral]
  47. Valera S, Ackermann K, Pliotas C et al (2016) Accurate extraction of nanometer distances in multimers by pulse EPR. *Chemistry* 22(14):4700–4703. <https://doi.org/10.1002/chem.201505143>  
[Crossref][PubMed][PubMedCentral]
  48. Hett T, Zbik T, Mukherjee S, Matsuoka H, Bönigk W, Klose D, Rouillon C, Brenner N, Peuker S, Klement R, Steinhoff H-J, Grubmüller H, Seifert R, Schiemann O, Kaupp UB (2021) Spatiotemporal resolution of conformational changes in

- biomolecules by combining pulsed electron–electron double resonance spectroscopy with microsecond freeze-hyperquenching.. J Am Chem Soc 143(18):6981-6989. <https://doi.org/10.1021/jacs.1c01081>
49. Schiemann O, Heubach CA, Abdullin D, Ackermann K, Azarkh M, Bagryanskaya EG, Drescher M, Endeward B, Freed JH, Galazzo L, Goldfarb D, Hett T, Esteban Hofer L, Fábregas Ibáñez L, Hustedt EJ, Kucher S, Kuprov I, Lovett JE, Meyer A, Ruthstein S, Saxena S, Stoll S, Timmel CR, Di Valentin M, Mchaourab HS, Prisner TF, Bode BE, Bordignon E, Bennati M, Jeschke G (2021) Benchmark test and guidelines for DEER/PELDOR experiments on nitroxide-labeled biomolecules. J Am Chem Soc 143(43):17875–17890. <https://doi.org/10.1021/jacs.1c07371>
  50. Jeschke G (2019) DeerAnalysis 2019. <https://epr.ethz.ch/software.html>. Accessed 11 Jan 2021
  51. Worswick SG, Spencer JA, Jeschke G et al (2018) Deep neural network processing of DEER data. Sci Adv 4(8):eaat5218. <https://doi.org/10.1126/sciadv.aat5218> [Crossref][PubMed][PubMedCentral]
  52. Vicino MF, Wuebben C, Kerzhner M, Famulok M, Schiemann O. Spin labeling of long RNAs via click reaction and enzymatic ligation. Methods Mol Biol
  53. Bruker BioSpin GmbH, EPR Division (2014) Product description: ER 5106QT-2 Q-Band Probehead
  54. Hett T (2020) PELDOR\_PulseSPEL. [https://github.com/TobiasHett/PELDOR\\_PulseSPEL](https://github.com/TobiasHett/PELDOR_PulseSPEL). Accessed 11 Jan 2021
  55. Ghimire H, McCarrick RM, Budil DE et al (2009) Significantly improved sensitivity of Q-band PELDOR/DEER experiments relative to X-band is observed in measuring the intercoil distance of a leucine zipper motif peptide (GCN4-LZ). Biochemistry 48(25):5782–5784. <https://doi.org/10.1021/bi900781u> [Crossref][PubMed]
  56. Endeward B, Butterwick JA, MacKinnon R et al (2009) Pulsed electron-electron double-resonance determination of spin-label distances and orientations on the tetrameric potassium ion channel KcsA. J Am Chem Soc 131(42):15246–15250. <https://doi.org/10.1021/ja904808n> [Crossref][PubMed][PubMedCentral]
  57. Jeschke G (2007) Instrumentation and experimental setup. In: ESR spectroscopy in membrane biophysics, vol 27. Springer US, Boston, MA, pp 17–47 [Crossref]
  58. Hett T (2020) ScanState. <https://github.com/TobiasHett/ScanState>. Accessed 11 Jan

2021

59. Hett T (2020) DTA2dat. <https://github.com/TobiasHett/DTA2dat>. Accessed 11 Jan 2021
60. Stoll S, Schweiger A (2006) EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J Magn Reson* 178(1):42–55. <https://doi.org/10.1016/j.jmr.2005.08.013>  
[Crossref][PubMed]
61. Altenbach C (2020) LongDistances. <https://sites.google.com/site/altenbach/labview-programs/epr-programs/long-distances>. Accessed 22 Jan 2021
62. Steinhoff HJ, Radzwill N, Thevis W et al (1997) Determination of interspin distances between spin labels attached to insulin: comparison of electron paramagnetic resonance data with the X-ray structure. *Biophys J* 73(6):3287–3298. [https://doi.org/10.1016/S0006-3495\(97\)78353-X](https://doi.org/10.1016/S0006-3495(97)78353-X)  
[Crossref][PubMed][PubMedCentral]
63. Jeschke G, Stoll S (2019) DeerAnalysis User Manual - Version 2019. <http://www.epr.ethz.ch/software/index>. Accessed 11 Jan 2021
64. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Res* 28:235–242. <https://www.rcsb.org>  
[Crossref]
65. Hagelueken G, Ward R, Naismith JH et al (2012) MtsslWizard: in silico spin-labeling and generation of distance distributions in PyMOL. *Appl Magn Reson* 42(3):377–391. <https://doi.org/10.1007/s00723-012-0314-0>  
[Crossref][PubMed][PubMedCentral]
66. Abdullin D, Brehm P, Fleck N, Spicher S, Grimme S, Schiemann O (2019) Pulsed EPR dipolar spectroscopy on spin pairs with one highly anisotropic spin center: the low-spin Fe(III) case. *Chem Eur J* 25:14388–14398. <https://github.com/dinarabdullin/SnrCalculator>  
[Crossref]
67. Schmidt T, Jeon J, Okuno Y et al (2020) Submillisecond freezing permits cryoprotectant-free EPR double electron-electron resonance spectroscopy. *ChemPhysChem* 21(12):1224–1229. <https://doi.org/10.1002/cphc.202000312>  
[Crossref][PubMed][PubMedCentral]
68. Wuebben C, Vicino MF, Mueller M et al (2020) Do the P1 and P2 hairpins of the guanidine-II riboswitch interact? *Nucleic Acids Res* 48(18):10518–10526. <https://doi.org/10.1093/nar/gkaa431>

[org/10.1093/nar/gkaa703](https://doi.org/10.1093/nar/gkaa703)

[Crossref][PubMed][PubMedCentral]

69. Abdullin D, Florin N, Hagelueken G et al (2015) EPR-based approach for the localization of paramagnetic metal ions in biomolecules. *Angew Chem Int Ed Engl* 54(6):1827–1831. <https://doi.org/10.1002/anie.201410396>  
[Crossref][PubMed]
70. Weber RT (2005) ELEXSYS E580 User's Manual. Bruker BioSpin Corporation, Manual Version 2.0
71. Larsen RG, Singel DJ (1993) Double electron–electron resonance spin–echo modulation: spectroscopic measurement of electron spin pair separations in orientationally disordered solids. *J Chem Phys* 98(7):5134–5146. <https://doi.org/10.1063/1.464916>  
[Crossref]
72. Ward R, Schiemann O (2013) Interspin distance determination by EPR. In: Roberts GCK (ed) Encyclopedia of biophysics. Springer, Berlin, pp 1116–1123  
[Crossref]
73. Marko A, Margraf D, Yu H et al (2009) Molecular orientation studies by pulsed electron-electron double resonance experiments. *J Chem Phys* 130(6):64102. <https://doi.org/10.1063/1.3073040>  
[Crossref]
74. Prisner TF, Marko A, Sigurdsson ST (2015) Conformational dynamics of nucleic acid molecules studied by PELDOR spectroscopy with rigid spin labels. *J Magn Reson* 252:187–198. <https://doi.org/10.1016/j.jmr.2014.12.008>  
[Crossref][PubMed]
75. Jeschke G (2007) Dipolar spectroscopy - double-resonance methods. In: Harris RK, Wasylissen RL (eds) eMagRes. Wiley, Chichester, pp 1459–1476
76. Reijerse E, Savitsky A (2007) Electron paramagnetic resonance instrumentation. In: Harris RK, Wasylissen RL (eds) eMagRes. Wiley, Chichester, pp 187–206

# **Part V**

## **Use of DNAzymes**

## 17. Sensing Metal Ions with Phosphorothioate-Modified DNAzymes

Po-Jung Jimmy Huang<sup>1</sup> and Juewen Liu<sup>1</sup>✉

(1) Department of Chemistry, Waterloo Institute for Nanotechnology,  
University of Waterloo, Waterloo, ON, Canada

✉ **Juewen Liu**  
Email: [liujw@uwaterloo.ca](mailto:liujw@uwaterloo.ca)

### Abstract

Phosphorothioate (PS) modification refers to replacing one of the nonbridging oxygen atoms in nucleic acids with sulfur. PS modifications can be easily introduced during solid-phase DNA synthesis. It has been extensively used in ribozyme and DNAzyme research to achieve a bioinorganic understanding of metal binding, bioanalytical applications of metal detection, and chemical biology of DNA modification. It allows for the access of new chemistry, not available to natural DNA. Since each PS modification is accompanied by the production of a chiral phosphorus center, a key technical challenge is to separate the two diastereomers called  $R_p$  and  $S_p$ . In this chapter, we describe our methods of HPLC-based separation followed by ligation to generate a long and fluorescently modified DNAzyme substrate. Subsequently, the use of the modified substrate for activity assay to understand metal binding and for metal ion detection is also described.

**Key words** DNAzymes – Deoxyribozymes – RNA – Phosphorothioate – Biosensors

---

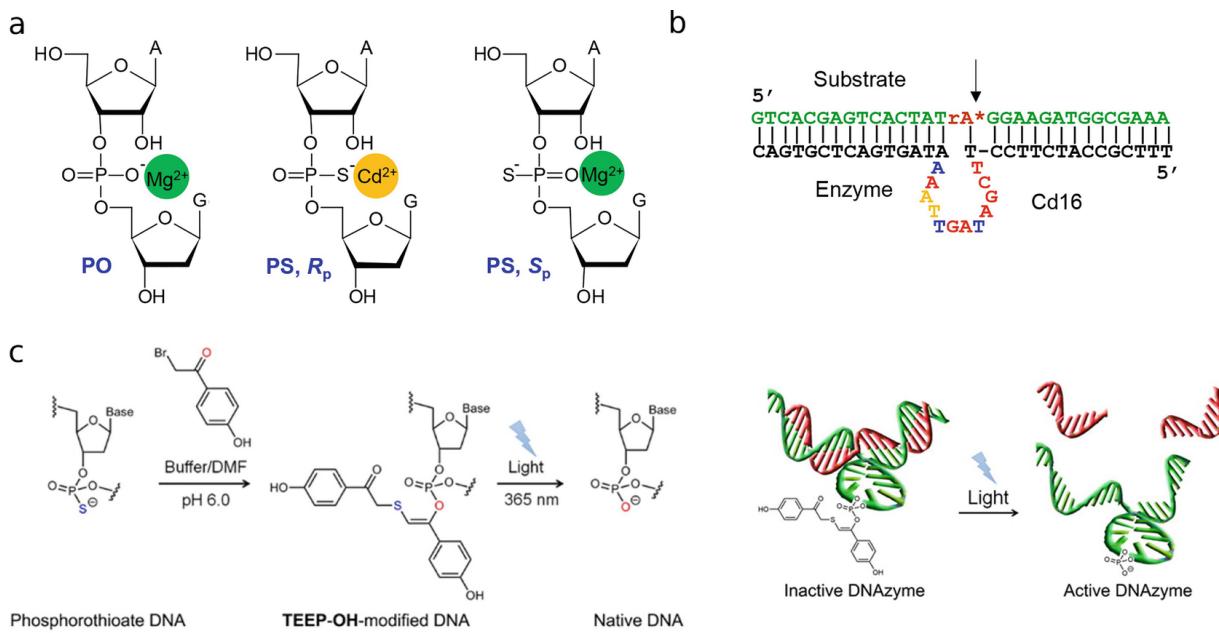
# 1 Introduction

DNAzymes are DNA-based catalysts [1]. Due to the excellent stability, ease of modification and programmability of DNA, DNAzymes have attracted extensive interest since its initial report in 1994 [2]. Among the various types of DNAzymes, those with RNA cleavage activity have been most extensively studied. They were originally developed for performing antiviral applications to selectively cleave viral RNA [3]. Later, it was realized that these DNAzymes require metal ions for activity, and some showed excellent specificity for certain metal ions [4, 5]. This has led to the selection of new DNAzymes for the detection of metal ions [6–8]. Other applications such as the use of DNAzymes for controlled drug release, assembly of stimuli-responsive materials, and building logic gates have also been explored [9–11].

Naturally, polynucleotides or ribonucleotides are linked by phosphodiester bonds. When replacing one of the nonbridging oxygen atoms in the backbone phosphate with sulfur, this is called a phosphorothioate (PS) modification [12]. The early work on PS-modified DNA aimed to increase their stability against nucleases for gene therapy applications. Later, taking advantage of the strong interactions between PS and thiophilic metal ions, PS DNA was also used for the functionalization of various metal and semiconducting nanoparticles. When a PS modification is introduced, the phosphorus becomes a chiral center [13]. Synthetically, a racemic mixture is typically obtained, and for the above applications, separation of them is not required in most cases.

For RNA-cleaving DNAzymes, PS modifications have been mainly used for three applications. First, by modifying the scissile phosphate (the cleavage site), it is possible to understand the mechanism of metal binding [14, 15]. The two nonbridging oxygens are called  $\text{pro-}R_p$  and  $\text{pro-}S_p$  in a normal phosphate linkage (Fig. 1a). Typically, a metal ion cofactor interacts with one of them during the transition state to neutralize the accumulated negative charges. So far, all known ribozymes and DNAzymes use the  $\text{pro-}R_p$  oxygen to bind metal ions, where replacing this oxygen with sulfur can result in 100-fold or more loss of activity. If such loss of activity can be rescued by adding a thiophilic metal ion, this is a strong evidence of inner-sphere metal binding [16]. On the other hand, if the  $\text{pro-}S_p$  oxygen is replaced by sulfur, the effect on activity is smaller. There are also examples

where neither of the oxygen can be replaced or the activity is nearly fully inhibited, and this may imply that both oxygen atoms are involved in metal binding [17]. We have seen some examples where two and even three metal ions were involved to activate DNAzymes, and PS modification can provide further evidence to support the metal binding mechanisms [18–20].



**Fig. 1** (a) A scheme of metal binding to a phosphodiester linkage, and  $Cd^{2+}$  binding to the sulfur and  $Mg^{2+}$  binding to oxygen in PS-modified RNA. (b) The secondary structure of the Cd16 DNAzyme uses  $Cd^{2+}$  to cleave the PS (asterisk) modified substrate [23]. (c) Grafting a photolabile modification with a PS site to cage the DNAzyme, which can be uncaged by UV light irradiation. (Reprinted from Ref. [24] with permission. Copyright © 2016 American Chemical Society)

Second, PS-modified DNAzymes can favor the detection of certain thiophilic metal ions [21]. Since stabilization of the transition state is a key role of metal ions, for a typical scissile phosphate, it is hard to find DNAzymes that can use thiophilic metals such as  $Cu^{2+}$  and  $Cd^{2+}$ . We introduced a PS modification into the library to perform in vitro selections and obtained highly specific DNAzymes for these metal ions [22]. An example of the  $Cd^{2+}$ -specific Cd16 DNAzyme is shown in Fig. 1b, and it was designed into a fluorescent sensor for the detection of  $Cd^{2+}$  [23]. Since Cd16d cleaves the  $R_p$  substrate about 100-fold faster than the  $S_p$  substrate,

separation of these two can increase the sensitivity and decrease the background of the sensor.

Third, the PS sites can be used to introduce further modifications. Phosphate is chemically quite inert, which is ideal for protecting DNA from chemical attacks. On the other hand, it makes it difficult to do chemistry on the phosphate backbone. By introducing PS modifications, it is possible to graft other chemicals. For example, Xiang and coworkers introduced a photolabile modification on the 8–17 DNAzyme (Fig. 1c) [24]. Upon light irradiation, a native DNA phosphodiester linkage was produced and the DNAzyme became active.

Compared to many other modifications on DNA, PS is very cost-effective. Each PS costs only a few extra dollars, while the cost of a thiol modification can be up to 50 times higher. Therefore, it is possible to make tandem PS modifications. For DNAzyme research, the majority of applications involve only single PS modifications. A key aspect is the separation of the two diastereomers. For short oligonucleotides, HPLC can be used for this purpose [13]. For long oligonucleotides, it is possible to synthesize two pieces separately where the shorter piece contains the PS modification. After HPLC purification, the two purified products are respectively ligated to form the full strands. Of note, synthetic chemistry has also been developed to prepare chirally pure PS oligos, both DNA and RNA. This would allow the synthesis of oligos with multiple PS modifications, for which it would be nearly impossible for HPLC to separate and assign the products.

In this chapter, we describe the detailed methods for preparing chirally pure PS-modified DNAzyme substrates using HPLC followed by ligation. Then, this substrate was used for performing biochemical assays and biosensor applications for detecting Cd<sup>2+</sup>.

---

## 2 Materials

All reagents are prepared using ultrapure water (resistivity of 18.2 MΩcm at 25 °C) and stored at 4 °C (unless indicated otherwise). DNA sequences details are listed in Table 1. Each lyophilized pellet is resuspended in ultrapure water in a concentration of 100 μM and stored at –20 °C.

**Table 1** DNA sequences for HPLC separation and sensing of Cd<sup>2+</sup>

<b>DNA name</b>	<b>Sequence and modifications</b>
HPLC-5'	5'-GAG TCA CTA TrA*G G-3'
HPLC-3'	5'-/5Phos/AAG ATG GCG AAA/36-FAM/-3'
HPLC-splint	5'-AAA AAA AAA ATT TCG CCA TCT TCC TAT AGT GAC TC-3'
PS-Sub	5'-GTC ACG AGT CAC TAT rA*GG AAG ATG GCG AAA/36-FAM/-3'
PO-Sub	5'-GTC ACG AGT CAC TAT rAGG AAG ATG GCG AAA/36-FAM/-3'
Cd16	5'-TTT CGC CAT CTT CCT TCG ATA GTT AAA ATA GTG ACT CGT GAC-3'
Q-Cd16	5'-/5IABkFQ/GTT CGC CAT CTT CCT TCG ATA GTT AAA ATA GTG ACT-3'
17E	5'-TTT CGC CAT CTT CTC CGA GCC GGT CGA AAT AGT GAC TCG TGA C-3'

Note: The cleavage site riboadenine is denoted by rA with asterisk (\*) for phosphorothioate modification. /5Phos/ is 5' phosphorylation for the ligation reaction. /36-FAM/ has 6-FAM modification at the 3'-end while /5IABkFQ/ is Iowa Black quencher labeled at the 5'-end

## 2.1 HPLC Purification of the PS Substrate

Prepare all solvents used for HPLC by passing through a 0.2 µm vacuum filtration system and degas them via 30 min sonication.

1. Agilent 1260 Infinity Quaternary HPLC system with an Inspire C18 column (5 µm, 250 mm × 4.6 mm) is used for the diastereomers separation. Elutions are monitored with a UV detector at 260 nm.
2. UV-Vis spectrophotometer to determine the concentration of small DNA volumes.
3. Solvent A: 1 L of 0.1 M ammonium acetate.
4. Solvent B: 1 L of 20% 0.1 M ammonium acetate and 80% acetonitrile.

5. A box of 100 HPLC vials.
6. 5 mL of 5 mM 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES), pH 7.5.
7. 5 mL of annealing buffer C: 10 mM MgCl<sub>2</sub>, 50 mM Tris–HCl, pH 7.5.
8. 50 mL of extraction buffer D: 1 mM EDTA, 10 mM Tris–HCl, pH 7.0.
9. 10 mL of digestion buffer E: 2 mM MgCl<sub>2</sub>, 100 mM Tris–HCl, pH 8.5.
10. T4 DNA ligase—20,000 units.
11. Phosphodiesterase I from *Crotalus adamanteus* venom.
12. Large 15% denaturing polyacrylamide gel (dPAGE).
13. Waters Sep-Pak C18 plus light cartridge.
14. 1 mL and 10 mL syringes.
15. 200 mL of 95% acetonitrile.
16. 200 mL of 1:1:1 methanol–acetonitrile–ultrapure water mixture.
17. 500 mL of ultrapure water.
18. 200 mL of 2 M ammonium acetate.
19. Liquid nitrogen or dry ice.

## 2.2 DNzyme Cleavage Activity Assays

Gel-based assays are performed with FAM labeled substrate in 15% of dPAGE.

1. Gel-imaging system with appropriate evaluation software.
2. 50 mL of reaction buffer F: 25 mM NaCl in 50 mM 2-(N-morpholino)ethanesulfonic acid (MES), pH 6.0.
3. 50 mL of reaction buffer G: 25 mM NaCl in 50 mM 3-(N-morpholino)propanesulfonic acid (MOPS), pH 7.5.
4. 5  $\mu$ M FAM-labeled DNAzyme stock: Mix 250 pmol of FAM-labeled substrate (either  $P_{\text{Mix}}^-$ ,  $R_p^-$ , or  $S_p$ -substrate) with 750 pmol of enzyme strand (Cd16 or 17E) in reaction buffer F. Anneal DNAzyme samples at 95 °C for 1 min. Then cool the complexes gradually to room temperature and store at –20 °C.
5. 15% dPAGE stock solution: For 100 mL stock, the mixture contains 48 g of urea, 37.5 mL of 40% acrylamide/bis acrylamide (29:1) and 10 mL of 10 $\times$  TBE. Gently heat and stir until the solution becomes clear. Store the stock solution in the dark at room temperature.
6. 500 mL of 1 $\times$  TBE running buffer in deionized (DI) water.
7. 10% (w/v) ammonium persulfate (APS) solution in ultrapure water: Prepare 1 mL of solution and aliquot into multiple 0.6 mL microcentrifuge tubes before storing at –20 °C. For most effective polymerization, the solution will need to be used up within few days.
8. *N,N,N,N'*-Tetramethyl-ethylenediamine (TEMED) is stored at 4 °C.
9. 10 mL of 8 M urea dissolved in ultrapure water.
10. 10 mL Syringe with needle.
11. 1 $\times$  Gel loading dye solution: Mix 200  $\mu$ L of 6 $\times$  blue gel loading dye with 1000  $\mu$ L 8 M urea in 1.75 mL microcentrifuge tube. This solution is used to quench the enzymatic reaction.

## 2.3 Sensing of Cd<sup>2+</sup>

The biosensor assays are performed using DNAzyme complex containing purified FAM-labeled  $R_p$ -substrate and quencher-labeled enzyme.

1. Microplate reader with wavelength  $Ex = 485$  nm;  $Em = 520$  nm; cut-off = 515 nm at 25 °C. The sample plate is monitored continuously for 1 h with time intervals set for every 20 s.
  2. Sensing buffer F: 50 mL of 1 mM HEPES buffer, pH 7.5.
  3. 5 µM FAM-Q DNAzyme sensor stock: Mix 250 pmol of full-length  $R_p$ -substrate with 500 pmol of Q-Cd16 in reaction buffer F and anneal at 95 °C for 1 min. The sensor complex is then gradually cooled to room temperature and stored at -20 °C.
- 

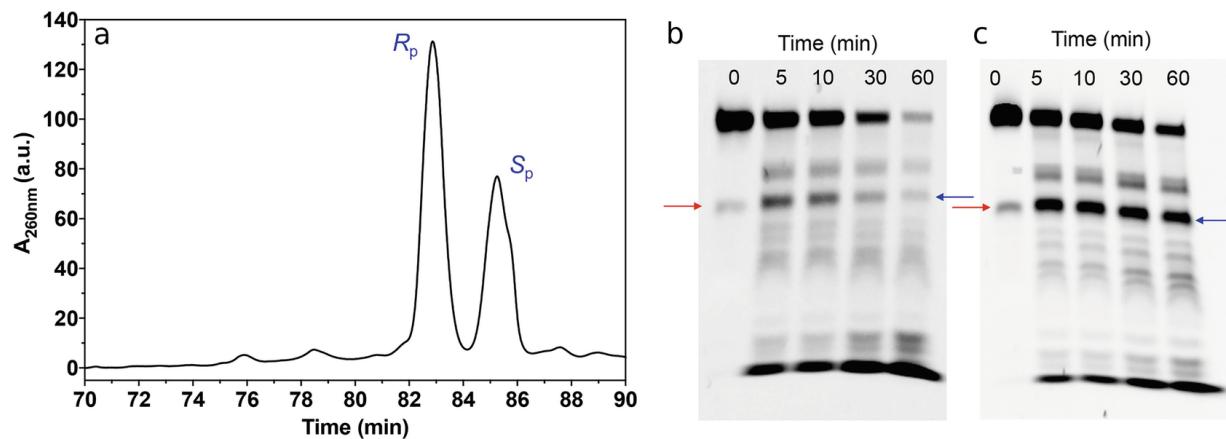
## 3 Methods

### 3.1 Separation of PS Diastereomers

#### 3.1.1 HPLC Purification of the PS Substrate

1. Transfer 100 µM of HPLC-5' into HPLC vial containing micro volume insert. Then place vial in sampler with temperature set at 4 °C.
2. Set the autosampler to inject 20 µL per run with a flow rate of 1 mL/min. Set column temperature at 45 °C to improve the resolution of the diastereomers.
3. In the solvent system set up, 97% of solvent A and 3% of solvent B are used from time 0 to 10 min. From 10 to 90 min, the solvent A is gradually decreased to 90% while solvent B is increased to 10%.
4. To collect the fraction, an initial test run is required to define the proper parameters like peak slope, peak duration, and threshold (see Note 1). The system automatically collects  $R_p$  and  $S_p$  DNA in separate HPLC vials (Fig. 2a).  
After multiple runs, combine the identical fractions and lyophilize

- overnight. Resuspend the dry pellets in 5 mM HEPES (pH 7.5) and determine the concentrations using a spectrophotometer.



**Fig. 2** (a) A representative HPLC trace of diastereomer separation with two well-resolved peaks (see Note 2). (b, c) Gel micrographs showing snake venom phosphodiesterase digestion of the separation products. Since the phosphodiesterase cleaves the *S<sub>p</sub>* diastereomer less efficiently compared to the *R<sub>p</sub>* diastereomer [25], and the first elution (b) is digested more quickly than the second elution (c), (b) is assigned as *R<sub>p</sub>* and (c) as *S<sub>p</sub>*. The blue arrowhead corresponds to the rA\*G junction position where enzymatic digestion stops, and the red arrow corresponds to the cleavage product. This red band runs slightly faster than the blue band, suggesting that these two bands differ by just a single nucleotide. (Adapted from Ref. [23])

### 3.1.2 Formation of Full-Length *R<sub>p</sub>* and *S<sub>p</sub>* Substrate

- To form the full-length *R<sub>p</sub>*- and *S<sub>p</sub>*- substrates, perform ligation reaction (Fig. 3a). Mix 200 pmol of HPLC-5', 300 pmol of HPLC-3' and 300 pmol HPLC-splint in 10 µL buffer C. Then anneal mixture at 95 °C for 1 min and gradually cool to room temperature.
- Follow the T4 ligation protocol provided by the vendor; carry out the reaction for 1 h at room temperature. Quench the reaction with equal volume of 1× gel loading dye.
- Purify the DNA product with large 15% dPAGE at 650 V for 90 min. For a 16.5× 28 cm gel plate, add 50 µL of 10% APS and 20 µL of TEMED in 22 mL of 15% gel stock in a flask. Swirl the flask for each addition to ensure thorough mixing. Note that polymerization has begun as soon as TEMED is added.

4.

After polymerization, secure the gel plates in the chamber and fill the upper reservoir with 1× TBE to cover the wells. Also, fill the lower reservoir with enough 1× TBE to cover bottom of the plates.

5.

Remove the comb and wash the wells a few times with the syringe to remove the excess unpolymerized mixture. Caution: without washing the wells, samples will sit unevenly between the wells after loading.

6.

Prerun the gel for at least 30 min.

7.

Wash the wells again before loading the samples. Run the gel at a constant 650 V for 90 min.

8.

Visualize the FAM-labeled DNA under an imaging system (Fig. 3b). Extract the ligated DNA from the gel with a razor blade and place it in the 1.75 mL microcentrifuge tube.

9.

Crush the excised gel and extract with a total of 1 mL buffer C for 2 h. Note that crushing the gel into a fine paste before adding 250 µL buffer and incubating the extraction solution at 60 °C for 30 min can facilitate DNA extraction process.

10.

Centrifuge the sample at 21,300 × *g* for 10 min and collect the supernatant. Repeat the incubation and centrifugation steps at least 3 more times or until no visible fluorescence is observed in supernatant.

11. Desalt the combined supernatant with a Sep-Pak C18 cartridge as shown in Fig. 4.

(a)

Condition the cartridge with a 10 mL syringe in the following order before loading the extracted DNA into the cartridge (see **Note 3**):

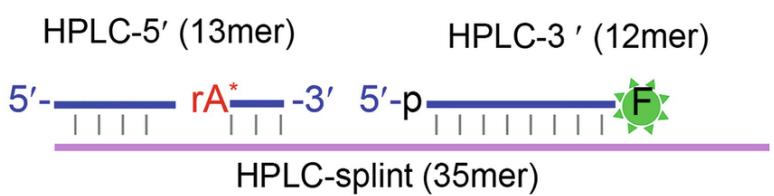
- 10 mL 95% acetonitrile
- 10 mL 1:1:1 acetonitrile–methanol–ultrapure water
- 20 mL ultrapure water
- 10 mL 2 M ammonium acetate

- TO THE -3' AMINATED ACCURATE
- (b) Pass the extracted DNA through the cartridge using a 1 mL syringe at least three times. Discard the passthrough at the end.
  - (c) Wash the cartridge with 20 mL water to remove impurities and salt.
  - (d) Elute the purified DNA with 2 mL of 1:1:1 mixture of acetonitrile–methanol–water. Collect the elute in two 1.5 mL microcentrifuge tubes.
  - (e) Flash-freeze the collected DNAs with liquid nitrogen and lyophilize the samples overnight.

12.

Redissolve the dried DNA substrate pellets in 20  $\mu$ L 5 mM HEPES, pH 7.5 and determine the concentration using a spectrophotometer.

a

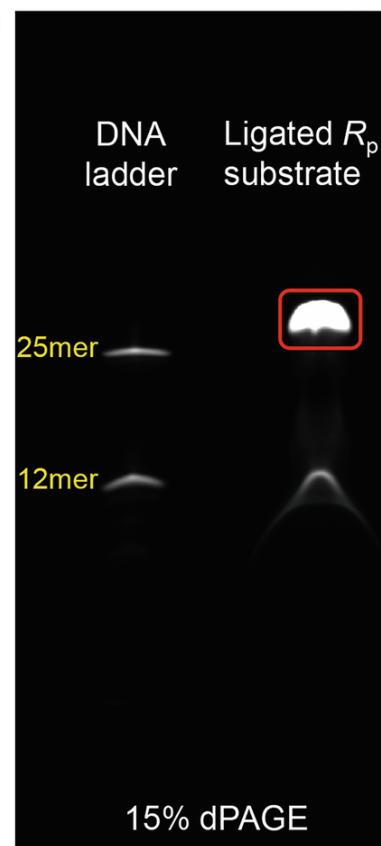


T4 DNA ligase

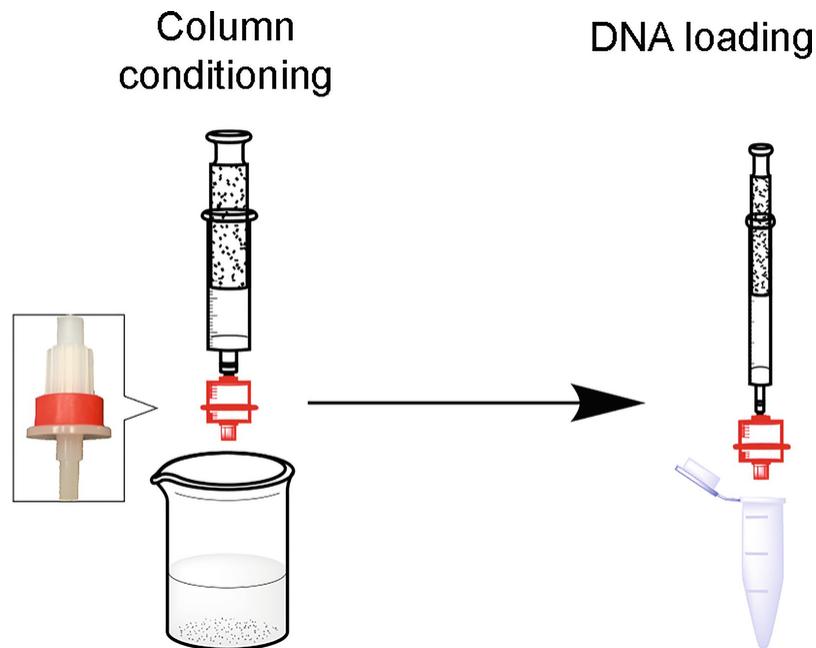
Full-length PS-substrate (25mer)



b



**Fig. 3** (a) The full-length PS-substrate is prepared via ligation using T4 DNA ligase. (b) A gel micrograph showing full-length  $R_p$ -substrate containing FAM (red box) and un-ligated HPLC-3' (12mer). Although complementary strand (HPLC-splint) is not labeled, it is 10 bases longer than the substrate. Thus, the full-length  $R_p$ -substrate is well separated from the other DNA strands in 15% dPAGE and can be extracted easily



**Fig. 4** Scheme of DNA desalting and purification using a C18 Sep-Pak cartridge. The cartridge is first conditioned via a 10 mL syringe. The DNA samples are then loaded with a 1 mL syringe. The column is then washed with water, and the DNA is eluted with 1:1:1 acetonitrile–methanol–water

### 3.1.3 Assignment of $R_p$ and $S_p$ Diastereomers by Enzyme Digestion

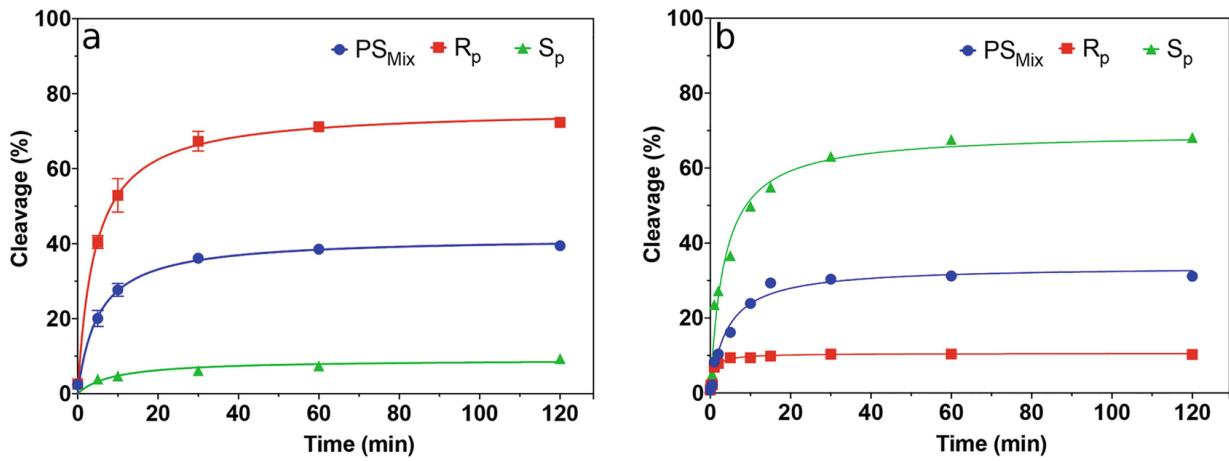
Once the two peaks are separated via HPLC,  $R_p$  form is eluted first followed by the  $S_p$  (Fig. 2a). Their configurations can also be confirmed by snake venom phosphodiesterase I digestion.

1. Transfer 50 pmol of full-length  $R_p$ - or  $S_p$ -substrate in 50  $\mu$ L buffer E.
2. Add 1  $\mu$ L of 0.01 unit/mL venom phosphodiesterase I and incubate at 37 °C up to an hour using PCR instrument.
3. At each designated time point, take 10  $\mu$ L of the digestion sample and

- quench with 1× gel loading buffer.
- Samples are separated in 15% dPAGE at a constant 200 V for 1 h and the gels are analyzed in an imaging system (Fig. 2b, c).

### 3.2 DNAzyme Cleavage Activity Assays

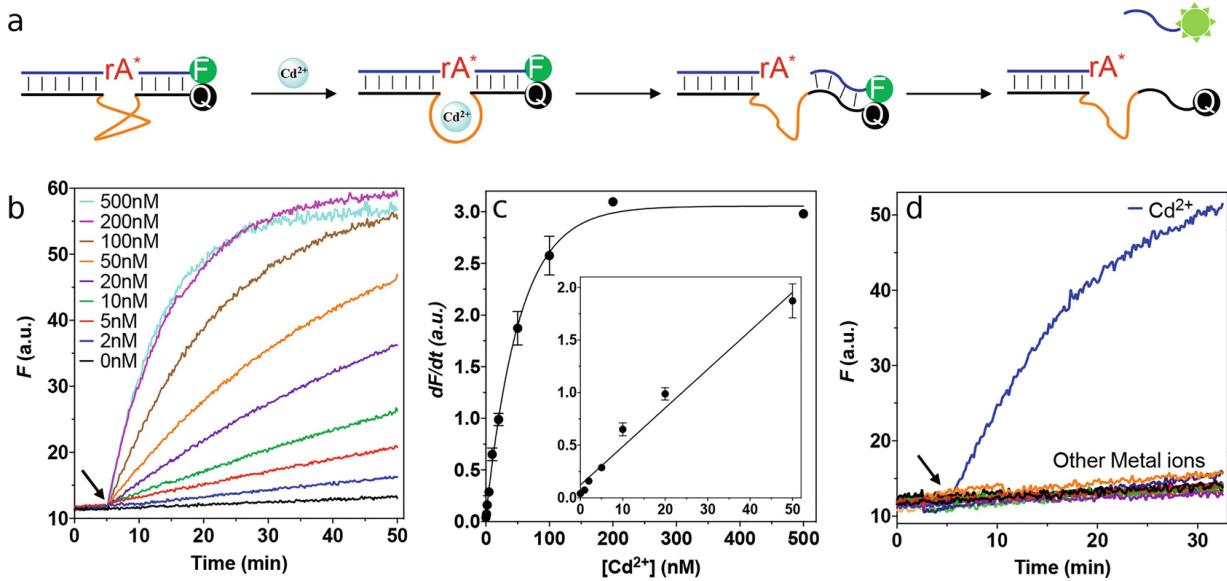
- For each 7 µL reaction, dilute the 5 µM FAM-labeled DNAzyme in buffer F into 0.7 µM with 4 µL of an appropriate reaction buffer, buffer F for Cd16 and buffer G for 17E (*see Note 4*).
- Add 2 µL of metal ion to initiate the reaction at the designated time point.
- Add 7 µL of 1× gel loading dye at the end of the incubation period. Place samples on an ice bath until it is time for electrophoresis.
- For a set of 10 × 10 cm gel plates, add 64 µL of 10% APS and 10 µL of TEMED in 14 mL of 15% gel stock in a flask. Swirl the flask for each addition to ensure thorough mixing.
- Following the same steps as setting up large gel described above and allow the gel to polymerize for approximately 45 min.
- Wash the wells and prerun the gel for at least 30 min.
- Rewash the wells again and load the samples. Run the gel at constant 200 V for 80 min.
- Visualize the FAM-labeled DNA under an imaging system and analyze the percentage of the cleaved bands using appropriate software (Fig. 5a, b).



**Fig. 5** Cleavage kinetics of the un-purified  $\text{PS}_{\text{Mix}}$ -, and purified  $R_p$ -, and  $S_p$ -substrate by (a) Cd16 with  $10 \mu\text{M}$   $\text{Cd}^{2+}$  in buffer F; or (b) 17E with  $10 \text{ mM}$   $\text{Mg}^{2+}$  in buffer G at room temperature.  $\text{Cd}^{2+}$  has high affinity toward sulfur. Cd16 cleaves the  $R_p$ -substrate  $\sim 100$ -fold faster than the  $S_p$ -substrate. With  $\text{Mg}^{2+}$ , 17E has higher cleavage activity for the  $S_p$ -substrate. In both cases, the kinetics trace of the un-separated  $\text{PS}_{\text{Mix}}$ -substrate is cleaved between the two separated ones. (Adapted from Ref. [23])

### 3.3 Sensing of $\text{Cd}^{2+}$

1. Before sensing, dilute a portion of the FAM-Q DNAzyme (Fig. 6a) stock  $100\times$  with buffer F to form a  $50 \text{ nM}$  sensor solution (see Note 5).
2. Pipette  $100 \mu\text{L}$  of the sensor into each sample well in a 96-well microplate. Monitor continuously the background fluorescence signal of the sensor until the sensor is equilibrated at  $25^\circ\text{C}$ . Typically, it takes 10–15 min.
3. Once the sensor is stabled, collect the background signal kinetics for an additional 5 min to establish a baseline (Fig. 6b, d).
4. Add  $1 \mu\text{L}$  metal ion to initiate cleavage and continuously measure the signaling kinetics for additional 45 min.



**Fig. 6** (a) Schematic design of the fluorescent DNAzyme sensor for Cd<sup>2+</sup> detection. (b) Fluorescence kinetics of Cd16 sensor containing  $R_p$ -substrate in the presence of various Cd<sup>2+</sup> concentration. (c) A plot of initial rate of fluorescence enhancement (from 1 to 10 min after Cd<sup>2+</sup> addition) as a function of Cd<sup>2+</sup> concentration. Inset: the linear response of the sensor at low Cd<sup>2+</sup> concentration. (d) Sensor response to 100 nM of various divalent and trivalent metals. (Adapted from Ref. [23])

## 4 Notes

- If the fraction collection parameters are not set properly or when the HPLC trace shows that the two diastereomers are not well separated (e.g., the bases of the peaks are overlapping), there are usually a small fraction of  $R_p$  presence in the  $S_p$  collection. In this case, a second HPLC purification can be performed on the  $S_p$  collection under the same condition to achieve even higher purity.
- Although it was reported by numerous groups that  $R_p$  usually elutes before the  $S_p$  [25, 26], the elution order may be reversed if phosphorothioate modification is placed at different locations or other modifications are also incorporated into the sequence design [27, 28]. Enzymatic digestion or other assays are necessary for confirmation of the diastereomers, especially for the first time of separation.
- During the cartridge rehydration and DNA loading steps, it is critical to

slowly push the solvent through to achieve dropwise flow and avoid pushing air into the cartridge. In addition, detach the syringe from the cartridge before removing the plunger to avoid disturbing the column.

4. The 17E DNAzyme is very active with many divalent metal ions in pH 7.5 buffers. Preparing the 17E DNAzyme stock at pH 6 in buffers strictly devoid of divalent metal ions is recommended, which can drastically reduce the unintended cleavage reaction and lower background signal can be achieved. For example, 10 concentrated DNAzyme can be annealed in MES buffer, pH 6.0. When ready to use, it can be diluted tenfold to 25 mM pH 7.5 buffer.
  5. As shown in Fig. 5a, only ~50% of substrate can be cleaved when a racemic mixture is used. Since  $R_p$ -substrate is ~100-fold more active than the  $S_p$ -substrate, it is ideal to use purified  $R_p$ -substrate to prepare the Cd16 sensor complex to achieve maximum fluorescence enhancement for sensing.
- 

## References

1. Silverman SK (2016) Catalytic DNA: scope, applications, and biochemistry of deoxyribozymes. *Trends Biochem Sci* 41(7):595–609  
[Crossref]
2. Breaker RR, Joyce GF (1994) A DNA enzyme that cleaves RNA. *Chem Biol* 1(4):223–229  
[Crossref]
3. Santoro SW, Joyce GF (1997) A general purpose RNA-cleaving DNA enzyme. *Proc Natl Acad Sci U S A* 94(9):4262–4266  
[Crossref]
4. Li J, Lu Y (2000) A highly sensitive and selective catalytic DNA biosensor for lead ions. *J Am Chem Soc* 122(42):10466–10467  
[Crossref]
5. Hwang K, Hosseinzadeh P, Lu Y (2016) Biochemical and biophysical understanding of metal ion selectivity of DNAzymes. *Inorg Chim Acta* 452:12–24  
[Crossref]

6. Zhou W, Saran R, Liu J (2017) Metal sensing by DNA. *Chem Rev* 117:8272–8325  
[[Crossref](#)]
7. Lake RJ, Yang ZL, Zhang JL, Lu Y (2019) DNAzymes as activity-based sensors for metal ions: recent applications, demonstrated advantages, current challenges, and future directions. *Acc Chem Res* 52(12):3275–3286  
[[Crossref](#)]
8. Huang P-JJ, Liu J (2021) In vitro selection and application of lanthanide-dependent DNAzymes. *Methods Enzymol* 650:373–396  
[[Crossref](#)]
9. Liu M, Chang D, Li Y (2017) Discovery and biosensing applications of diverse RNA-cleaving DNAzymes. *Acc Chem Res* 50(9):2273–2283  
[[Crossref](#)]
10. Moon WJ, Liu J (2020) Interfacing catalytic DNA with nanomaterials. *Adv Mater Interfaces* 7:2001017  
[[Crossref](#)]
11. Peng H, Newbigging AM, Wang Z, Tao J, Deng W, Le XC, Zhang H (2018) DNAzyme-mediated assays for amplified detection of nucleic acids and proteins. *Anal Chem* 90(1):190–207  
[[Crossref](#)]
12. Saran R, Huang Z, Liu J (2021) Phosphorothioate nucleic acids for probing metal binding, biosensing and nanotechnology. *Coord Chem Rev* 428:213624  
[[Crossref](#)]
13. Frederiksen JK, Piccirilli JA (2009) Separation of RNA phosphorothioate oligonucleotides by HPLC. *Methods Enzymol* 468:289–309  
[[Crossref](#)]
14. Forconi M, Herschlag D (2009) Use of phosphorothioates to identify sites of metal-ion binding in RNA. *Methods Enzymol* 468:311–333  
[[Crossref](#)]
15. Thaplyal P, Ganguly A, Hammes-Schiffer S, Bevilacqua PC (2015) Inverse thio effects in the hepatitis delta virus ribozyme reveal that the reaction pathway is controlled by metal ion charge density. *Biochemistry* 54(12):2160–2175  
[[Crossref](#)]
16. Misra VK, Draper DE (1998) On the role of magnesium ions in RNA stability. *Biopolymers* 48(2–3):113  
[[Crossref](#)]

17. Zhou W, Liu J (2018) Multi-metal-dependent nucleic acid enzymes. *Metallooms* 10:30–48  
[[Crossref](#)]
18. Zhou W, Saran R, Huang P-JJ, Ding J, Liu J (2017) An exceptionally selective DNA cooperatively binding two  $\text{Ca}^{2+}$  ions. *Chembiochem* 18(6):518–522  
[[Crossref](#)]
19. Zhou W, Zhang Y, Huang P-JJ, Ding J, Liu J (2016) A DNAzyme requiring two different metal ions at two distinct sites. *Nucleic Acids Res* 44:354–363  
[[Crossref](#)]
20. Huang P-JJ, Vazin M, Matuszek Ž, Liu J (2015) A new heavy lanthanide-dependent DNAzyme displaying strong metal cooperativity and unrescuable phosphorothioate effect. *Nucleic Acids Res* 43(1):461–469  
[[Crossref](#)]
21. Huang P-JJ, Liu J (2020) In vitro selection of chemically modified DNAzymes. *ChemistryOpen* 9:1046–1059  
[[Crossref](#)]
22. Huang P-JJ, Liu J (2016) An ultrasensitive light-up  $\text{Cu}^{2+}$  biosensor using a new DNAzyme cleaving a phosphorothioate-modified substrate. *Anal Chem* 88:3341–3347  
[[Crossref](#)]
23. Huang P-JJ, Liu J (2015) Rational evolution of  $\text{Cd}^{2+}$ -specific DNAzymes with phosphorothioate modified cleavage junction and  $\text{Cd}^{2+}$  sensing. *Nucleic Acids Res* 43(12):6125–6133  
[[Crossref](#)]
24. Wang X, Feng M, Xiao L, Tong A, Xiang Y (2016) Postsynthetic modification of DNA phosphodiester backbone for photocaged DNAzyme. *ACS Chem Biol* 11(2):444–451  
[[Crossref](#)]
25. Burgers PMJ, Eckstein F (1978) Absolute configuration of the diastereomers of adenosine 5'-O-(1-thiotriphosphate): consequences for the stereochemistry of polymerization by DNA-dependent RNA polymerase from *Escherichia coli*. *Proc Natl Acad Sci U S A* 75(10):4798–4800  
[[Crossref](#)]
26. Slim G, Gait MJ (1991) Configurationally defined phosphorothioate-containing oligoribonucleotides in the study of the mechanism of cleavage of hammerhead

ribozymes. *Nucleic Acids Res* 19(6):1183–1188

[[Crossref](#)]

27. Stec WJ, Zon G (1984) Synthesis, separation, and stereochemistry of diastereomeric oligodeoxyribonucleotides having a 5'-terminal internucleotide phosphorothioate linkage. *Tetrahedron Lett* 25(46):5275–5278  
[[Crossref](#)]
28. Stec WJ, Zon G, Egan W (1984) Automated solid-phase synthesis, separation, and stereochemistry of phosphorothioate analogs of oligodeoxyribonucleotides. *J Am Chem Soc* 106(20):6077–6079  
[[Crossref](#)]

# 18. Enhancement of Activity for Peroxidase-Mimicking DNAzyme by Covalent Attachment of Hemin to G-Quadruplex-Forming Oligonucleotide Using Click Chemistry

Joanna Kosman<sup>1</sup>✉

(1) Department of Bioanalytical Chemistry, Faculty of Chemistry, Adam Mickiewicz University, Poznań, Poland

✉ Joanna Kosman  
Email: [kosman@amu.edu.pl](mailto:kosman@amu.edu.pl)

## Abstract

Peroxidase-mimicking DNAzymes are used in the development of new bioanalytical assays due to their advantages like thermal stability, simple synthesis and purification, and ability to hybridize with the complementary strand of nucleic acid. Here, we describe the method of covalent attachment of hemin to DNA oligonucleotide using click chemistry that allows good yield (60–70%) of the final conjugate product. The activity of obtained DNAzymes is monitored using chromogenic and fluorogenic substrates.

**Key words** DNAzyme – G-quadruplex – Hemin – Peroxidase activity – Click chemistry

---

## 1 Introduction

DNAzymes are single stranded DNA molecules able to catalyze chemical reaction [1]. Several DNAzymes with various activity have been developed but one that found the widest application is peroxidase-mimicking DNAzyme [2]. This DNAzyme catalyzes the reaction between hydrogen peroxide and organic substrate chosen to produce colorimetric, chemiluminescent, or fluorescent signal. It can successfully replace horseradish peroxidase in many applications and new assays are developed thanks to many advantages of this DNAzymes in comparison to protein enzymes, like thermal stability, simple synthesis, purification, and modification as well as hybridization between complementary strands [3]. Peroxidase-mimicking DNAzyme is composed of a G-quadruplex-forming oligonucleotide and a hemin molecule. The resulting complex is stabilized by end-stacking interaction between the porphyrin ring of hemin and the G-quartet of a G-quadruplex, formed by four guanosines connected by Hoogsteen hydrogen bonds [4]. Analytical methods based on peroxidase-mimicking DNAzymes were developed for detection of DNA sequences, single nucleotide polymorphism, proteins, metal ions, and other small molecules [5].

The research on peroxidase-mimicking DNAzymes extends in two main directions: (a) development of new applications and (b) enhancement of DNAzyme activity and stability. The second trend involves modification of DNAzyme and is focused on the variation of DNA oligonucleotide sequence [6], the usage of DNA analogues [7, 8] and the covalent attachment of hemin to DNA oligonucleotide [9–13]. Covalent attachment of hemin to DNA oligonucleotide (DNAzyme conjugate) on the one hand enhances its peroxidase activity and on the other hand decreases the signal of blank probe, which in the case of conventional association DNAzyme may approach high value. In conventional DNAzyme system free hemin is present and can alone catalyze oxidation of the substrate. Additionally, attachment of hemin can be helpful in development of heterogenic assays or applications in cells, because it prevents hemin dissociation from DNAzyme complex. The first attempts of covalent attachment of hemin to DNA oligonucleotide were performed using N-hydroxysuccinimide ester of hemin and amino-modified oligonucleotide [9–12]. Two approaches were developed using this method: post-synthetic and on-column conjugation. The main drawback of this method of hemin-DNA oligonucleotide conjugation is low reaction yield (ca. 10%) [10]. Our group developed the

method of hemin conjugation to DNA oligonucleotide using click reaction [13]. Synthesis of hemin-DNA conjugates through Cu-free strain-promoted alkaline-azide cycloaddition proceeded with 60–70% yield. Here we present the protocol for conjugation of hemin to DNA oligonucleotide using click chemistry and examples of monitoring catalytic activity using three substrates: 2,2'-azinobis(3-ethylbenzothiazoline)-6-sulfonic acid (ABTS, chromogenic), Amplex Red, and 4-(N-methylhydrazino)-7-nitrobenzofurazan (MNBDH, fluorogenic).

---

## 2 Materials

All solution should be prepared using ultrapure water (Milli-Q quality, resistivity of  $18 \text{ M}\Omega\cdot\text{cm}$  at  $25^\circ\text{C}$ ). All reagents should be of analytical grade.

### 2.1 Synthesis of Hemin Azide

1. Hemin solution in anhydrous dimethylformamide (DMF): Dissolve 65.2 mg of hemin (>97% purity) in 0.8 ml of anhydrous DMF in a 1.5 mL reaction tube.
2. *N,N,N',N'*-tetramethyl-O-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate: Store at  $4^\circ\text{C}$ .
3. *N,N*-diisopropylethylamine: Store at  $4^\circ\text{C}$ .
4. Azide-(polyethylene glycol)<sub>3</sub>-amine: Store at  $4^\circ\text{C}$  (*see Note 1*).
5. Methylene chloride.
6. Silica gel 60, 70–230 mesh American Standard Sieve Series (ASTM).
7. Solvents for gradient column separation: hexane; hexane–ethyl acetate mixtures (1:9, 1:4, 3:7, 2:3, 1:1, 3:2, 7:3, 4:1, 9:1); ethyl acetate; ethyl acetate–methylene chloride mixtures (1:9, 1:4, 3:7, 2:3, 1:1, 3:2, 7:3, 4:1, 9:1).
8. Methanol–Methylene chloride mixtures v/v from 0.1–3%. Prepare the

solutions by mixing appropriate volumes of methanol and methylene chloride with 0.1% increase. For 20 mL of 0.1% solution mix

0.02 mL of methylene chloride and 19.98 mL of methanol. For each next solution increase the methylene chloride volume by 0.02 mL and make up to 20 mL.

9. Dimethyl sulfoxide (DMSO).
10. Magnetic stirrer.
11. Plates for thin layer chromatography (TLC).
12. Argon gas.
13. Ice cube.
14. Anhydrous sodium sulfate.
15. Funnel for gravity filtration, glass flask.
16. Filter paper.
17. Glass column for column chromatography.
18. Rotary evaporator.
19. Lyophilizer.

## 2.2 Click Reaction

1. Dissolve dibenzocyclooctyl-modified oligonucleotide in water to obtain 150  $\mu$ M solution (volume of the water depends on the amount of oligonucleotide provided by the company). Store at  $-30^{\circ}\text{C}$  (*see Note 2*).
2. Hemin azide. Store at  $-30^{\circ}\text{C}$  (*see Note 3*, Subheading 3.1).

3. Plastic tube with volume of 1.5 mL.
4. Round-bottom glass flask.
5. Thermomixer for 1.5 ml plastic tubes.
6. Preparative ODS (octadecylsilyl or C18) column for HPLC.
7. Solutions for HPLC chromatography: A: 95% 0.1 M triethylammonium acetate (TEAA) in water pH 7.4, 5% acetonitrile (ACN), B: 20% 0.1 M TEAA pH 7.4, 80% ACN.
8. HPLC instrument equipped with heater and UV-Vis detector (diode array).

## 2.3 Catalytic Reaction Catalyzed by DNAzyme

1. Reaction buffer: 100 mM Tris–HCl pH 8.0. Weigh 0.6055 g Tris(hydroxymethyl)aminomethan (Tris) base and transfer to a glass beaker. Add 45 mL of water and mix the solution to complete dissolution of Tris. Adjust pH with 1 M HCl. Make up to 50 mL in volumetric flask and store at 4 °C.
2. 1 M KCl solution: Weigh 3.7275 g of KCl and transfer to volumetric flask and make up to 50 mL.
3. Hemin solution: Weigh 0.0065 g of hemin and transfer it to flask and make up to 1 mL with DMSO. As-prepared 0.01 M solution dilute next to 0.1 mM solution by withdrawing 10 µL to separate tube and adding to it 990 µL of DMSO (*see Note 3*).
4. 10 mM Hydrogen peroxide: Place 10 µL of 30% H<sub>2</sub>O<sub>2</sub> in a volumetric flask and add water to 10 mL (*see Note 4*).
5. 10 mM ABTS: Weigh 0.0051 g of ABTS and transfer to a 1.5 mL reaction tube. Add 1 mL of water (*see Note 4*).

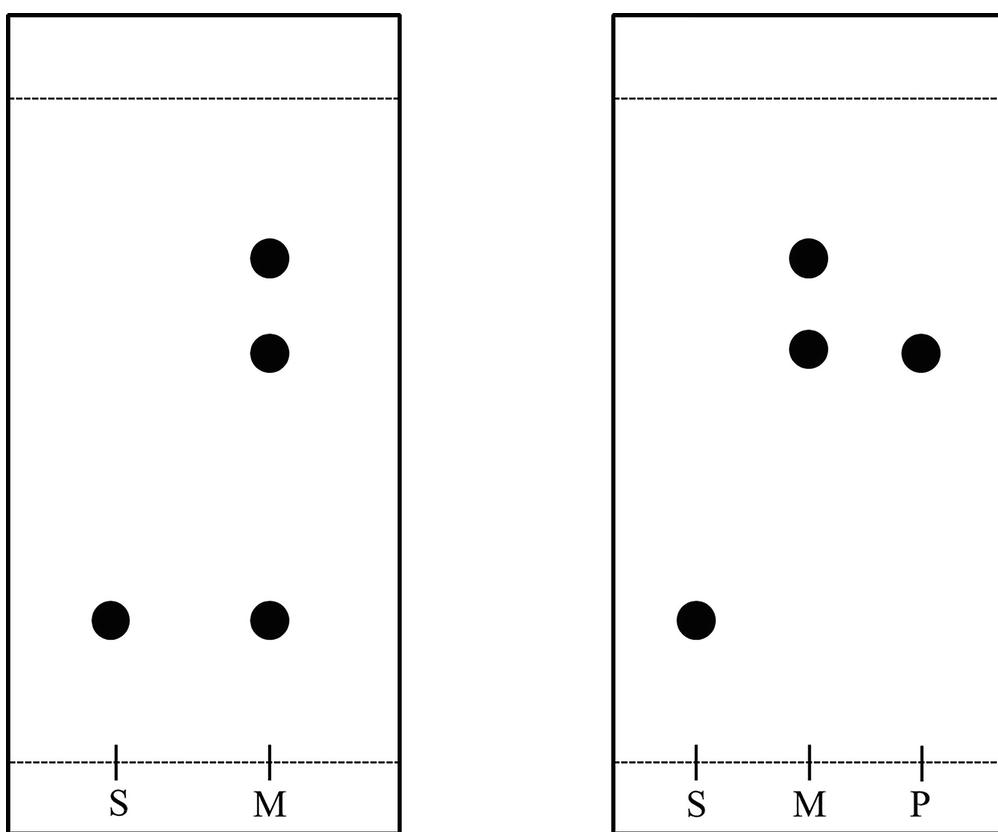
6. Plastic tubes 0.5 mL and 1.5 mL volumes.
  - 7.
  - Thermoblock for 0.5 mL plastic tubes.
  8. Clear microplates.
  9. Microplate reader.
- 

### 3 Methods

#### 3.1 Hemin Azide Synthesis

1. Add 0.8 mL of hemin solution in DMF, 0.0417 g of *N,N,N',N'*-tetramethyl-O-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate and 0.036 mL of *N,N*-diisopropylethylamine to a round-bottom glass flask. Add magnetic stirrer to the solution.
2. Degas the solution with argon for 5 min by bubbling argon gas through solution. Incubate the solution while stirring for 30 min at room temperature.
3. Add 20  $\mu$ L of azide-(polyethylene glykol)<sub>3</sub>-amine to the solution. Degas the solution for 5 min with Argon. Stir the reaction mixture for 24 h at room temperature. Control the reaction progress with thin layer chromatography (TLC) (Fig. 1, see Note 5).
4. Pour the mixture on ice cube placed in a beaker (see Note 6). After melting of the ice cube filtrate the mixture using the set under reduced pressure using a filter paper.
5. Dissolve the precipitate in 20 mL of methylene chloride and add as drying agent anhydrous sodium sulfate (see Note 7). After 15 min filter the solution by gravity through a funnel with filter paper. Transfer filtrate to a round-bottom glass flask and evaporate solvent under reduced pressure.

6. Purification of the product: Prepare column for column chromatography. Secure the column outlet with cotton wool. Add hexane on the column. During the very slow flow of the solvent from the column, add slurry of silica gel in hexane (1:50) and wait for the settlement of silica (*see Note 8*).
7. Dissolve hemin azide in minimal volume of methylene chloride and add on the column. Rinse the column with 20 mL portions of the following solvents: hexane, ethyl acetate–hexane (1:9, 1:4, 3:7, 2:3, 1:1, 3:2, 7:3, 4:1, 9:1), ethyl acetate, methylene chloride–ethyl acetate (1:9, 1:4, 3:7, 2:3, 1:1, 3:2, 7:3, 4:1, 9:1), methylene chloride. Rinse the column with 20 mL of eluent containing increasing percentage of methanol in methylene chloride starting with 0.1% and increasing the methanol concentration with 0.1% steps.
8. Monitor the collected fractions with TLC. Hemin azide is typically obtained from fraction containing 1.6–2.5% of methanol in methylene chloride. Combine the fractions of similar purity and evaporate the solvents. Lyophilize the product and store at –30 °C.

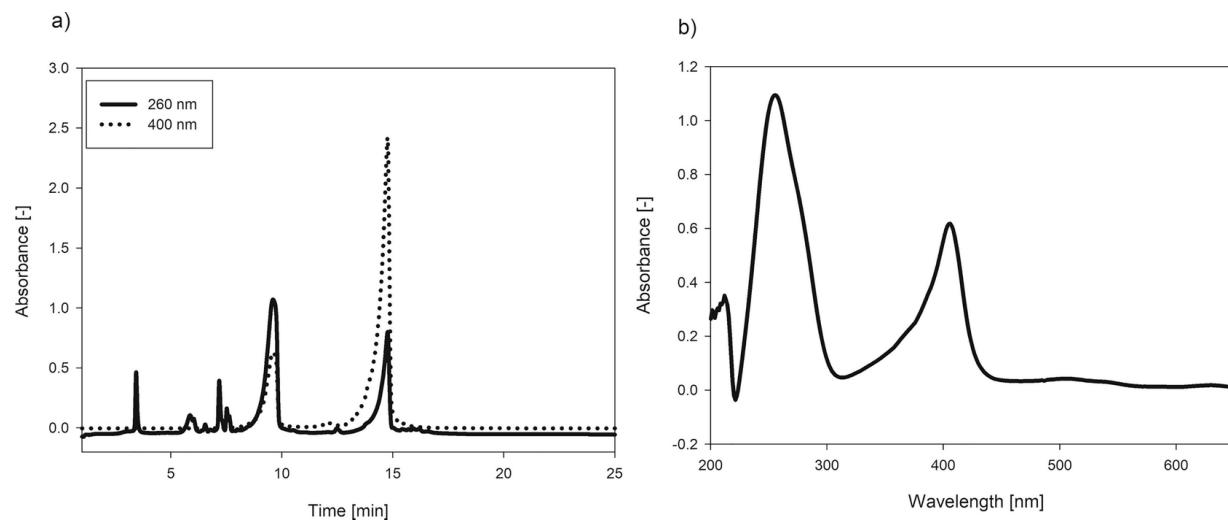


**Fig. 1** The scheme of TLC plates performed to observe hemin azide synthesis progress. (a) Analysis of the reaction mixture, (b) Comparison of purified product with substrate and reaction mixture. Designations: S-substrate (hemin), M-mixture, P-product (hemin azide)

### 3.2 Click Reaction

1. Add 30  $\mu$ L of 150  $\mu$ M DNA oligonucleotide solution to the 1.5 mL reaction tube. Add 10  $\mu$ L of 30 mM hemin azide solution. Add 200  $\mu$ L of DMSO. Incubate the tube for 24 h at 25 °C with stirring (800 rpm) in thermomixer.
2. Purification of hemin-DNA oligonucleotide using HPLC equipped with heater and UV-Vis detector on an octadecylsilyl (ODS) column (*see Note 9*). Use gradient method for purification (0 min 100% A; 5 min 80% A, 20% B, 15 min 100% B, 25 min 100% A). Set the heater at 55 °C. After the injection of, monitor the separation by UV-Vis detector and collect fractions containing the product (two absorption bands at around 260 and 400 nm). An example of chromatogram and UV-Vis spectrum of product shown on Fig. 2.

3. Combine all fractions, evaporate the solvents and lyophilize the product.
  
4. Dissolve the product in 50  $\mu$ L of water and determine the concentration of DNA using UV-Vis spectrophotometry at 260 nm [14]. Dilute the solution to obtain 100  $\mu$ M solution of hemin-DNA oligonucleotide conjugate.



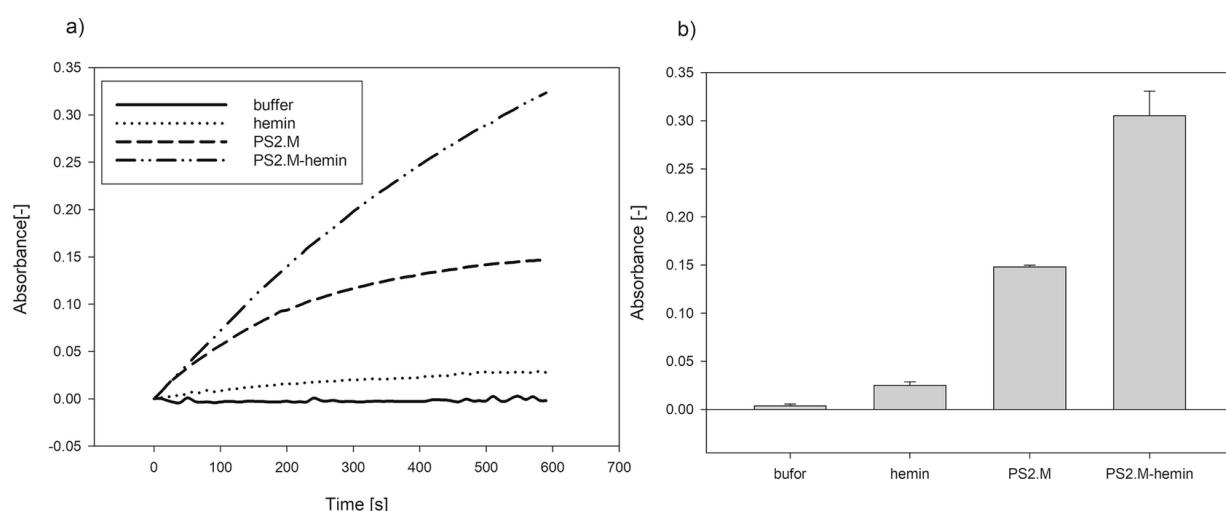
**Fig. 2** (a) HPLC chromatogram of PS2.M-hemin synthesis mixture,  $R_t = 9.1$  min corresponds to PS2.M-hemin and  $R_t = 14.5$  to unreacted hemin azide. (b) UV-Vis spectrum of reaction product with  $R_t = 9.1$  min

### 3.3 DNAzyme Activity Measurement

1. To the 0.5 ml reaction tube add 20  $\mu$ L Tris-HCl buffer, 20  $\mu$ L KCl, 2  $\mu$ L 100  $\mu$ M DNAzyme (hemin-DNA conjugate), and 126.3  $\mu$ L water. For DNA oligonucleotide reference add 20  $\mu$ L Tris-HCl buffer, 20  $\mu$ L KCl, 2  $\mu$ L DNA oligonucleotide, 2  $\mu$ L 100  $\mu$ M hemin, and 124.3  $\mu$ L water.
2. Mix the solutions. Preheat thermoblock to 90 °C. Incubate the tube for 5 min in thermoblock.
3. Transfer the tube to ice and incubate for 15 min. After that incubate the solution for 15 min at room temperature.

## 4.

Add 20 µL of ABTS. Mix the solution and transfer 140 µL to a clear microplate. Add 10 µL H<sub>2</sub>O<sub>2</sub> using dispenser. Measure the absorbance change at 414 nm for 10 min in a microplate reader (see Note 10). Figure 3 shows dependence of absorbance at 414 nm on time for hemin, PS2.M oligonucleotide and PS2.M-hemin conjugate obtained by click chemistry.



**Fig. 3** Comparison of peroxidase activity of DNA–hemin complex and DNA–hemin conjugate in reaction with ABTS. (a) Change of absorbance at 414 nm over time. (b) Absorbance of each sample at 414 nm measured after 10 min

## 4 Notes

1. In this method polyethylene glycol connector with three ethylene glycol repeats was used. The connectors with different length and chemistry are also commercially available.
2. Here we used oligonucleotide based on PS2.M sequence which was one of the first sequences of peroxidase-mimicking DNAzyme: 5'-GTG GGT AGG GCG GGT TGG-3'. Oligonucleotide without modifications was used for comparison of activity enhancement for PS2.M-hemin conjugate.

3.

We recommend storage of the hemin and hemin azide in a solid form and dissolving of required amount in DMSO only when needed. The stock solution is stable for 3 weeks at -30 °C.

4.

ABTS has to be freshly prepared before the measurement of DNAzyme activity. We also recommend to prepare fresh solution of H<sub>2</sub>O<sub>2</sub>.

5.

We use aluminum plates covered with silica gel 60 F<sub>254</sub> 0.2 mm thick (Merck). For TLC experiments we routinely cut out 10 × 5 cm plates and on the narrow edge we mark the 0.5 cm distance from the edge where the samples are applied. The plate is next placed in glass container with required solvents mixture (here chloroform–methanol 7:1 v/v). The level of the solvent must be below the line of samples application. Next, the glass container is closed and chromatographic separation is observed by movement of the solvent along the plate. The plate has to be removed from container before the solvent reach the opposite edge of the plate. Next, the plate is dried and the position of the samples is analyzed. In the case of hemin and its derivatives the intensive color allows for visual observation.

6.

Use 7–10 g of ice for 1 mL of DMF after reaction.

7.

The drying agent needs be added when clumping is no longer observed.

8.

Before adding silica gel on the column it should be mixed to prevent formation of bubbles. Next the air bubbles that formed in the gel should be removed by careful strokes. The silica gel must not be allowed to dry so a level of solvent should always be above gel. The gel–sample ratio and diameter of the column affects the speed of the separation. Greater amount of gel and smaller diameter provide better separation but extend the separation time. For faster separation external pressure can be applied.

For the separation of oligonucleotide-hemin conjugates we are using

9. preparative column Waters X-terra MS C18 ( $d_p = 2.5 \mu\text{M}$ ,  $4.6 \times 100 \text{ mm}$ ).
10. DNAzyme activity can be also measured using fluorogenic substrates like Amplex Red and MNBDH. For Amplex Red the procedure is as follows:

Add to the 0.5 mL reaction tube 11  $\mu\text{L}$  of Tris–HCl buffer, 11  $\mu\text{L}$  KCl, 5.5  $\mu\text{L}$  1  $\mu\text{M}$  DNA oligonucleotide and 60.4  $\mu\text{L}$  water. Incubate the tube for 5 min in thermoblock at 90 °C. Transfer the tube to ice and incubate for 15 min. After that incubate the solution for 15 min in room temperature. Add 1.1  $\mu\text{L}$  1 mM Amplex Red. Incubate the solution in dark for 30 min. Transfer 90  $\mu\text{L}$  the solution to black microplate. Add 10  $\mu\text{L}$   $\text{H}_2\text{O}_2$  using dispenser. Measure the fluorescence change at 595 nm with excitation at 560 nm for 10 min in microplate reader.

For the measurements with MNBDH the following procedure is recommended:

Add to the 0.5 mL reaction tube 11  $\mu\text{L}$  Tris–HCl buffer, 11  $\mu\text{L}$  KCl, 5.5  $\mu\text{L}$  1  $\mu\text{M}$  DNA oligonucleotide, and 60.4  $\mu\text{L}$  water. Incubate the tube for 5 min in thermoblock at 90 °C. Transfer the tube to ice and incubate for 15 min. After that incubate the solution for 15 min in room temperature. Add 1.1  $\mu\text{L}$  1 mM MNBDH. Incubate the solution in dark for 30 min. Transfer 90  $\mu\text{L}$  the solution to black microplate. Add 10  $\mu\text{L}$   $\text{H}_2\text{O}_2$  using dispenser. Measure the fluorescence change at 540 nm with excitation at 470 nm for 10 min in microplate reader. The measurement of DNAzyme activity can be also performed in cuvettes using spectrophotometer (in the case of measurement with ABTS) or spectrofluorometer (in the case of measurement with MNBDH or Amplex Red) but such measurements require approximately 10 times higher amount of reagents.

## Acknowledgments

I wish to acknowledge help and mentorship provided by Prof. Bernard Juskowiak from Adam Mickiewicz University, Poznan, Poland. This work

was supported by National Science Center (Poland) grant no. 2013/10/M/ST4/00490.

---

## References

1. Breaker RR (1997) DNA enzymes. *Nat Biotechnol* 15:427–431  
[[Crossref](#)]
2. Travascio P, Li YF, Sen D (1998) DNA-enhanced peroxidase activity of a DNA aptamer-hemin complex. *Chem Biol* 5:505–517  
[[Crossref](#)]
3. Silverman SK (2004) Deoxyribozyme: DNA catalysts for bioorganic chemistry. *Org Biomol Chem* 2:2701–2706  
[[Crossref](#)]
4. Burge S, Parkinson GH, Hazel P, Todd AK, Neidle S (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res* 34:5402–5415  
[[Crossref](#)]
5. Kosman J, Juskowiak B (2011) Peroxidase-mimicking DNAzymes for biosensing applications: a review. *Anal Chim Acta* 707:7–17  
[[Crossref](#)]
6. Zhou YY, Zhou X (2012) Studies of the activity of peroxidase-like DNAzyme by modifying 3'- or 5'-end of aptamers. *Chem Biodivers* 9:170–180  
[[Crossref](#)]
7. Li C, Zhu L, Fu H, Jenkins G, Wang C, Zou Y, Lu X, Yang CJ (2012) Backbone modification promotes peroxidase activity of G-quadruplex based DNAzyme. *Chem Commun* 48:8347–8349  
[[Crossref](#)]
8. Kosman J, Żukowski K, Juskowiak B (2019) Increasing the activity of DNAzyme based on telomeric sequence: 2'-OMe-RNA and LNA modifications. *Open Chem* 17:1157–1166  
[[Crossref](#)]
9. Thirstrup D, Baird GS (2010) Histochemical application of peroxidase DNAzyme with covalently attached hemin cofactor. *Anal Chem* 8:2498–2504  
[[Crossref](#)]
10. Nakayama S, Wang J, Sintim OH (2011) DNA-based peroxidation catalyst – what is

the exact role of topology on catalysis and is there a special binding site for catalysis?  
Chem Eur J 17:5691–5698  
[\[Crossref\]](#)

11. Gribas AV, Korolev SP, Zatsepin TS, Gottikh MB, Sakharov IY (2015) Structure-activity relationship study for design of highly active covalent peroxidase-mimicking DNAzyme. RSC Adv 5:51672–51677  
[\[Crossref\]](#)
12. Fruk L, Muller J, Niemeyer CM (2006) Kinetic analysis of semisynthetic peroxidase enzymes containing a covalent DNA-heme adducts as cofactor. Chem Eur J 12:7448–7457  
[\[Crossref\]](#)
13. Kosman J, Stanislawska A, Gluszynska A, Juskowiak B (2017) Conjugation of hemin to G-quadruplex forming oligonucleotide using click chemistry. Int J Biol Macromol 101:799–804  
[\[Crossref\]](#)
14. Fasman GD (ed) (1975) Handbook of biochemistry and molecular biology, Nucleic Acids, vol 1, 3rd edn. CRC Press, Boca Raton

## 19. Generation of RNA with 2', 3'-Cyclic Phosphates by Deoxyribozyme Cleavage in Frozen Solutions

Kristian K. Le Vay<sup>1</sup> and Hannes Mutschler<sup>1, 2</sup>✉

- (1) Biomimetic Systems, Max Planck Institute of Biochemistry,  
Martinsried, Germany  
(2) Faculty of Chemistry and Chemical Biology, TU Dortmund University,  
Dortmund, Germany

✉ **Hannes Mutschler**

Email: [hannes.mutschler@tu-dortmund.de](mailto:hannes.mutschler@tu-dortmund.de)

### Abstract

The generation of terminal 2', 3'-cyclic phosphates on RNA oligomers is an important process in the study of tRNA splicing and repair, ribozyme catalysis, and RNA circularization. Here, we describe a simple method for producing 2', 3'-cyclic phosphate functionalized RNA by the deoxyribozyme-catalyzed cleavage of a short 3'-RNA overhang in frozen solution. This method avoids the nonspecific modification and degradation of RNA and attached functional groups (e.g., fluorophores) inherent in other methods, and the use of frozen conditions enables cleavage at very low divalent metal ion concentrations, limiting RNA hydrolysis.

**Key words** Site-specific RNA cleavage – Cyclic phosphates – 8–17 deoxyribozyme – Ice – Ribozyme substrate synthesis

---

## 1 Introduction

RNA 2',3'-cyclic phosphate ends play important roles in RNA metabolism as substrates for tRNA splicing and general RNA repair [1–3]. Furthermore, the chemical energy stored in the strained cyclic phosphate enables direct nonenzymatic and nucleic acid enzyme–catalyzed ligation of RNA [4–8]. In these reactions, the 5'-OH and the terminal 2',3'-cyclic phosphate of two oligomers are brought into proximity by a complementary substrate binding strand, then nucleophilic attack by the 5'-OH results in the opening of the strained cyclic phosphate and formation of a new 2'- or 3'-5' linkage. This reaction is postulated to have played a role during an early nucleic acid-based biology, as it provides a plausible route toward the formation of long, complex, and functional RNA via the ligation of short oligonucleotide fragments [9, 10].

A commonly used methodology for producing 2', 3'- cyclic phosphate is treatment of 2'- or 3'-phosphate modified RNA with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) [10, 11]. While effective, EDC modification is not site specific to the terminal phosphate, and in particular may affect any exposed carboxyl groups, which can then crosslink with primary amines. A notable consequence is that common fluorophore labels used in nucleic acid research, such as 6-carboxyfluorescein, react with EDC resulting in a significant reduction of fluorescence in the 2', 3'-cyclic phosphate functionalized product. EDC has also been shown to modify unpaired T and G bases [12, 13].

Here, we present a simple method for the formation of 2', 3'-cyclic phosphates on RNA oligomers using a derivative of the 8–17 deoxyribozyme, which cleaves a short 3'-tail sequence from the target RNA oligomer [14]. The reaction is performed in frozen solution, which allows the cleavage to proceed at very low divalent metal ion concentrations, thereby limiting degradation of the target RNA oligomer by hydrolysis. In this regard, we have chosen  $\text{Ca}^{2+}$  as the divalent metal ion in this reaction, which leads to less cyclic phosphate hydrolysis than other metal ions (e.g.,  $\text{Mn}^{2+}$ ), while permitting deoxyribozyme activity in frozen solution [15]. Freezing has previously also been shown to accelerate the cleavage of RNA by the EtNa deoxyribozyme [16].

---

## 2 Materials

### 2.1 Nucleic Acids

1. RNA and DNA oligomers are purchased with standard desalting and used without further purification. The nucleic acid oligomers used in this example are:
  - (a) E1111 8–17 deoxyribozyme: 5'-CAG ATG TCA GCG ACA CGA AGT CAT TTT G-3'.
  - (b) RNA substrate: 5'-FAM-rCrArA rArArU rGrArC rArGrU rCrUrG-3'.
2. Dissolve lyophilized nucleic acids in RNase-free water to a final concentration of 100 µM and store at –80 °C.

### 2.2 Tris–Borate–Urea (TBU) Acrylamide Stock and Tris–Borate–Ethylenediaminetetraacetic Acid (EDTA) (TBE) Running Buffer

1. Tris(hydroxymethyl)aminomethan (Tris)-borate-ethylenediaminetetraacetic acid (EDTA) (TBE) running buffer (10× stock): 0.89 M Tris base, 0.89 M boric acid, 20 mM EDTA (pH 8). Weigh 108 g Tris base and 55 g boric acid. Add 40 mL 500 mM EDTA solution (pH 8) and 500 mL water, then stir until dissolved. Make up to 1 L in a graduated cylinder, vacuum filter using a cellulose acetate membrane and store at room temperature.
2. 8 M urea–20% acrylamide stock: 20% acrylamide–bis solution (19:1), 8 M urea, 1× TBE buffer. Weigh 240.24 g urea into a 1 L glass beaker. Add 250 mL 40% acrylamide–bis (19:1) and 50 mL of 10× TBE buffer prepared in **step 1**. Heat with stirring in a fume hood to dissolve urea. Vacuum filter using a cellulose acetate membrane (*see Note 1*), then store at room temperature away from direct sunlight.
3. 10% (w/v) ammonium persulfate (APS) in water, stored at –20 °C.

4. *N,N,N,N'-tetramethyl-ethylenediamine* (TEMED), stored at 4 °C.

## 2.3 Gel Loading Buffers

1. For 10 mL loading buffer, dissolve 2–5 mg bromophenol blue in 9.99 mL formamide and 10 µL EDTA (0.5 M, pH 8). Store at room temperature.

## 2.4 Additional Reagents and Materials

1. 30 mM CaCl<sub>2</sub> stock solution.
2. 5 M NaCl stock solution.
3. 3 M Sodium acetate solution pH 5.
4. 500 mM Tris–HCl buffer stock solution pH 7.4.
5. 5 mg/mL Glycogen solution.
6. Ethanol.
7. Isopropanol.
8. 0.22 µm Cellulose Acetate SpinX columns.

## 2.5 Equipment

1. Slab gel casting kits: 8 × 8 cm mini-gel or 20 × 20 cm maxi-gel, 1 mm thick, and 20 × 20 cm maxi-gel, 2 mm thick.
2. PCR cycler.
3. Programmable cooling water bath (−30 to 37 °C). Alternatively: solid

state cold block ( $-9\text{ }^{\circ}\text{C}$ ), heat block ( $37\text{ }^{\circ}\text{C}$ ) and freezer ( $-80\text{ }^{\circ}\text{C}$ ).

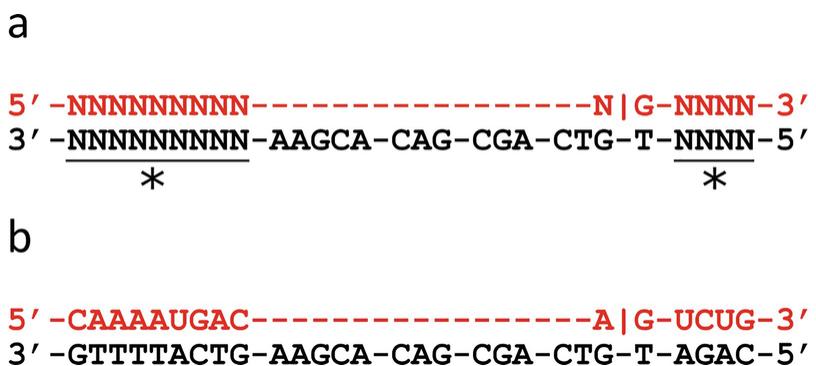
4. Tube rotator for 2 mL reaction tubes in cold room ( $4\text{ }^{\circ}\text{C}$ ).
  5. 254 nm UV-lamp, autofluorescent silica-coated glass plate and protective equipment.
  6. Scalpel.
  7. Cooling centrifuge (ideally  $20,900 \times g$ , cooling to  $4\text{ }^{\circ}\text{C}$  or lower).
  8. Fluorescent gel imaging system.
- 

### 3 Methods

#### 3.1 Nucleic Acid Design

1. In this example, we use the E1111 variant of the 8–17 deoxyribozyme to cleave an RNA oligomer at an rA|rG junction, leaving a 2',3'-cyclic phosphate at the terminal 3'-A nucleotide [14]. This junction was chosen as cleavage proceeds quickly to completion, and because a 2',3'-cyclic phosphate on the terminal 3'-A nucleotide is commonly used in our lab for hairpin ribozyme ligation reactions. The length of substrate can be freely varied, provided that there is sufficient electrophoretic separation between the size of substrate, cleaved products, and deoxyribozyme to enable gel purification of the product with the 2',3'-cyclic phosphate (>5–10 nucleotides (nt)).
2. The E1111 deoxyribozyme is sufficiently versatile to readily cleave rC|rG, rU|rG, and rG|rG junctions with similar efficiency and therefore produce a 2',3'- cyclic phosphate on any terminal nucleotide. Although we have not tested them in our method, other 8–17 variants are likely to be suitable for cleavage of different rN|rN junctions (see Note 2) [14].
3. To adapt the deoxyribozyme to be able to cleave a specific RNA

sequence, alter the two binding regions that flank the deoxyribozyme catalytic core to be complementary to the regions on either side of the cleavage site. An overview of the system used here is shown in Fig. 1.

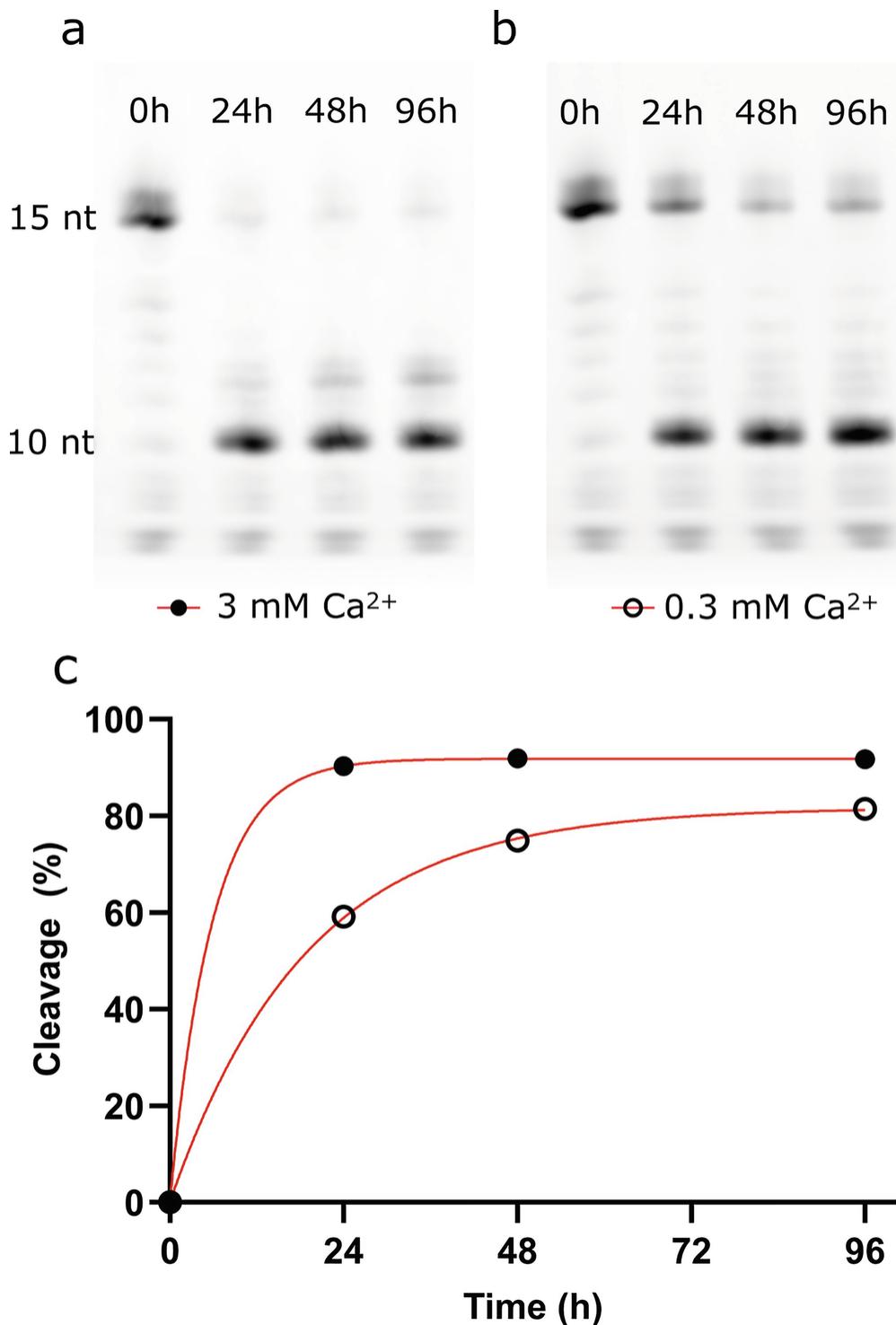


**Fig. 1** A schematic showing (a) a generalized deoxyribozyme-RNA substrate complex structure and (b) the specific sequences used here. The RNA substrate and deoxyribozyme are shown in red and black respectively. The variable binding regions, which may be freely varied depending on RNA substrate sequence, are marked with an asterisk

### 3.2 Deoxyribozyme-Catalyzed RNA Cleavage and Cyclic Phosphate Formation

1. The final reaction composition is: 20  $\mu$ M 8–17 deoxyribozyme, 10  $\mu$ M RNA substrate, 75 mM NaCl, 0.3 mM CaCl<sub>2</sub>, 50 mM Tris–HCl (pH 7.4).
2. First, anneal the E1111 deoxyribozyme and RNA substrate by adding both to a polymerase chain reaction (PCR) tube and incubating at 92 °C for 5 min, and then cool gradually to 23 °C over 10 min in a thermal cycler.
3. Once cooled, add water, NaCl, CaCl<sub>2</sub>, and Tris–HCl buffer (pH 7.4) to obtain the final reaction composition (see **step 1**).
4. Place the reaction mixture to a –80 °C freezer for 30 min, then transfer to a cooling bath or block set at –9 °C (see **Note 3**).
5. Incubate for 1–4 days (see Fig. 2 and **Note 4**).
6. Quench the reaction by adding twice the volume of gel loading buffer.

To assess reaction progress, run aliquots of the reaction mixture on a 20% TBU acrylamide gel (*see* Fig. 2 and Note 4).



**Fig. 2** Analytical urea-PAGE gel (20% acrylamide) showing the 8–17 deoxyribozyme-catalyzed cleavage of the FAM-tagged RNA substrate over time with (a) 3 mM and (b) 0.3 mM  $\text{Ca}^{2+}$ . (c) Plot of RNA substrate cleavage yield over time at 3 mM (filled circles)

and 0.3 mM (empty circles)  $\text{Ca}^{2+}$ . Yields were determined from the integrated intensities of substrate and product bands

### 3.3 Quantification of Cleavage by Urea–Polyacrylamide Gel Electrophoresis (PAGE)

1. To assess cleavage, separate small aliquots of the reaction mixture on a 1 mm thick  $20 \times 20$  cm 20% TBU acrylamide gel (*see Note 5*).
2. For each  $20 \times 20$  cm maxi gel, mix 29.67 mL of TBU gel stock with 30  $\mu\text{L}$  TEMED and 300  $\mu\text{L}$  10% APS (w/v) to initiate polymerization.
3. Quickly pour the gel mixture into a preassembled cast and insert the comb, then leave to polymerize for a minimum of 2 h (ideally overnight) at room temperature.
4. Quench a 5  $\mu\text{L}$  reaction mixture aliquot with 20  $\mu\text{L}$  of gel loading buffer and denature by heating to 80 °C for 5 min.
5. Preheat the gel by running at 25 W for 10 min, then use a syringe and needle to flush the wells with running buffer to remove urea at least twice.
6. Load 5  $\mu\text{L}$  of the quenched sample, then run at low power (15 W) until the blue loading dye has completely entered the gel, then increase power to 25 W and run until the blue loading dye band is close to the end of the gel.
7. Image the gel using a fluorescent gel imager and determine the progress of the reaction by the ratio of the integrated intensities of the cleaved and uncleaved bands.

### 3.4 Purification of Cleaved Substrate by Urea-PAGE

1. To purify the cleaved RNA fragment, separate the reaction mixture on a 2 mm thick  $20 \times 20$  cm 20% TBU acrylamide gel (*see Note 5*).
2. For each gel, mix 59.34 mL of TBU gel stock, 60  $\mu\text{L}$  TEMED and

- 600 µL 10% APS (w/v) to initiate polymerization.
3. Quickly pour the gel mixture into a preassembled cast and insert a wide toothed comb to produce large wells (10 mm × 15 mm × 2 mm), then leave to polymerize for a minimum of 2 h (ideally overnight) at room temperature.
4. Quench 50–200 µL of reaction mixture with an equal volume of gel loading buffer and denature by heating to 80 °C for 5 min.
5. Preheat the gel by running at 25 W for 10 min, then use a syringe and needle to flush the wells with running buffer at least twice to remove urea.
6. Divide the quenched reaction mixture between the wells of the gel. Generally, the sample should be loaded into as few wells as possible to maximize the amount of RNA in each band, approximately 100–200 µL per lane.
7. Run at low power (15 W) until the blue loading dye has completely entered the gel, then increase power to 25 W and run until the blue loading dye band is close to the end of the gel.
8. Once the run is finished, remove the gel from the casing and wrap in clingfilm or saran wrap.
9. Determine the position of the cleaved RNA fragment using the fluorescent tag (e.g., either by illumination of the FAM-tag with blue light and a yellow filter, or simply by eye if the band is concentrated enough).
10. If the substrate has no fluorescent tag, identify the band using UV shadowing. Place the wrapped gel on an autofluorescent background (*see Note 6*) and illuminate with a 254 nm lamp. Dark regions caused by absorption of UV light appear dark. Make sure suitable personal protection equipment is worn to prevent UV light exposure and limit the UV exposure of the sample to no more than 30 s to prevent photodamage.
11. Mark the outline of the band visualized with either method using a

permanent marker, then excise the band using a scalpel. Take care to avoid excess gel material as this will reduce overall yields.

12. Place each gel slice in a 2 mL reaction tube and crush (*see Note 7*).
13. Add 1–2 volumes of sodium acetate (0.3 M, pH 5) to the gel slice (e.g., if the mass of the gel slice is 300 mg, add 300–600 µL of elution buffer) and rotate overnight at 4 °C.
14. The following day, remove the gel debris by loading onto 0.22 µm Cellulose Acetate SpinX columns and centrifuging at 16,000 × *g* for 1 min. Discard the filter containing the gel debris.
15. Add the following to each tube: 10 µL sodium chloride solution (5 M) 1 µL glycogen solution (5 mg/mL), one volume of isopropanol and two volumes of ethanol.
16. Mix well by inverting the tubes, then freeze at –80 °C for at least 1 h.
17. Centrifuge at 20,900 × *g* at 4 °C or colder for 45 min to precipitate the RNA, which will form a small but visible pellet.
18. Using a pipette, carefully remove the supernatant without agitating the pellet.
19. Add 1 mL ice-cold 80% ethanol, incubate at 4 °C for 10 min, then centrifuge at 20900 × *g* at 4 °C or colder for 10 min. Use a pipette to carefully remove the supernatant again.
20. Repeat **step 19**.
21. Place the tubes in a heat block set to 37 °C, lids removed, to evaporate residual ethanol, then resuspend the purified RNA in small volume of RNase-free water (10–50 µL).
22. Determine the concentration by measuring the absorbance at 260 nm and using a sequence specific extinction coefficient. Store RNA at –80 °C until use.

---

## 4 Notes

1. The solvation of urea is endothermic and proceeds very slowly without gentle heating. In addition, the final mixture is quite viscous and filters very slowly when cold, so this step should be carried out when it is still warm.
2. In particular, the E2112 variant of the 8–17 deoxyribozyme is capable of cleaving A|rA, C|rA, G|rA, and U|rA with high efficiency.
3. Freezing initially at –80 °C ensures that the mixture is completely frozen before eutectic phase formation and reaction at –9 °C. Omitting this step may prevent freezing, which is essential for DNAzyme activity with low divalent metal ion concentrations. For cleavage of more complex substrates or less active deoxyribozymes, temperature cycling can be used to promote activity. In an example cycle, the sample is first heated to 37 °C for 30 min, which allows for mild reannealing and rescue of inactive complexes. The sample is then cooled to –30 °C for 2 h, ensuring freezing, then incubated at –9 °C for 9 h 30 min before repeating the cycle. In our lab this is achieved using Grant R3 and R4 refrigerated water baths filled with 50% ethylene glycol, which are controlled by TXF150 thermostatic controllers capable of programmed cycling. However, this process can be performed manually with less specialized equipment, for example by manually transferring the sample between a heat block or PCR cycler for incubation at 37 °C, a –80 °C freezer to ensure freezing, and a –9 °C cooling block or bath for eutectic phase formation and reaction.
4. The reaction may be removed from the bath to take aliquots to follow reaction progress, but the reaction mixture must be refrozen at –80 °C before returning to –9 °C. It is generally easier to aliquot some of the reaction mixture into separate tubes at the start of the reaction that can be removed and quenched later to assess progress.
5. Either mini- (8 × 8 cm) or maxi- (20 × 20 cm) gels between 15% and 20% acrylamide are suitable for assessing reaction progress, with larger gels offering improved resolution. Longer substrates will require lower acrylamide concentrations. For the following purification step, a 2 mm

thick maxi-gel is required to achieve sufficient separation between the bands for effective purification.

6.

Silica-coated glass thin layer chromatography plates are a cheap and readily available option. High-quality printer paper can suffice if no brighter autofluorescent backgrounds are available.

7.

A small spatula or pipette tip can be used to crush the gel. In particular, the plunger from a 1 mL syringe is an effective tool. The protocol now proceeds with the classical “freeze and squeeze” method of nucleic extraction, but more modern alternatives such as electroelution may be used if available.

---

## References

1. Tanaka N, Chakravarty AK, Maughan B, Shuman S (2011) Novel mechanism of RNA repair by RtcB via sequential 2',3'- cyclic phosphodiesterase and 3'-phosphate/5'-hydroxyl ligation reactions. *J Biol Chem* 286(50):43134–43143. <https://doi.org/10.1074/jbc.M111.302133>  
[Crossref][PubMed][PubMedCentral]
2. Tanaka N, Shuman S (2011) RtcB is the RNA ligase component of an Escherichia coli RNA repair operon. *J Biol Chem* 286(10):7727–7731. <https://doi.org/10.1074/jbc.C111.219022>  
[Crossref][PubMed][PubMedCentral]
3. Tanaka N, Meineke B, Shuman S (2011) RtcB, a novel RNA ligase, can catalyze tRNA splicing and HAC1 mRNA splicing in vivo. *J Biol Chem* 286(35):30253–30257. <https://doi.org/10.1074/jbc.C111.274597>  
[Crossref][PubMed][PubMedCentral]
4. Lutay AV, Chernolovskaya EL, Zenkova MA, Vlasov VV (2005) Nonenzymatic template-dependent ligation of 2',3'-cyclic phosphate-containing oligonucleotides catalyzed by metal ions. *Dokl Biochem Biophys* 401:163–166. <https://doi.org/10.1007/s10628-005-0060-4>  
[Crossref][PubMed]
5. Lutay AV, Chernolovskaya EL, Zenkova MA, Vlassov VV (2006) The nonenzymatic template-directed ligation of oligonucleotides. *Biogeosciences* 3:243–249  
[Crossref]

6. Dolinnaya NG, Sokolova NI, Ashirbekova DT, Shabarova ZA (1991) The use of BrCN for assembling modified DNA duplexes and DNA-RNA hybrids; comparison with water-soluble carbodiimide. *Nucleic Acids Res* 19(11):3067–3072. <https://doi.org/10.1093/nar/19.11.3067>  
[Crossref][PubMed][PubMedCentral]
7. McLeod AC, Lilley DMJ (2004) Efficient, pH-dependent RNA ligation by the VS ribozyme in trans. *Biochemistry* 43(4):1118–1125. <https://doi.org/10.1021/bi035790e>  
[Crossref][PubMed]
8. Hoadley KA, Purtha WE, Wolf AC et al (2005) Zn<sup>2+</sup>-dependent deoxyribozymes that form natural and unnatural RNA linkages. *Biochemistry* 44(25):9217–9231. <https://doi.org/10.1021/bi050146g>  
[Crossref][PubMed]
9. Mutschler H, Wochner A, Holliger P (2015) Freeze-thaw cycles as drivers of complex ribozyme assembly. *Nat Chem* 7:502–508. <https://doi.org/10.1038/nchem.2251>  
[Crossref][PubMed][PubMedCentral]
10. Mutschler H, Holliger P (2014) Non-canonical 3'-5' extension of RNA with prebiotically plausible ribonucleoside 2',3'-cyclic phosphates. *J Am Chem Soc* 136:5193–5196. <https://doi.org/10.1021/ja4127714>  
[Crossref][PubMed][PubMedCentral]
11. Hertel KJ, Uhlenbeck OC, Herschlag D (1994) A kinetic and thermodynamic framework for the hammerhead ribozyme reaction. *Biochemistry* 33(11):3374–3385. <https://doi.org/10.1021/bi00177a031>  
[Crossref][PubMed]
12. Obianyor C, Newnam G, Clifton BE et al (2020) Towards efficient nonenzymatic DNA ligation: comparing key parameters for maximizing ligation rates and yields with carbodiimide activation. *Chembiochem* 21(23):3359–3370. <https://doi.org/10.1002/cbic.202000335>  
[Crossref][PubMed]
13. Shabarova ZA, Dolinnaya NG, Drutsa VL et al (1981) DNA-like duplexes with repetitions. III. Efficient template-guided chemical polymerization of d(TGGCCAAGCTp). *Nucleic Acids Res* 9(21):5747–5761. <https://doi.org/10.1093/nar/9.21.5747>  
[Crossref][PubMed][PubMedCentral]
14. Cruz RPG, Withers JB, Li Y (2004) Dinucleotide junction cleavage versatility of 8-17 deoxyribozyme. *Chem Biol* 11:57–67. [https://doi.org/10.1016/S1074-5521\(03\)00284-9](https://doi.org/10.1016/S1074-5521(03)00284-9)

[Crossref][PubMed]

15. Breslow R, Huang DL (1991) Effects of metal ions, including Mg<sup>2+</sup> and lanthanides, on the cleavage of ribonucleotides and RNA model compounds. Proc Natl Acad Sci U S A 88:4080–4083. <https://doi.org/10.1073/pnas.88.10.4080>  
[Crossref][PubMed][PubMedCentral]
16. Yu T, Zhou W, Liu J (2018) An RNA-cleaving catalytic DNA accelerated by freezing. Chembiochem 19:1012–1017. <https://doi.org/10.1002/cbic.201800049>  
[Crossref][PubMed]

# **Index**

## **A**

African Swine Fever Virus Polymerase X (*AsfvPolX*)  
Agarose gel electrophoresis  
Amber suite of programs  
Aptamer  
Azide-alkyne cycloaddition

## **B**

Binding constant  
Binding isotherms  
Binding sites  
Biosensors

## **C**

Capture-SELEX  
Catalysis  
Circular dichroism (CD) spectroscopy  
Cleavage assay  
Click chemistry  
Conformational dynamics  
Conformational transitions  
Crystallization  
Crystallography  
Cyclic phosphates  
Cycloaddition

## **D**

Degenerate FRET states  
Denaturation temperature  
Desalting  
Dextran  
DINAMelt  
Double electron-electron resonance

## **E**

Ectoine  
8-17 DNAzyme  
Electron paramagnetic resonance (EPR)  
Electrophoretic Mobility Shift Assay (EMSA)  
Enzymatic ligation of RNA  
Expanded genetic alphabet

## F

Fluorescent RNA labeling  
Folding of nucleic acid  
Fürster resonance energy transfer (FRET)

## G

Gold nanoparticles  
G-quadruplex  
Group II intron

## H

Hemin  
High performance liquid chromatography  
Homonuclear NMR

## I

In vitro selection  
In vitro transcription

## K

Kinetics  
measurements

## L

Ligation in ice

## M

Metalation  
Metal ion cofactors  
Metal ions  
 $Mn^{2+}$  titration

Molecular crowding

Molecular dynamics (MD) simulation

Multifunctional Analysis Software to Handle single-molecule FRET data (MASH-FRET)

## N

Nuclear magnetic resonance (NMR)

## O

Observed rate constant ( $K_{\text{obs}}$ )

## P

Paramagnetic relaxation enhancement (PRE)

Pearson hardness

Peroxidase activity

Phosphorothioate (PS)

Poland program

Polyacrylamide gel electrophoresis (PAGE)

Polyethylene glycol (PEG)

Polymerase chain reaction (PCR)

Porphyrin

Post-transcriptional labeling

Pro- $R_p$  and pro- $S_p$  oxygen

Pulsed dipolar spectroscopy (PDS)

Pulsed electron-electron double resonance

Purification of DNA

## R

Rate constant

Ribozymes

RNA

cleavage

dual-end labeling

labeling

purification

spin labeling

RNAfold

RNAplfold

**S**

Single-molecule FRET (smFRET)  
Site-specific labeling  
Site-specific RNA cleavage  
Size exclusion chromatography (SEC)  
Soaking  
Spin-labeled RNA  
Spin labeling  
ssDNA library recovery  
ssDNA production  
State transitions  
Structural dynamics  
Structure determination  
Systematic evolution of ligands by exponential

**T**

10-23 DNAzyme  
T7 transcription

**U**

U1 small nuclear ribonucleoprotein A  
Unnatural base pairs

**X**

X-ray crystallography