Chemical Methods for Stabilizing RNA in Crystallographic Research

Mgr Martyna Mateja-Pluta

Supervisor: dr hab. Agnieszka Kiliszek, Prof. IBCH

Department of RNA Structural Research

Poznań, 2025

I would like to express my heartfelt gratitude to all the people
who contributed to the completion of this dissertation:

My supervisor, ***Dr. hab. Agnieszka Kiliszek, prof. IBCH***
for her invaluable help, scientific guidance, and dedicated mentorship.
I am deeply grateful for the years of collaboration,
countless hours of discussions – both scientific and beyond,
and inspiring brainstorming sessions
that shaped both this work and my development as a researcher.

To my colleagues and friends from the
Department of RNA Structural Research *and beyond,*
Thank you for the wonderful atmosphere,
the many conversations, and all the help and support you have given me.
Especially to ***Magda, Kasia, Tosia, Alena, and Jagoda***
for being an incredible support group and even better company
throughout my time at the Institute, both at work and outside it.

To ***Mrs. Małgosia Dąbkiewicz****, for being a kind soul*
who supported me every step of the way.

To my ***grandparents****, who supported me from the very beginning.*
To my ***grandpa****, for passing on to me—surely in my genes—his passion for science.*
To my ***granny****, for her countless words of encouragement and unwavering belief in me.*
Thank you for shaping the person I have become and for standing by me at every step of this journey.

To my ***mom****, who never doubted me and always knew*
that I would achieve everything I set my mind to.

To my ***husband Mateusz****, for his immense heart, extraordinary support,*
and endless patience with my complaints and meltdowns.
My love, I would not have done this without you.

Table of contents

Funding

The research conducted in this dissertation was supported by grant:

NCN SONATA BIS 2017/26/E/NZ1/00950 - *Development of methodology aiming at stabilisation of RNA in a hairpin form for crystallographic studies*

2022/45/B/NZ7/03543 - *Crystallographic analysis of RNA-ligand complexes. Towards rational design oflead compounds for development of therapies against neurodegenerative disorders*

List of papers included in the doctoral dissertation

1. Marcin Ryczek, Martyna Pluta, Leszek Błaszczyk*, Agnieszka Kiliszek*

Overview of Methods for Large-Scale RNA Synthesis

*Applied Sciences **2022**, 12, 1543.*

IF_{2022} = 2.7, $MNiSW_{2022}$ = 40

2. Martyna Mateja-Pluta, Leszek Błaszczyk, Magdalena Bejger, Kazuhiko Nakatani, Agnieszka Kiliszek*

Naphthyridine carbamate dimer ligand induces formation of Z-RNA-like fold of disease-related RNA and exhibits a molecular glue characteristics in crystal lattice formation

Nucleic Acids Research, **2025**, 53, gkaf924

IF_{2024} = 13.1, $MNiSW_{2025}$ = 200

The studies described in this PhD dissertation have been included in the following article, which is currently under review:

Design and Synthesis of Aromatic Linkers for Inter-strand Cross-linking in RNA with Applications in Structural Biology

Martyna Mateja-Pluta, Agnieszka Kiliszek*

List of papers not included in the doctoral dissertation

1. Martyna Mateja-Pluta, Agnieszka Kiliszek

Amidofosforynowy związek aromatyczny, sposób jego wytwarzania, jego zastosowanie oraz związek pośredni do otrzymywania amidofosforynowego związku aromatycznego.

Patent Application no. 13P53836PL00, **2025**

2. Agnieszka Kiliszek*, Martyna Pluta, Magdalena Bejger, Wojciech Rypniewski*

Structure and thermodynamics of a UGG motif interacting with Ba2+ and other metal ions: accommodating changes in the RNA structure and the presence of a G(syn)–G(syn) pair

RNA, **2022**, 29:44–54

IF_{2022} = 4.5, MNiSW = 140

3. Leszek Błaszczyk†, Marcin Ryczek†, Biomolendu Das†, Martyna Mateja-Pluta, Magdalena Bejger, Joanna Śliwiak, Kazuhiko Nakatani, Agnieszka Kiliszek*

Antisense RNA C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms a triplex-like structure and binds small synthetic ligand

Nucleic Acids Research, **2024**, 52, 6707–6717

IF_{2024} = 13.1, MNiSW = 200

*equal authors of the publication

Papers currently under review:

4. Martyna Mateja-Pluta, Joanna Strzelec, Marcin K. Chmielewski*, Agnieszka Kiliszek*

Structural Characterization of Cu(II), Pd(II), and Rh(II) Complexes with the Aminoalcohol Ligand N-(2-pyridyl)aminoethanol: Towards the Design of Ionic Traps Incorporating Thermolabile Units

5. Joanna Strzelec, Bartłomiej Szarłan, Martyna Mateja-Pluta, Marcin K. Chmielewski

2-Aminopyridyl Complexes with d-Block Metal Ions: A Framework for Thermo-Activated Catalysts

Conference presentations and posters:

1. The 25th Annual Meeting of the RNA Society (2020)

Martyna Pluta, Agnieszka Kiliszek, Ryszard Kierzek

Poster: *Pseudouridine as a modification in RNA duplexes*

2. The 26th Annual Meeting of the RNA Society (2021)

Martyna Pluta, Aleksandra Januchowska, Agnieszka Kiliszek

Poster: *Synthesis of modified units stabilizing RNA molecules*

3. 23rd Heart of Europe Bio-Crystallography Meeting (HEC23) (2021)

Martyna Pluta

Presentation: *RNA stabilizing modification used in crystallographic research*

4. Doctoral Symposium On Chemistry In Lodz (2023)

Martyna Mateja-Pluta, Agnieszka Kiliszek

Poster: *RNA stabilizing modification used in crystallographic research*

5. XVI Copernican Seminar of Doctoral Students (2023)

Martyna Mateja-Pluta

Presentation: *Chemiczne metody stabilizacji struktur spinkowych RNA do badań strukturalnych*

6. 26th Congress and General Assembly of the Union of Crystallography (IUCr 2023) Melbourne (2023)

Martyna Mateja-Pluta

Presentation: *RNA stabilizing modification used in crystallographic research*

7. 25th Heart of Europe Bio-Crystallography Meeting (HEC25) (2023)

Martyna Mateja-Pluta

Presentation: *Stabilization of RNA structures using small molecules in crystallographic research*

8. The 29th Annual Meeting of the RNA Society (2024)

Martyna Mateja-Pluta, Agnieszka Kiliszek

Poster: *RNA Stabilizing modification used in crystallographic research*

9. 65th Polish Crystallographic Meeting (2024)

Martyna Mateja-Pluta

Presentation: *Stabilization of RNA structures using small molecules in crystallographic research*

10. 2nd Polish RNA Biology Meeting (2025)

Martyna Mateja-Pluta, Leszek Błaszczyk, Magdalena Bejger, Kazuhiko Nakatani, Agnieszka Kiliszek

Poster: *Naphthyridine carbamate dimer ligand induces formation of Z-RNA-like fold of disease-related RNA and serves as a molecular glue for crystal lattice formation*

11. 35th European Crystallographic Meeting (ECM35) (2025)

Martyna Mateja-Pluta, Leszek Błaszczyk, Magdalena Bejger, Kazuhiko Nakatani, Agnieszka Kiliszek

Presentation: *Naphthyridine carbamate dimer ligand induces formation of Z-RNA-like fold of disease-related RNA and serves as a molecular glue for crystal lattice formation*

List of Figures and Tables

List of abbreviations

A – adenosine

ATP – 5’-adenosine triphosphate

BNA – bridged nucleic acid

C – cytidine

C12 - 12-carbon spacer arm

CD – circular dichroism

$CDCl_3$ – chloroform-d

CEP – cyanoethylphosphoramidite

CPG – controlled pore glass

DMSO – dimethyl sulfoxide

DMT - dimethoxytrityl

DNA - deoxyribonucleic acid

DSC – differential scanning calorimetry

EDTA - ethylenediaminetetraacetic acid

ESI-MS - electrospray ionization mass spectrometry

G - guanosine

ICL – interstrand cross-linking

LNA – locked nucleic acid

MBL – mismatch binding ligand

miRNA – microRNA

mRNA - messenger RNA

NCD – naphthyridine carbamate dimer

NMR – nuclear magnetic resonance

nSBs – nuclear stress bodies

NTP – 5’-nucleoside triphosphate

PCR – polymerase chain reaction

PDB – protein data bank

PEG – polyethylene glycol

RBP – RNA binding protein

RMSD – root-mean-square deviation

RNA – ribonucleic acid

siRNA - small interfering RNA

snRNA - small nuclear RNA

SRL – sarcin-ricin loop

TAR – trans-activation response

TBDMS – tert-butyldimethylsilyl

TCEP – tris(2-carboxyethyl)phosphine

TEA·3HF - triethylamine trihydrofluoride

TLC - thin-layer chromatography

TMS - tetramethylsilane

TOM - triisopropylsilyloxymethyl

TRIS - tris(hydroxymethyl)aminomethane

tRNA - transfer RNA

U - uridine

UV - ultraviolet

1. Abstract

The structural characterization of RNA molecules remains challenging due to the high conformational flexibility of nucleotide residues, their propensity to adopt multiple secondary and tertiary structures, and their susceptibility to structural rearrangements induced, for example, by experimental conditions. Although X-ray crystallography is an effective method for determining high-resolution RNA structures, crystallizing RNA molecules is often difficult and becomes the limiting step in the structural analysis process. For instance, RNA hairpins tend to unfold and form alternative duplex structures at high RNA and salt concentrations typically used for crystallization. Consequently, many biologically important native RNA motifs cannot be captured in a homogeneous crystalline state without additional stabilization strategies.

The aim of this dissertation was to develop new strategies for stabilizing RNA structures. Two approaches were applied: covalent stabilization using non-nucleosidic linkers and stabilization using a small-molecule ligand.

In the first strategy, a series of aromatic linkers were designed, synthesized, and incorporated into the RNA chain to covalently connect the 3' and 5' ends while preserving the native RNA geometry. In the second approach, the ability of a naphthyridine carbamate dimer (NCD) to stabilize RNA was investigated. It was demonstrated that the binding of a small-molecule ligand can stabilize RNA structures through mechanisms fundamentally different from covalent methods. NCD acts as a clamp linking nucleotide residues localized either within the same RNA molecule or between two adjacent RNA molecules, serving as a "molecular glue".

The final part of this thesis focuses on the biophysical and biochemical evaluation of the modified RNA constructs, including thermal denaturation studies monitored by UV spectroscopy, circular dichroism measurements, and in vitro cleavage assays using Dicer nuclease. These methods were used to assess how different stabilization strategies influence the RNA structural integrity, folding behavior, and biological compatibility.

In summary, this dissertation presents two strategies for stabilizing RNA molecules for structural studies. The results obtained expand our understanding of how both covalent non-nucleosidic linkers and small-molecule ligands can be used to overcome the challenges associated with RNA crystallization.

2. Streszczenie

Charakterystyka strukturalna cząsteczek RNA pozostaje wyzwaniem ze względu na wysoką elastyczność konformacyjną reszt nukleotydowych skłonność do przyjmowania wielu struktur drugorzędowych i trzeciorzędowych oraz podatność na rearanżacje strukturalne indukowane np.: warunkami eksperymentalnymi. Chociaż krystalografia rentgenowska jest skuteczną metodą wyznaczania struktury RNA o wysokiej rozdzielczości, krystalizacja cząsteczek RNA jest często trudna i limitująca cały proces analizy strukturalnej. Przykładowo, struktury spinek RNA mają dużą tendencję do rozplatania się i tworzenia alternatywnych struktur dupleksów przy wysokich stężeniach RNA i soli, typowo stosowanych podczas krystalizacji. W konsekwencji, wielu biologicznie istotnych natywnych motywów RNA nie można uchwycić w jednorodnym stanie krystalicznym bez dodatkowych strategii stabilizacyjnych.

Celem niniejszej rozprawy było opracowanie nowych strategii stabilizacji struktur cząsteczek RNA. Zastosowano dwa podejścia: kowalencyjna stabilizacja za pomocą nienukleozydowych linkerów oraz stabilizacja z użyciem małocząsteczkowego ligandu.

W przypadku pierwszej strategii zaprojektowano, zsyntetyzowano i włączono do łańcucha RNA serię aromatycznych linkerów, mających na celu kowalencyjne połączenie końców 3' i 5' przy jednoczesnym zachowaniu natywnej geometrii RNA. W drugim podejściu przebadano zdolność stabilizacji RNA za pomocą dimeru karbaminianu naftyrydyny (NCD). Wykazano, że wiązanie małocząsteczkowego ligandu może stabilizować strukturę RNA poprzez mechanizmy zasadniczo odmienne od metod kowalencyjnych. NCD działa jak klamra spinająca reszty nukleotydowe zarówno w obrębie tej samej jak i dwóch sąsiednich cząsteczek RNA działając jako „klej molekularny".

Ostatnia część pracy koncentruje się na ocenie fizykochemicznej i biochemicznej konstruktów RNA, obejmującej badania denaturacji termicznej monitorowane spektroskopią UV, pomiary dichroizmu kołowego oraz analizy cięcia RNA in vitro z użyciem nukleazy Dicer. Techniki te pozwoliły ocenić wpływ zastosowanych strategii stabilizacji na integralność strukturalną RNA, jego właściwości fałdowania oraz aktywność biologiczną.

Podsumowując, niniejsza rozprawa przedstawia dwie strategie ukierunkowane na stabilizację cząsteczek RNA w badaniach strukturalnych. Uzyskane wyniki poszerzają wiedzę na temat możliwości wykorzystania zarówno kowalencyjnych łączników nienukleozydowych, jak i ligandów małocząsteczkowych w pokonywaniu barier krystalizacji RNA.

3. The aim of the study

The aim of this doctoral work is to develop, validate, and structurally characterize innovative chemical strategies for stabilizing RNA molecules, facilitating their use in high-resolution crystallographic studies. RNA molecules often exhibit significant structural flexibility, dynamic conformational equilibria, and a strong tendency to form alternative, secondary structures. All of these features severely limit their ability to crystallize and hinder reliable structural characterization. This PhD thesis aims to address these challenges by introducing covalently bonded non-nucleosidic linkers and ligand-mediated stabilization as complementary approaches to enforce RNA structural homogeneity.

Achieving these goals required the implementation of the following tasks:

1. Design and synthesis of a series of non-nucleotidic linkers derived from pyromellitic and naphthalene compounds and incorporate them into RNA oligonucleotides. The goal of these modifications is to achieve stabilization of RNA hairpin and duplex architectures, by covalent joining 3’ and 5’ ends of RNA strands, without perturbing the native base pairing or overall RNA topology.

2. Crystallization and structural analysis of stabilized RNA. Determination of the impact of incorporated non-nucleotidic linkers on RNA structure and formation of the crystal lattice constants.

3. Elucidation of RNA stabilization and crystal lattice formation by the naphthyridine carbamate dimer (NCD) molecule. This task includes determining crystal structures of RNA–ligand complexes and unbounded RNA to define the binding mode and structural effect on RNA upon ligand binding.

4. Biophysical and biochemical evaluation of stabilized RNA construct using thermal denaturation of RNA monitored by UV spectroscopy, circular dichroism and in vitro cleavage of modified RNA using Dicer nuclease.

4. Introduction

4.1. RNA Structure and Stability

RNA molecules exhibit extraordinary functional diversity within biological systems, ranging from catalytic activity to the regulation of gene expression, a versatility that critically depends on their ability to fold into complex three-dimensional structures (1, 2). This structural heterogeneity, which spans the primary sequence, secondary motifs, and tertiary interactions, enables RNA to participate in complex molecular recognition events and perform diverse cellular roles (3). Messenger RNAs (mRNAs) serve as essential carriers of genetic information, transferring the DNA code to ribosomes, where it is translated into proteins with the help of transfer RNAs (tRNAs) (4–7). In addition, small RNAs such as small nuclear RNAs (snRNAs), microRNAs (miRNAs), and small interfering RNAs (siRNAs) are key regulators of gene expression, influencing splicing, degradation, or translational repression of specific mRNAs (6, 8). Other classes of RNA further expand this repertoire: riboswitches in bacteria sense cations and small metabolites, adopting alternative conformations that inhibit protein synthesis (9, 10), whereas catalytic RNAs (ribozymes) mediate cleavage and ligation reactions during the processing of viral, bacterial, and eukaryotic RNAs (11). Collectively, these examples highlight the central role of RNA not only as an information carrier but also as a dynamic regulator and catalyst in fundamental biological processes.

With the growing interest in RNA biology has come the recognition that understanding RNA activities requires detailed knowledge of their structures and conformational dynamics (12). Such insights often combine mechanistic studies of catalytic reactions with structural characterizations of ground states, transition states, and functional groups involved in catalysis (13). RNA structure is typically described at three levels: the primary structure, corresponding to the nucleotide sequence; the secondary structure, consisting of Watson–Crick base pairs and common motifs such as helices, hairpins, bulges, internal loops, and junctions. Various types of single-stranded regions occur between helices, including hairpin loops (linking the two sides of a single helix), bulges and internal loops (connecting two helices), and multi-branched loops (joining three or more helices) (14). The tertiary structure arises from interactions between distinct secondary elements. While the secondary structure is relatively well understood, predicting or determining tertiary contacts has long posed a

major challenge, which has only recently been overcome by advances in structural methods such as X-ray crystallography and NMR spectroscopy (13, 15). High-resolution RNA structures, including tRNA (16–18), the hammerhead ribozyme (19–22), the P4–P6 domain of the Tetrahymena group I intron (23), and the hepatitis delta virus ribozyme (24), have demonstrated that tertiary interactions are crucial for shaping the global fold of the molecule. Thus, RNA folding can be viewed as a hierarchical process in which the primary sequence dictates the secondary structure, which in turn constrains the tertiary organization, with only minimal alteration of the initial base-pairing pattern (25).

4.2. Crystallography of RNA: Principles and Challenges

RNA plays a central role in diverse cellular processes, and in many cases, its biological activity is dictated by its three-dimensional structure. High-resolution RNA structures are most reliably obtained using X-ray crystallography, which enables comparisons between related RNAs and identification of conformational changes associated with their function. However, crystallization remains a major bottleneck in RNA structure determination, as obtaining well-ordered crystals suitable for diffraction is inherently challenging. The crystal structures of ribozymes and ribosomal subunits have provided crucial insights into the principles of RNA folding, including the packing of helices into compact structures that create catalytic centers and binding sites. Yet, much of RNA biology can only be elucidated through further crystallographic studies (12, 26).

Despite their importance, RNA 3D structures remain rare compared to proteins (27, 28). This scarcity reflects the intrinsic properties of RNA, such as conformational flexibility, misfolding, and the ability to adopt multiple folded states (29–31). Unlike proteins, whose chemically diverse side chains facilitate crystal contacts, RNA presents a largely uniform, negatively charged phosphate backbone that hinders lattice formation (12, 32, 33). Moreover, RNA molecules are often elongated and loosely packed, with high solvent content, and their relatively weak tertiary interactions further increase their flexibility and inter-domain motion. Consequently, many nucleic acid crystals diffract poorly and only at low resolutions (34). Although strategies such as the engineering of modular RNA domains can promote crystallization, RNA and RNA–protein complexes remain significantly more difficult to crystallize than proteins.

The crystallization of RNA molecules beyond simple oligonucleotide duplexes remains a major challenge for X-ray structure determination. Even biochemically, covalently, and conformationally homogeneous RNA molecules often fail to form sufficiently large crystals for data collection. When crystals are obtained, they frequently diffract to low resolution, which is insufficient for detailed structural analysis. This difficulty arises from the largely undifferentiated molecular surface, which is dominated by the negatively charged phosphate backbone. During crystal nucleation and growth, neighboring molecules may pack out of register, leading to premature termination of crystal growth and disorder (34). Additional factors, such as conformational heterogeneity, flexible termini and loops, and the chemical lability of the 2'-hydroxyl group, further hinder the formation of well-ordered crystals (35).

To mitigate these challenges, one strategy involves reducing large nucleic acid molecules to functionally important domains while preserving their biological activity (36, 37). Since RNA folding follows a hierarchical process from primary to secondary and subsequently to tertiary structures, such modular domains can serve as useful building blocks for both experimental and computational studies. Modern prediction tools, including RNAStructure and RNAComposer, exploit structural data deposited in the Protein Data Bank by assembling fragments of experimentally solved three-dimensional structures to model the architecture of novel RNAs (38–40).

RNA hairpin structures serve as illustrative example. However, these motifs often unfold under experimental conditions, such as those employed in crystallographic or nuclear magnetic resonance (NMR) studies. The primary reason is that hairpins are thermodynamically unstable at the high RNA and/or salt concentrations typically required for crystallization. Under such conditions, they tend to unfold and form duplexes. This process is thermodynamically driven, and sequence alterations do not resolve the issue, as any hairpin with complementary ends can readily form a duplex with another strand of the same sequence (Figure 1) (41–43) This intrinsic instability is reflected in the near-absence of RNA hairpin crystal structures reported in the literature.

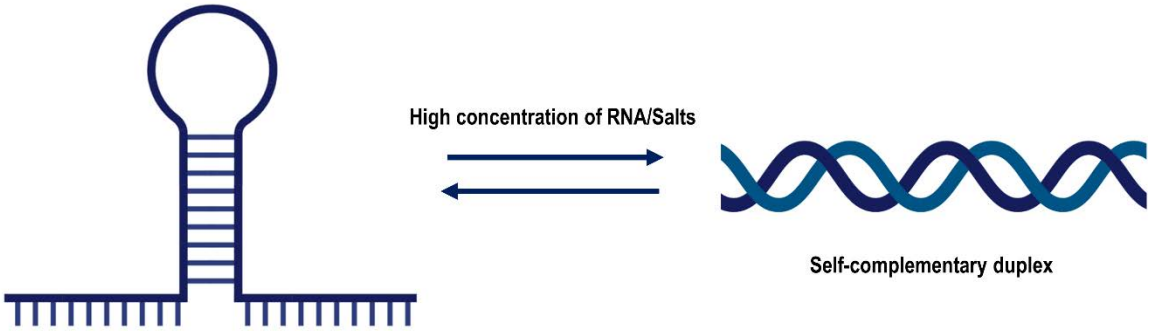


Figure 1. Hairpin destabilization under crystallization conditions. High RNA or salt concentrations promote unfolding of RNA hairpins and formation of intermolecular duplexes.

4.3. Stabilization methods - Chemical modifications that increase RNA stability

Chemical modifications of RNA have become indispensable tools for improving stability in therapeutic contexts and structural biology (44–47). A wide range of backbone, sugar, and nucleobase modifications, such as 2'-O-methyl, 2'-fluoro (2'-F), locked nucleic acids (LNA), phosphorothioates, bridged nucleic acids (BNA), and complete backbone replacements, are now routinely employed to increase resistance to nuclease degradation and enhance thermal stability (48–51). Moreover, site-specific chemical substitutions, such as selenium derivatization, offer distinct advantages for phasing in crystallographic studies (52–56). These modifications often enhance the hybridization affinity and overall molecular robustness. However, they can also introduce local structural perturbations that complicate high-resolution structural analyses. Therefore, it is crucial to carefully evaluate each modification method. The balance between biochemical stability and the risk of structural artifacts must be considered when designing RNA molecules for crystallographic or functional studies.

4.3.1. Non-nucleosidic linkers & covalent restraints: concepts and types

In structural studies of nucleic acids and their complexes, controlling molecular flexibility is often crucial to obtain well-ordered crystals and reproducible conformations suitable for high-resolution analysis. One of the most effective strategies for achieving this control is the introduction of non-nucleosidic linkers and covalent restraints, which help rigidify specific regions of the molecule or stabilize defined spatial arrangements (57–59).

These are known as interstrand crosslinking agents (ICL) (Figure 2A). Cross-linking agents are substances capable of reacting with two distinct sites on oligonucleotide sequences, leading to various lesions. These compounds facilitate interstrand cross-linking by covalently bonding to both ends of the same strand, preventing their separation (60). Their bifunctional groups can covalently link the 3'- and 5'-OH ends of nucleic acids, acting as interstrand linkers (Figure 2B). When they connect opposing strands, they are referred to as linkers (Figure 2C). Commercially available ICL compounds include 4,4'-dimethoxytrityl and (2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite groups, which are essential for their integration into RNA/DNA sequences through solid-phase chemical synthesis (61–63).

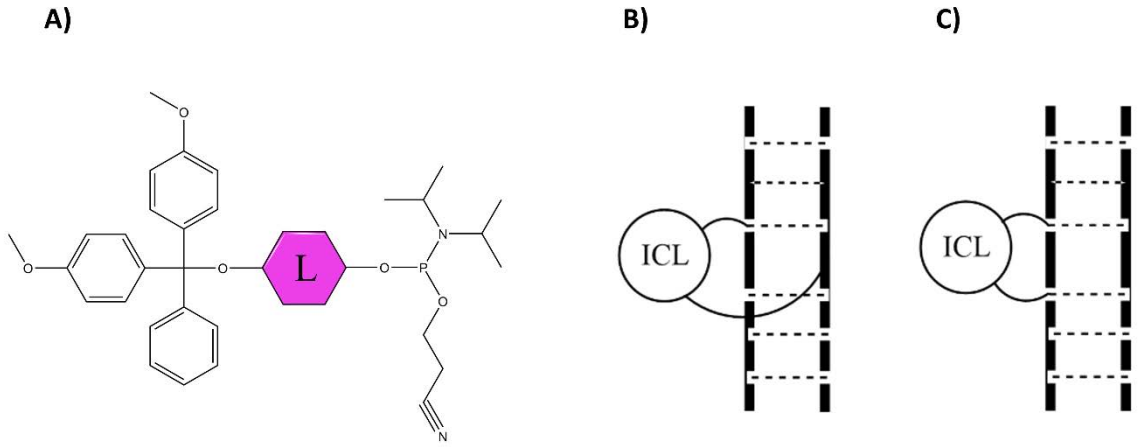


Figure 2. Schematic representation of interstrand and intrastrand linker compounds (ICLs). A) Bifunctional groups such as 4,4’-dimethoxytrityl and (2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite can covalently connect the 3’- and 5’-hydroxyl ends of nucleic acids. Symbol L corresponds to the core of the linker. B) interstrand linker C) intrastrand linker.

Cross-linking agents, through their capacity to stabilize molecular structures and modulate biological interactions, has emerged as a crucial tool in the development of novel therapeutic strategies. This approach relies on the formation of covalent bonds to modify molecular characteristics, enabling the design of biomaterials with tailored mechanical and functional properties for diverse healthcare applications (60, 64, 65). These compounds can be integrated at any point within an oligomer, enabling the precise manipulation of DNA or RNA properties. ICL compounds may exhibit specific physicochemical characteristics, such as fluorescence, which allow modified DNA/RNA oligomers to serve as molecular probes in diagnostics and as tools for monitoring biological processes. A notable application of these oligomers is in studying of the mechanisms of DNA repair systems within cells (36, 60, 66).

In the context of nucleic acids, cross-linking can elucidate and irreversibly stabilize secondary and tertiary structures. For instance, the discovery of a DNA–RNA hybrid-type G-quadruplex through copper-based cycloaddition illustrates how cross-linking can reveal complex structural motifs (60, 67–69). In chemotherapeutics, numerous cross-linking agents with intrinsic reactivity—particularly bifunctional alkylating agents and platinum-based compounds—are employed to inhibit cell replication and transcription by covalently modifying nucleic acids (60, 70–72).

This principle extends to various nucleic acid-based therapeutic modalities, including antisense therapy, transcription factor decoy oligonucleotides, and RNA interference, where precise molecular anchoring can enhance target specificity and therapeutic efficacy. Introducing cross-links between the strands of a decoy duplex—or closing the duplex on both ends to form cyclic or dumbbell structures—significantly increases stability toward endonuclease degradation (73–76). Such stabilization has been correlated with successful inhibition of target protein expression (73, 77, 78). Furthermore, cross-linking has been demonstrated to augment the inhibitory effects of oligonucleotides in RNA interference pathways, as exemplified by psoralen-modified analogues that exhibit superior RISC function inhibition compared to their unmodified counterparts (60, 79, 80).

4.3.1.1. Linear non-nuclosidic linkers

Linear non-nucleosidic linkers serve as crucial tethers in chemical biology, facilitating the controlled connection of molecular entities in both reversible and irreversible systems (81). In the context of synthetic oligonucleotides and nucleic acid analogues, these linkers act as versatile structural elements that introduce flexible, chemically stable spacers between nucleotide residues. This spacing allows the fine-tuning of molecular conformation, hybridization behavior, and the spatial orientation of functional groups.

Chemically, these linkers typically consist of an aliphatic carbon chain interspersed with amino or glycol groups (often termed glycol linkers). They are available in various lengths, typically ranging from 3 to 18 carbon atoms (Figure 3) (37, 82). Short carbon linkers are used in RNA-protein functional studies (83, 84) and as an effective inhibitor of DNA polymerase activity in PCR reactions (85, 86).

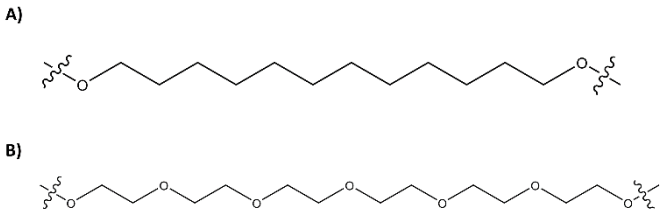


Figure 3. Examples of linear linkers. A) C12 linker B) Hexaethylene glycol linker

Glycol linkers (PEG) are hydrophilic and structurally flexible. They serve as substitutes for apical loops in RNA hairpins (Figure 4B), oligomer-based catalysts, and self-assembly systems (82, 87–90). They are commonly incorporated during solid-phase oligonucleotide synthesis to add specific functionalities or spatially separate domains within RNA constructs (91). This approach enables site-specific labeling, allowing for a detailed investigation of RNA structure, environment, and interactions in cellular contexts.

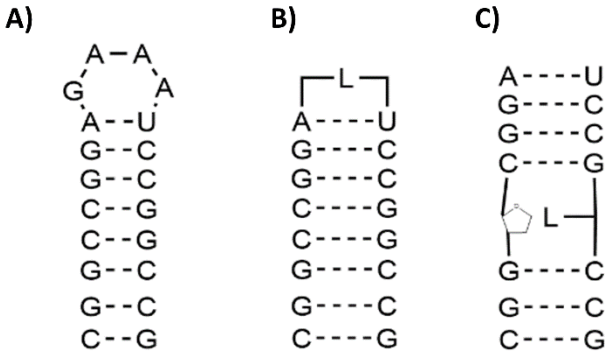


Figure 4. Examples of bifunctional linker applications: a) hairpin structure in native form b) replacement of the apical loop of the hairpin with a non-nucleotide linker c) intercalation of the RNA helix by the linker resulting from the absence of a nucleobase.

In the most direct strategy, crosslinks can substitute nucleotide sequences to connect distinct domains within higher-order nucleic acid structures. They may also function as tethers between independently hybridized regions or between a nucleic acid and an external ligand or reporter. Among the possible linkers, ethylene glycol chains offer a particular advantage: their hydrophilic ether–ethyl pattern favors hydration and an extended conformation in aqueous environments, unlike simple carbon chains that tend to collapse owing to hydrophobic effects. This property allows ethylene glycol–based linkers to efficiently bridge

distant sites within macromolecules, and their derivatives are readily available and easily modified for cross-linking applications.

Oligonucleotides incorporating such polyethylene linkers exhibit valuable properties, including enhanced stability in biological environments, improved cellular uptake, and optimized hybridization efficiency (92). In particular, oligonucleotides with modified apical loops are resistant to nuclease degradation in cellular media and often demonstrate greater thermodynamic stability than their native counterparts (37, 82, 86, 86, 89, 93). However, the stability of such RNA molecules is highly dependent on the length of the introduced linker and the type of counterions present in the aqueous buffer.

Furthermore, replacing nucleotide sequences with suitable linkers can significantly simplify the structure of functional nucleic acid domains without substantial loss of biological activity (36, 37). Another major advantage of substituting nucleotides with linkers is the substantial reduction in the number of synthesis steps required, which decreases production costs and accelerates the overall synthesis timeline, as a single extended linker can replace multiple nucleotide residues. Additionally, conjugation with small synthetic molecules enables precise modulation of oligonucleotide function, further broadening their potential in structural and therapeutic applications (94–97).

Linear non-nucleotide linkers have valuable applications in structural biology. They facilitate structural simplification of functional nucleic acid domains while maintaining biological activity, thus acting as stabilizing elements for RNA structures. For instance, hexaethylene glycol linker (commercial: spacer 18) have been employed to stabilize RNA in its hairpin form. These non-nucleotide linkers are positioned opposite to the apical loop, forming the base of the hairpin. The hairpin structure was circularized via an enzymatic approach, specifically through ligation using T4 RNA ligase, which connects the 5’-phosphate and 3’-OH groups. Although certain structural and sequential criteria must be met for efficient ligation, this method retains the native backbone of the hairpin RNA (Figure 5) (98).

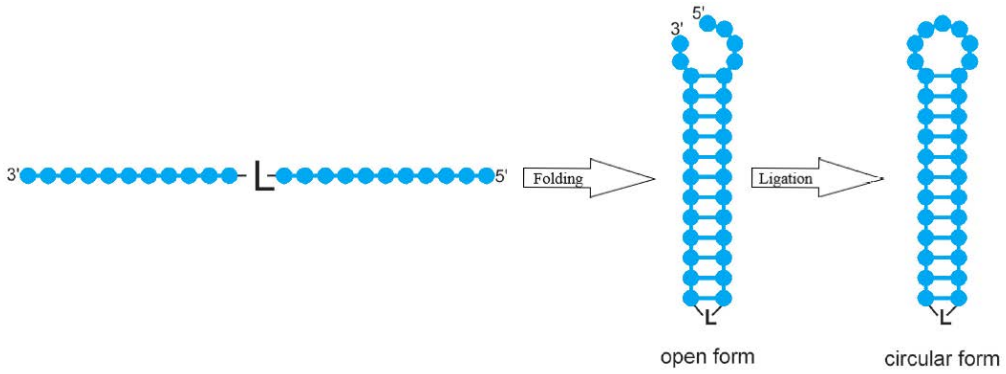


Figure 5. Stabilization of the hairpin structure by non-nucleotide linkers (-L-) following enzymatic circularization.

There is one reported structure of DNA containing C12-alkyl linker forming loop region. Alkane linkers derived from α,ω-alkane diols have been used in the synthesis of DNA mini-hairpins to probe short base-pair domains. Structural studies using NMR revealed that an n-dodecane (C12) linker adopts a curved conformation parallel to adjacent base pairs, resulting in a shortened P–P distance (~15.4 Å) relative to B-DNA. Despite this compression, canonical Watson–Crick pairing was preserved, suggesting that the linker enhances structural stability through hydrophobic interactions while maintaining native duplex geometry (99).

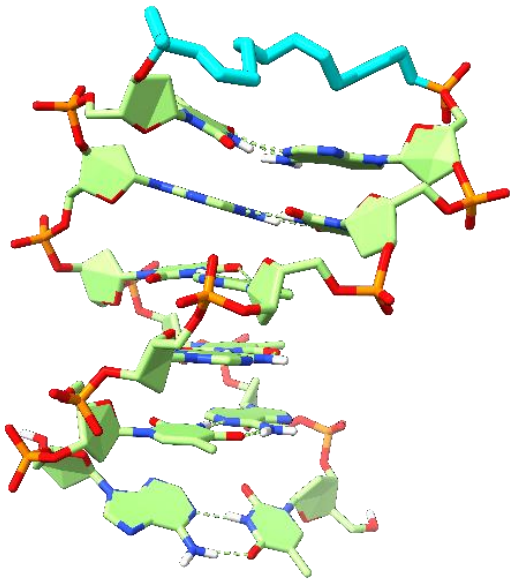


Figure 6. Hairpin of A-T basepairs having a C12-alkyl linker (cyan) forming the loop region (pdb code: 2L13).

4.3.1.2. Aromatic non-nucleosidic linkers

The second type of non-nucleosidic linker is comprised of aromatic compounds. Unlike linear linkers, they contain at least one aromatic ring substituted with a phosphate group on one side and a hydroxyl group on the other connected through an aliphatic chain. These aromatic linkers exhibit lower conformational flexibility and greater structural compatibility with nucleic acids, making them attractive for constructing the tertiary structures of DNA and RNA. The aromatic ring enables stacking interactions with nitrogenous bases (100), supporting the maintenance of native nucleic acid structure and biological functionality. Compounds that serve as suitable representatives of this group of linkers include a) stilbene (93, 101–103), b) phenanthrene, and c) pyrene and its derivatives (Figure 7) (104–109).

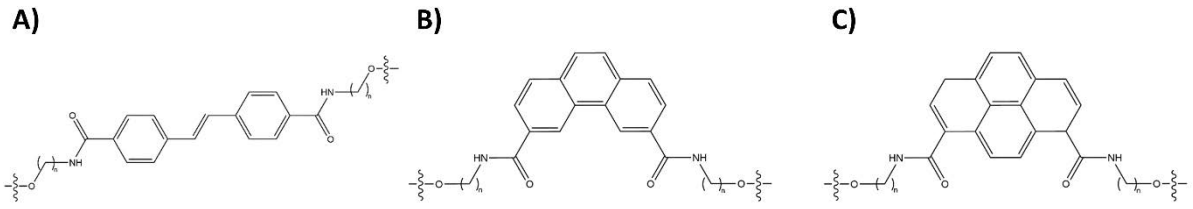


Figure 7. Aromatic linkers: A) Stilbene, B) Phenanthrene, C) Pyrene

4.3.1.2.1. Stilbene linkers

The stilbenedicarboxamide linker was particularly effective in stabilizing the structures of short RNA/DNA duplexes and triplexes. It features a relatively rigid core that restricts the structural flexibility of the linker, along with a fluorescent photoreactive chromophore that acts as a reporter group or induces photoinduced conformational changes in the linker, thereby influencing the hybridization properties of oligonucleotides (101–103, 110). Similarly to hexaethylene glycol, stilbene linker was used in stabilization of RNA hairpin structure (Figure 5) (98).

Stilbene linkers, particularly stilbenediether linkers, have been successfully employed in crystallography to investigate the structural properties of nucleic acid constructs, such as DNA hairpins (101, 103, 111, 112). These linkers are crucial for stabilizing molecular architecture and studying specific interactions. The crystal structure of a DNA hairpin containing a bis(2-hydroxyethyl)stilbene-4,4'-diether linker, resolved at 1.5 Å, revealed that the linker exhibits notable conformational flexibility, adopting both face-to-face and edge-to-face orientations with adjacent base pairs (101). These linkers promote π-stacking interactions and facilitate

unique lattice contacts that are distinct from those in the canonical B-form DNA (111). The compact arrangement of the stilbene chromophore, with a plane-to-plane separation of approximately 3.25 Å from the neighboring G:C base pair, enhanced the structural stability of the hairpin (103). These findings underscore the role of stilbene linkers in modulating the nucleic acid architecture and intermolecular interactions (Figure 8).

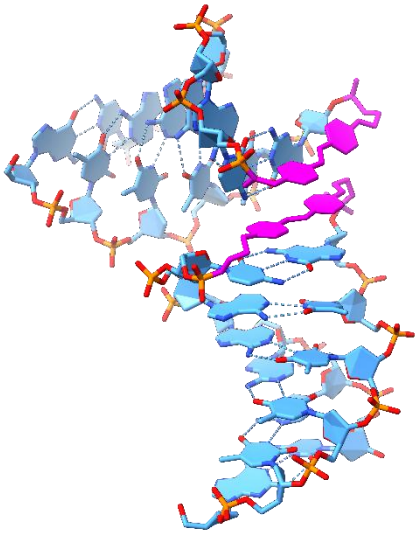


Figure 8. Crystal structure of the hairpin d(GT_4G)-Sd_2-d(CA_4C) determined at 1.5 Å resolution. The asymmetric unit contains two molecules differing in both the linker (magenta) and stem conformations. In one molecule, the planar stilbene moiety stacks directly on the adjacent G:C base pair, whereas in the other, the stilbene exhibits a significant rotation between the phenyl rings and adopts an unusual edge-to-face orientation relative to the base pair.

4.3.1.2.2. Phenanthrene and Pirene linkers

Aromatic linkers based on phenanthrene and pirene compounds are essential non-nucleosidic components in nucleic acid chemistry, used to introduce specific functionalities, spatially separate domains, and enhance structural stability. As in the case of stilbene linkers, its relatively rigid aromatic core constrains the conformational flexibility of the linker. Common examples include porphyrin, azo-benzene, perylene, phenanthrene, pyrene, phenanthroline, stilbene, and tetrathiafulvalene derivatives, typically incorporated into oligonucleotides via phosphoramidite chemistry (113). Their defining feature is the ability to engage in strong π–stacking interactions with neighboring bases and with each other, markedly increasing duplex stability (113). Phenanthroline and phenanthrene derivatives, for instance, form particularly

stable hybrids, with phenanthrolines yielding greater duplex stability owing to their higher dipole moment and stronger interstrand stacking (105, 106).

Linkers based on phenanthrene and pyrene have also been used in the synthesis of therapeutic oligomers that act as molecular “prostheses” for damaged DNA (114). During depurination or depyrimidization, mutations involving the loss of nitrogenous base—DNA strands are prone to cleavage and degradation, ultimately leading to cell death (115–118). When opposite apurinic or apyrimidinic sites are incorporated, phenanthrene and pyrene linkers can intercalate with neighboring bases, thereby stabilizing the double helix and protecting DNA from further degradation (Figure 4C) (119, 120). The length and flexibility of these aromatic linkers remain crucial, as sufficient conformational freedom enables optimal stacking and alignment within the duplex, particularly in constructs containing abasic sites (105, 106).

Aromatic linkers also serve as scaffolds for the ordered arrangement of aromatic systems, such as pyrenes, within DNA or RNA frameworks, facilitating the design of complex molecular architectures (121). They have been employed to replace hairpin loops in RNA, forming stable mimics with enhanced hybridization properties, and expanding the utility of modified nucleic acids across diagnostics, materials science, and medicinal chemistry (104–106, 113).

Crystallographic analysis revealed distinct aromatic stacking interactions between the pyrene and phenanthrene moieties incorporated within the B-DNA duplex framework. In duplexes containing a single modification in each strand near the terminal region, the two polyaromatic hydrocarbons adopted a face-to-face stacking orientation, indicative of strong π–π interactions. However, due to crystal packing constraints and steric effects, the terminal GC base pair is disrupted in all three crystal structures, resulting in a non-ideal stacking geometry of the aromatic chromophores in two of the complexes. In the hybrid containing three pyrene modifications, crystal lattice-induced end-to-end stacking of individual DNA duplexes gave rise to an extended π-stack comprising four co-axially aligned pyrene units, with their aromatic planes oriented in a parallel fashion. These observations underscore the ability of hydrophobic polyaromatic systems to drive supramolecular DNA assembly, and demonstrate the utility of co-crystallization with the recombinant binding domain of the restriction enzyme BpuJI to elucidate the structural organization of DNA-assembled oligochromophores at atomic resolution (107).

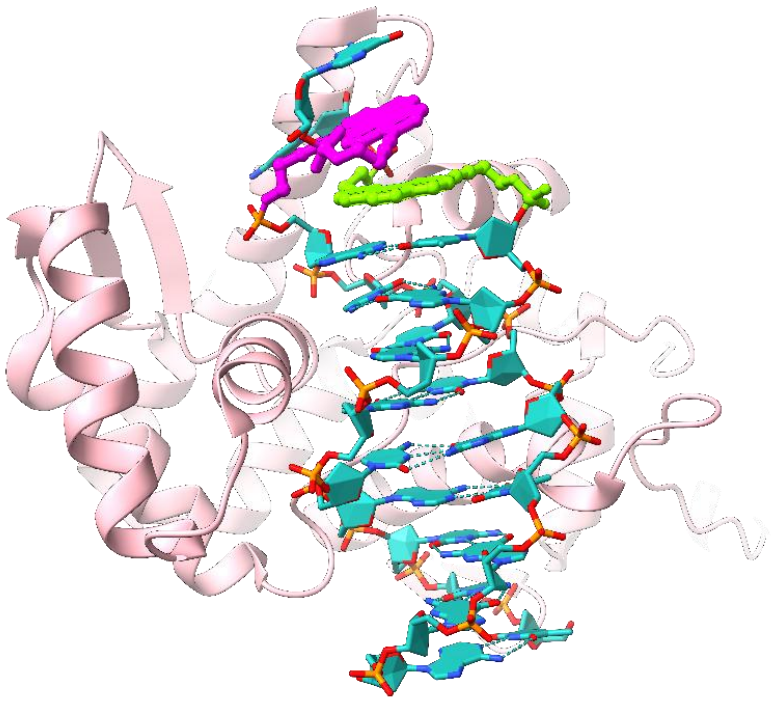


Figure 9. Structure of the BpuJI:DNA complex (pdb code: 5HNF). The protein is represented in light pink. The phosphate backbone of the DNA is linked to the phenanthrene (lime) and pyrene (magenta) modifications.

In structural research 8-aminomethyl-2-quinolinecarboxylic acid and 8-amino-2-quinolinecarboxylic acid dimer was used as a both a B-DNA duplex hairpin turn and a single helical (MQ)n extension, thereby placing the grooves and arrays of negative charges of the two segments in register (122, 123). Single-crystal X-ray diffraction studies revealed that the diamide quinoline–based linker effectively connects DNA and aromatic foldamer segments, while preserving the structural alignment between their helices. Despite their different chemistries, the linker maintains the helical register and groove continuity characteristics of B-DNA. Circular dichroism confirmed that the foldamer's M-handedness was directed by the d-stereochemistry of DNA, and molecular modeling supported the observed compatibility. These findings demonstrate that the linker enables stable DNA hairpin architecture and provides a structural basis for developing DNA–foldamer hybrids as selective inhibitors of DNA-binding proteins (124).

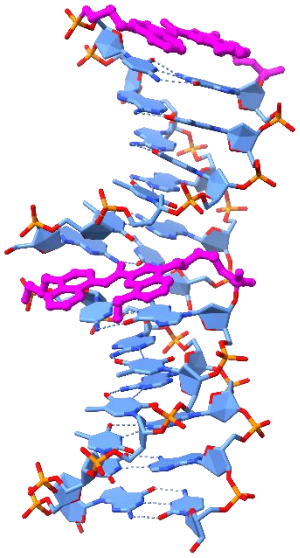


Figure 10. Racemic crystal structure of a synthetic DNA Hairpin (blue) with diquinoline linker (magenta) (pdb code: 8Q60).

4.3.2. Ligand-Induced Stabilization of RNA Structures

RNA forms intricate three-dimensional structures that play diverse functional roles in human biology and disease. Efforts to therapeutically target these RNA structures with small molecules are actively advancing and supported by key developments in the field. These include the creation of computational tools that predict evolutionarily conserved RNA structures, and the formulation of strategies that broaden the mode of action and enhance interactions with cellular machinery (125–130). Gaining a deeper insight into the fundamental principles of molecular recognition remains essential concerning RNA conformational dynamics and the biophysical intricacies of RNA-ligand interactions (131–135). These interactions, which are central to biological processes such as gene regulation via riboswitches, involve high-specificity recognition mediated by hydrogen bonding, electrostatic interactions, van der Waals forces, and stacking interactions, often engaging nearly all functional groups of the ligand (136–141). The intrinsic flexibility of non-coding RNAs, enabling multiple conformations, is crucial for their regulatory roles yet complicates small-molecule design (142–147).

Targeting nucleic acids with small molecules has recently emerged as a promising therapeutic strategy, particularly for cancer and bacterial and viral diseases. Modulating gene expression at the mRNA level offers a method to influence otherwise undruggable proteins without altering the genome. Despite this potential, most molecular docking tools remain optimized for proteins (144, 148–152). RNA's inherent structural flexibility of RNA complicates the design of selective small-molecule binders and underscores the importance of elucidating how ligand interactions modulate their conformational landscape (132, 139, 144, 152, 153). The negatively charged nature of RNA favors interactions with positively charged groups, metal ions, or solvating water molecules, whereas its binding pockets are typically shallow, polar, and hydrated, which are distinct from those in proteins (134, 139, 148, 154–158). A deeper structural understanding of RNA–ligand recognition is therefore essential for the rational design of RNA-targeting therapeutics, particularly given the limited number of high-resolution RNA–ligand complex structures currently available (134, 135, 151, 157, 159–161).

4.3.2.1. Mismatch binding ligands in RNA Crystallization

Accurate determination of RNA three-dimensional structures remains a significant challenge due to inherent difficulties in crystallization, including low yields of pure, properly folded RNA, limited sequence diversity for crystal contacts (29). A promising strategy to address these challenges involves the use of small molecules known as mismatch-binding ligands (MBLs), which selectively recognize non-canonical base pairs or mismatches in RNA (159). MBLs are synthetic molecules that were originally designed in silico based on the concepts of hydrogen bonding and semi-intercalation with nucleic acids. Most MBLs consist of two heterocyclic aromatic units possessing hydrogen-bonding surfaces that are fully or partially complementary to those of the nucleotide bases. In this design, each heterocycle interacts with one of the mismatched bases via hydrogen bonding to form pseudo-base pairs that stack with the adjacent canonical base pairs (162). The aromatic components of MBLs are typically engineered to complement the hydrogen-bonding patterns of nucleobases (163, 164). Structural analyses using NMR spectroscopy and X-ray crystallography have demonstrated that these ligands interact directly with nucleotides, forming Watson-Crick-like pairs (163, 165–168). In certain cases, ligand binding is associated with conformational rearrangements in the RNA structure (163, 166). By stabilizing specific RNA conformations, MBLs promote

crystal formation, reduce structural heterogeneity, and increase the probability of generating well-ordered crystal lattices (35).

4.3.2.1.1. Naphthyridine Carbamate Dimer (NCD)

Naphthyridine carbamate dimers (NCDs) were developed from first-generation naphthyridine derivatives to enhance binding affinity toward G-rich sequences and improve chemical stability by replacing the labile amide linkage with a more robust carbamate linkage (Figure 9A) (162, 164). This structural modification significantly strengthened binding to G–G homomismatches and increased molecular stability, establishing the carbamate linkage as a superior guanine recognition motif (162, 169–171). The optimized NCD scaffold enables effective recognition of 5′-CGG-3′/5′-CGG-3′ sequences, which frequently occur in slip-out structures of expanded trinucleotide repeats that are notoriously difficult to target therapeutically (162, 167, 172–174). Moreover, previous studies have demonstrated that NCD functions as a molecular glue, promoting the assembly and stabilization of higher-order nucleic acid architectures. NCD and its multimeric derivatives have been employed to induce DNA folding into tetrahedral structures, modulate RNA aptamer activity, and enhance the catalytic efficiency of hammerhead ribozymes (175–178).

Recently, the naphthyridine carbamate dimer (NCD) was shown to specifically recognize pathogenic RNA repeats 5′-UGGAA-3′/5′-UGGAA-3′ associated with spinocerebellar ataxia type 31 (SCA31) (179). SCA31 is an autosomal neurodegenerative disorder characterized by the abnormal expansion of r(UGGAA) pentanucleotide repeats, which sequester various RNA-binding proteins (RBPs) through the formation of toxic nuclear foci (174, 180–184). In vitro pulldown assays demonstrated that NCD disrupted the interaction between r(UGGAA)n repeat RNA and several RBPs. Consistent with these findings, NCD treatment suppressed the formation of RNA foci in HeLa cells overexpressing r(UGGAA)n RNA and inhibited the assembly of endogenous nuclear stress bodies (nSBs) upon thermal stress, thereby affecting nSB-dependent splicing regulation (159, 174, 185, 186). Furthermore, administration of NCD to Drosophila larvae expressing toxic r(UGGAA)n repeats led to marked alleviation of the disease phenotype, likely resulting from the inhibition of RBP–RNA interactions, as confirmed both in vitro and in cellular models. These studies demonstrate the therapeutic potential of small molecules that directly target toxic repeat RNAs. Nuclear magnetic resonance (NMR) analysis of the NCD–UGGAA complex (Figure 9C) revealed a distinct bound RNA conformation

characterized by hydrogen bonding between the naphthyridine and guanine residues, zigzag binding orientation of two NCD molecules at the pentad site (Figure 9B), adenine base flipping, and adoption of a syn glycosidic conformation by all NCD-bound guanines (159, 186).

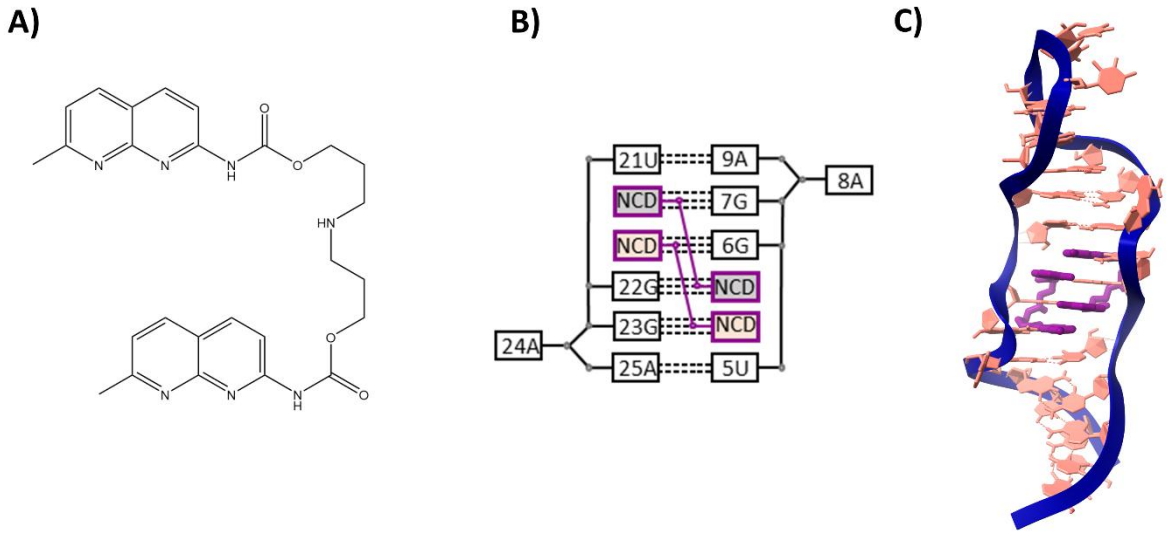


Figure 11. Interactions between NCD ligand and nucleic acids. A) Chemical structure of NCD molecule. B) Secondary structure diagrams and C) tertiary structures of interactions between the UGGAA motif and NCD molecules (purple) in NMR models.

5. Materials and methods

5.1. Materials

5.1.1. Chemical and biochemical reagents

Most reagents were purchased from Merck, Biosynth, Fluorochem, Acros Organics, Bio-Rad, and Honeywell. All reagents were of the highest quality and free from ribonucleases. All filters were purchased from Millipore or Sartorius, and NAP-25 columns from GE Healthcare. Phosphoramidites for chemical synthesis with a 2'-O-TriisopropylsilylOxyMethyl protection group were obtained from Chemegene. Spacer 18 was purchased from LGC Link. Crystallization screens were obtained from Hampton Research.

5.1.2. Enzymes

T4 RNA ligase, T7 DNA polymerase, FastAP^{TM} alkaline phosphatase, T4 Polynucleotide Kinase were purchased from Thermo Fisher.

5.1.3. Aqueous solutions

Electrophoresis buffer under denaturing conditions

10× TBE 890 mM Tris, 890 mM boric acid, 20 mM EDTA, pH 8.3

40% polyacrylamide gel solution (29:1)

193.33 g acrylamide, 6.66 g bisacrylamide, H_2O up to 500 mL

Gel for RNA electrophoresis under denaturing conditions

	8%	10%	12%	15%	20%
40% ACRYLAMIDE/BISACRYLAMIDE SOLUTION, 29:1 (W/W)	10 ml	12,5 ml	15 ml	18,75 ml	25 ml
10× TBE (FINAL CONCENTRATION 1×)	5 ml	5 ml	5 ml	5 ml	5 ml
UREA (FINAL CONCENTRATION 8 M)	21 g	21 g	21 g	21 g	21 g
H_2O (TO 50ML)	14 ml	11,5 ml	9,5 ml	5,25 ml	-

Buffer for UV melting and CD spectra

0,1 M sodium chloride, 40 mM sodium cacodylate pH 7

In vitro transcription buffer

0,4M HEPES pH 7,5, Spermidine 0,1M, 0,1% TRITON X-100, 0,2M $MgCl_2$, 10mM TCEP Tris(2-carboxyethyl)phosphine

Ligation buffer

0,5M TRIS-HCl pH 7,5, 0,1M $MgCl_2$, 0,1M dithiothreitol

Buffer for RNA cleavage by Dicer

0,25mM NaCl, 20mM Tris-HCl pH 7,5

Crystallization buffer

10mM Sodium Cacodylate

Loading-denaturing buffer

BB: 8 M urea; 0.2% (w/v) bromophenol blue; 0.2% (w/v), XC: 8 M urea, xylene cyanol; 0.2% (w/v)

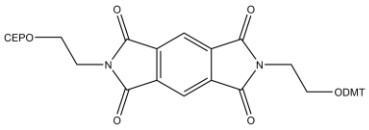
5.1.4. Oligonucleotides

SUMMARY OF RNA SEQUENCES USED IN THE WORK

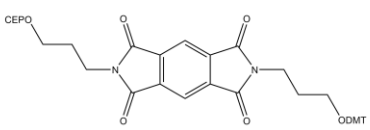
NAME	Sequence 5'→ 3'
T1	CUGCUGCUG
T2	CUGCUGCUG-3a-CUGCUGCUG
T3	CUGCUGCUG-3b-CUGCUGCUG
T4	CUGCUGCUG-3c-CUGCUGCUG
T5	CUGCUGCUG-3d-CUGCUGCUG
T6	CUGCUGCUG-3e-CUGCUGCUG
T7	GCAGCUGCUG-3a-CUGCUGCUG
T8	GACGCUGCUG-3d-CUGCUGCUG
T9	GCAGCUGCUG-3f-CUGCUGCUG
T10	CGCUG-3a-CUGCG
T11	CGCUG-3d-CUGCG
T12	CGCUG-3f-CUGCG
T13	GCUGC
T14	GCUGdC-3a-dGCUGC
T15	GCUGdC-3b-dGCUGC
T16	GCUGdC-3c-dGCUGC
T17	GCUGdC-3d-dGCUGC
T18	GCUGdC-3e-dGCUGC
T19	GCUGdC-3f-dGCUGC
T20	GCUGdC-3g-dGCUGC

T21	CCGUGdG-3a-dCCGUGG
T22	CCGUGdG-3b-dCCGUGG
T23	CCGUGdG-3d-dCCGUGG
T24	CCGUGdG-3e-dCCGUGG
T25	UGCUCCUAGUACGAGAGGACCGGAGUG SRL
T26	AGGACCGGAGCdG-3a-dUGCUCCUAGUACGAG
T27	AGGACCGGAGCdG-3d-dUGCUCCUAGUACGAG
T28	GGAGCCUGGGAGCUCC TAR
T29	GAGCCdC-3a-dGGGCCUGG
T30	GAGCCdC-3d-dGGGCCUGG
T31	GAGCCdC-3f-dGGGCCUGG
T32	GAGCCdC-S18-dGGGCCUGG
T33	GGUCGGUCCCAGACGACC R06
T34	CAGACGCdC-3a-dGGCGGUCC
T35	CAGACGCdC-3d-dGGCGGUCC
T36	CAGACGCdC-3f-dGGCGGUCC
T37	CAGACGCdC-S18-dGGCGGUCC
T38	UAGCUUAUCAGACUGAUGUUGACUGCCGAAUCUCAUGGCAACACCAGUC GAUGGGCUGU premir-21
T39	UCUCAUGGCAACACCAGUCGAUGGGCUGU-S18-UAGCUUAUCAGACUGAUGUU GACUGUUGAA Pre mir-21 S18
T40	GGGACUGGAAGUGCUAGCUUAUCAGACUGAUGUUGACUGUUGAAUCUCAUGGCA ACACCAGUCGAUGGGCUGUGCACUGGAAGUCC PREMIR 21 NCD
T41	GCACUCCGGUAGCUUAUCAGACUGAUGUUGACUGUUGAAUCUCAUGGCAACA CCAGUCGAUGGGCUGUCCAUU premir U1A
T42	AGGACCGGAGCG-S18-UGCUCCUAGUACGAG

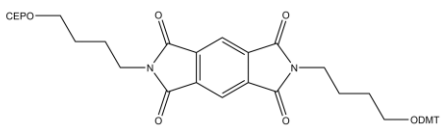
3A



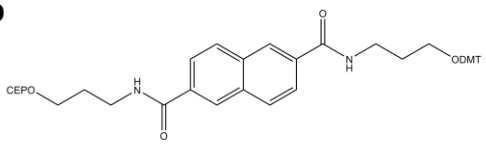
3B



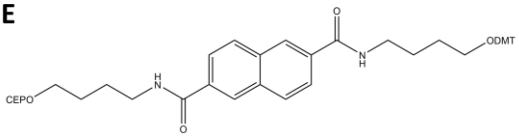
3C



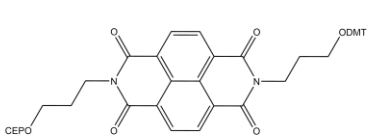
3D



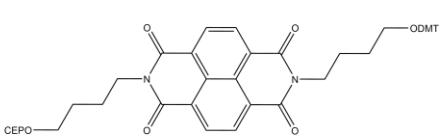
3E



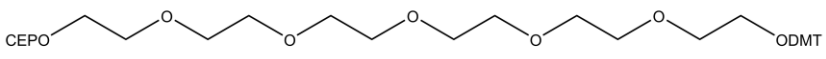
3F



3G



S18



5.1.5. Aparature

- Laboratory centrifuge MIKRO 200 R Hettich
- Vortex Bender and Hobein
- Thermoblock BioSan TS-100C
- Mini shaker IKA MS3 basic
- Nanophotometer IMPLEN N60
- MicroCal PEAQ-DSC Malvern
- 392 ABI DNA/RNA Synthesizer
- Amersham™ Typhoon™ biomolecular imager
- Spectrometer NMR 400 MHz (9.39 T) AVANCE II Bruker
- Jasco J-815 Circular Dichroism (CD) Spectrometer
- Jasco V-650 UV VIS spectrophotometer
- Gryphon crystallization robot Art Robbins Instruments
- Advanced laboratory rotary evaporator - Rotavapor® R-300

5.2. Methods

5.2.1. Thin layer and column chromatography

The progress of the chemical reactions was monitored using thin-layer chromatography (TLC). Chromatograms were developed on glass plates coated with a 0.25 mm layer of silica gel 60 F_{254}. Column chromatography was performed using Merck silica gel 60 (particle size 0.063–0.200 mm).

5.2.2. NMR spectroscopy and mass spectrometry

Nuclear magnetic resonance (NMR) spectra were recorded at the NMR Laboratory of the Institute of Bioorganic Chemistry in Poznań using Bruker Avance spectrometers operating at 400 mHz Chemical shifts are reported in parts per million (ppm) and coupling constants in hertz (Hz). The chemical shifts were referenced to tetramethylsilane (TMS) for both ^{1}H and ^{13}C spectra, and to H_3PO_4 in D_2O for ^{31}P spectra.

All samples were prepared by dissolving approximately 20 mg of the compound in deuterated solvents, either DMSO-d_6, $CDCl_3$ or methanol-d_4.

High-resolution mass spectrometry (HRMS) analyses were conducted using electrospray ionization (ESI) and MALDI-TOF techniques at the Mass Spectrometry Laboratory of the Institute of Bioorganic Chemistry and the European Centre for Bioinformatics and Genomics in Poznań.

5.2.3. Synthesis of aromatic linkers based on pyromellitic and naphthalene derivatives

5.2.3.1. Synthesis of N,N'-bis(2-hydroxyethyl)-1,2,4,5-benzenetetracarboxylic diimide (1a)

A suspension of 1 g (4.58 mmol) of benzene-1,2,4,5-tetracarboxylic dianhydride in 5 ml dimethylformamide with the addition of 2 eq. of 2-aminoethanol (553 µl, 9.16 mmol) was heated under reflux at 140 °C for 4 h. After the reaction, the mixture was poured onto ice. Lowering the temperature led to the formation of a precipitate. The precipitate was filtered on a Buchner funnel and left in a desiccator to dry. The precipitate was recrystallized from dimethylformamide to obtain N,N'-bis(2-hydroxyethyl)-1,2,4,5-benzenetetracarboxylic diimide TLC (DCM/MeOH 9:1): Rf 0.67. (1.224g, 88% yield). ^{1}H NMR (400 MHz, DMSO, 25°C, TMS, ppm): δ = 8.16 (s; 2H), 4.35 (t, J=5,3Hz; 2H), 3.69 (t, J = 5.5 Hz; 4H), 3.61 (t, J=5.5Hz, 4H); ^{13}C NMR (101 MHz, DMSO) δ 166.36, 136.98, 116.99, 57.83, 40.90.

5.2.3.2. Synthesis of N,N'-bis(3-hydroxypropyl)-1,2,4,5-benzenetetracarboxylic diimide (1b)

A solution of 0.5 g (2.29 mmol) of benzene-1,2,4,5-tetracarboxylic dianhydride in dimethylformamide with the addition of 2 eq. of 3-aminopropanol (357 µl, 4.58mmol) was heated under reflux at 140 °C for 4 h. After the reaction, the mixture was poured onto ice. Lowering the temperature led to the formation of a precipitate. The precipitate was filtered on a Buchner funnel and left in a desiccator to dry. The precipitate was recrystallized from dimethylformamide to give N,N'-bis(3-hydroxypropyl)-1,2,4,5-benzenetetracarboxylic diimide TLC (DCM/MeOH 9:1): Rf 0.76 (660.9 mg, 87% yield). ^{1}H NMR (400 MHz, DMSO, 25°C, TMS, ppm): δ = 8.1 (s, 2H), 3.67 (t, J = 6.4 Hz, 4H), 3.45 (t, J = 6.2 Hz 4H), 1.76 (q, 4H).^{13}C NMR (101 MHz; DMSO) δ 166.47, 137.08, 117.13,; 58.73, 35.95, 30,92.

5.2.3.3. Synthesis of N,N'-bis(4-hydroxybutyl)-1,2,4,5-benzenetetracarboxylic diimide (1c)

A solution of 1 g (4.58 mmol) of benzene-1,2,4,5-tetracarboxylic dianhydride in dimethylformamide with the addition of 2 eq. of 4-aminobutanol (845 µl, 9.16 mmol) was heated under reflux at 140 ˚C for 4 h. After the reaction, the mixture was poured onto ice. Lowering the temperature led to the precipitation of a precipitate. The precipitate was filtered on a Buchner funnel and left in a desiccator to dry. The precipitate was recrystallized from dimethylformamide to obtain N,N'-bis(4-hydroxybutyl)-1,2,4,5-benzenetetracarboxylic diimide. TLC (DCM/MeOH 9:1): Rf 0.64 (1.5 g, 91% yield). ^{1}H NMR (400 MHz, DMSO, 25°C, TMS, ppm): δ = 8.15 (s, 2H, 4.11 (t, 2H), 3.62 (t, J = 7.0 Hz, 4H), 3.40 (t, J = 6.2 Hz, 4H), 1.65 (q, 4H), 1.43 (q, 4H). ^{13}C NMR (101 MHz, DMSO) δ 166.30, 136.89, 117.08, 60.13, 37.95, 29.70, 24.59.

5.2.3.4. Synthesis of N2,N6-bis(3-hydroxypropyl)naphthalene-2,6-dicarboxamide (1d)

A suspension of 300 mg (1.22 mmol) of dimethyl 2,6-naphthalene dicarboxylate in dimethylformamide with 2 eq of 3-aminopropanol (235 µl, 2.44 mmol) was heated to 160 ˚C under reflux for 4 h. After the reaction, the solution was evaporated under reduced pressure, and the precipitate was re-dissolved in dimethylformamide for recrystallization. N2,N6-bis(3-hydroxypropyl)naphthalene-2,6-dicarboxamide TLC (Hexane/AcOEt 8:2): Rf 0.044 (265.2 mg, 60%) was obtained. ^{1}H NMR (400 MHz; DMSO) δ 8.69 (d, J = 5.6 Hz, 2H), 8.22 (dd, J = 8.6, 2.2 Hz; 2H), 8.03 (dd, J = 8.6, 1.5 Hz, 2H), 4.36 (t, J = 6.3 Hz, 2H), 3.41 (t, J = 6.3 Hz, 4H), 3.18 – 3.08 (m, 4H), 1.55 (p, J = 6.6 Hz, 4H). ^{13}C NMR (101 MHz, DMSO) δ 167.22,; 134.28, 134.16, 130.17, 129.99, 126.16, 58.33, 34.37, 32.27.

5.2.3.5. Synthesis of N2,N6-bis(4-hydroxybutyl)naphthalene-2,6-dicarboxamide (1e)

A suspension of 300 mg (1.22 mmol) of dimethyl 2,6-naphthalene dicarboxylate in dimethylformamide with 2 eq of 4-aminobutanol (226 µl, 2.44 mmol) was heated to 160 ˚C under reflux for 4 h. After the reaction, the solution was evaporated under reduced pressure, and the precipitate was re-dissolved in dimethylformamide for recrystallization. N2,N6-bis(3-hydroxypropyl)naphthalene-2,6-dicarboxamide TLC (Hexane/AcOEt 8:2): Rf 0.04 (242.15 mg, 55%) was obtained. ^{1}H NMR (400 MHz, DMSO-d6, 25°C, TMS, ppm): δ = 8.70 (s, 2H, NH), 8.26 (d, J = 8.6 Hz, 2H, H-1, H-5), 8.20 (dd, J = 8.5, 4.3 Hz, 2H, H-4, H-8), 8.01 (dd, J = 37.7, 11.0 Hz, 2H, H-3, H-7), 3.93 (s, 2H, OH), 3.50−3.32 (m, 4H, CH_2OH), 3.06 (q, J = 11.6, 5.6 Hz, 4H, $NHCH_2$), 1.46−1.36 (m, 8H, $CH_2CH_2CH_2CH_2$). ^{13}C NMR (101 MHz, DMSO-d6, 25°C, TMS. ppm): δ = 165.9, 134.2, 130.2, 130.0, 129.1, 125.7, 60.3, 35.8, 29.8, 25.7.

5.2.3.6. Synthesis of 2,7-bis(3-hydroxypropyl)benzo[3,8]phenanthroline-1,3,6,8(2H,7H) (1f)

A suspension of 200 mg (0.746 mmol) of 1,4,5,8-naphthalenetetracarboxylic dianhydride in 4 ml anhydrous pyridine with the addition of 4 eq. of 3-aminopropanol (230 µl, 3 mmol) and 1,25eq. zinc acetate dihydrate (100 µl, 0.954 mmol) was heated under reflux at 120 °C for 16 h. After cooling to room temperature, the pyridine was removed under reduced pressure, and the crude product was dry-loaded onto a silica column and purified by elution with a dichloromethane:methanol solvent mixture. TLC (DCM/MeOH 9:1): Rf 0.69 (211,3 mg, 74% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.79 (s, 4H), 4.38 (t, J = 6.3 Hz, 4H), 3.65 (dd, J = 11.4, 5.8 Hz, 4H), 2.59 (t, J = 6.5 Hz, 2H), 2.07 – 1.92 (m, 4H).

5.2.3.7. Synthesis of 2,7-bis(4-hydroxybutyl)benzo[3,8]phenanthroline-1,3,6,8(2H,7H)-tetraone) (1g)

A suspension of 423 mg (1.6 mmol) of 1,4,5,8-naphthalenetetracarboxylic dianhydride in 7 ml anhydrous pyridine with the addition of 4 eq. of 4-aminobuthanol (650 µl, 6.32 mmol) and 1,25eq. zinc acetate dihydrate (210 µl, 2 mmol) was heated under reflux at 120 °C for 16 h. After cooling to room temperature, the pyridine was removed under reduced pressure, and the crude product was dry-loaded onto a silica column and purified by elution with a dichloromethane:methanol solvent mixture. TLC (DCM/MeOH 9:1): Rf 0.43 (530 mg, 82% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.76 (s, 4H), 4.26 (t, J = 7.4 Hz, 4H), 3.75 (dd, J = 11.0, 5.6 Hz, 4H), 1.95 – 1.76 (m, 4H), 1.71 (q, J = 6.6 Hz, 4H).

5.2.3.8. Synthesis of 2-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)-6-N,N'-bis(2-hydroxyethyl)-1,2,4,5-benzenetetracarboxylic diimide (2a)

A 200 mg of N,N'-bis(2-hydroxyethyl)-1,2,4,5-benzenetetracarboxylic diimide (1a) (0.66 mmol) was dissolved in pyridine. The mixture was stirred until the precipitate dissolved. To the stirring solution was added 1.2eq of dimethoxytrityl chloride (267 mg, 0.79 mmol) and a catalytic amount of 4-dimethylaminopyridine. The mixture was stirred for 2.5 h. After the reaction, the mixture was extracted from methylene chloride in an aqueous solution of sodium bicarbonate. The organic layer was dried over anhydrous magnesium sulfate and then evaporated under reduced pressure. After evaporation, the residue was dissolved in methylene chloride and subjected to isolation by column chromatography using silica gel in an eluent of hexane:ethyl acetate. Concentration of the eluate resulted in the precipitation of a yellow solid of 2-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)-6-N,N'-bis(2-hydroxyethyl)-1,2,4,5-benzenetetracarboxylic diimide TLC (DCM/MeOH 9:1): Rf 0.65 (290 mg, 76%). ^{1}H NMR (400 MHz, $CDCl_3$, 25°C, TMS, ppm): δ = 8.23 (s, 2H), 7.32 (d, J = 8.3 Hz, 9H), 7.2 (dd, J = 9.0, 2.6 Hz, 4H), 5.28 (s, 1H), 3.73 (s, 6H, OMe), 3.44 (dd, J = 6.4, 3.8 Hz, 4H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.56, 158.3, 144.45, 137.20, 136.97, 135.70, 129.84, 127.93, 127.74, 126.70, 113, 86.24, 60.24, 58.33, 55.09, 41.14, 38.85.

5.2.3.9. Synthesis of 2-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-6-N,N'-bis(3-hydroxypropyl)-1,2,4,5-benzenetetracarboxylic diimide (2b)

A 100 mg of N,N'-bis(3-hydroxypropyl)-1,2,4,5-benzenetetracarboxylic diimide (1b) (0.3 mmol) was dissolved in pyridine. The mixture was stirred until the precipitate dissolved. To the stirring solution was added 1.2 eq of dimethoxytrityl chloride (122.8 mg, 0.36 mmol) and a catalytic amount of 4-dimethylaminopyridine. The mixture was stirred for 2.5 h. After the reaction, the mixture was extracted from methylene chloride in an aqueous solution of sodium bicarbonate. The organic layer was dried over anhydrous magnesium sulfate and then evaporated under reduced pressure. After evaporation, the residue was dissolved in methylene chloride and subjected to isolation by column chromatography using silica gel in an eluent of hexane:ethyl acetate. Concentration of the eluate resulted in the precipitation of a yellow precipitate of 2-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-6-N,N'-bis(3-hydroxypropyl)-1,2,4,5-benzenetetracarboxylic diimide TLC (DCM/MeOH 9:1): Rf 0.71 (82.7 mg, 43%). ^{1}H NMR (400 MHz, $CDCl_3$, 25°C, TMS, ppm): δ = 8.18 (s, 2H, H-4, H-8), 7.37-7.08 (m; 5H; DMT), 6.82 (d, J = 8.9 Hz, 4H) 6.71 (d, J = 8.9 Hz, 4H, H_{arom}), 3.91 (t, J = 6,5 Hz; 1H; OH), 3.75 (s, 6H, OMe), 3.67 (t, J = 5.7 Hz, 4H, $OHCH_2CH2$), 3.15 (t, J = 5.6 Hz, 4H, CH_2ODMT), 2.06-1.99 (m, 2H, CH_2CH_2ODMT), 1.97-1.9 (m, 2H, $OHCH_2CH_2$). ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.7, 158.2, 144.5, 137.4, 136.9, 129.9, 129.1, 128.1, 127.7, 126.6, 118, 86.2, 61.6, 55.1, 36.9, 35.2, 31, 28.4.

5.2.3.10. Synthesis of 2-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)butyl)-6-N,N'-bis(4-hydroxybutyl)-1,2,4,5-benzenetetracarboxylic diimide (2c)

200 mg of N,N'-bis(4-hydroxybutyl)-1,2,4,5-benzenetetracarboxylic diimide **(1c)** (0.56 mmol) was dissolved in pyridine. The mixture was stirred until the precipitate dissolved. To the stirring solution was added 1.2 eq of dimethoxytrityl chloride (228 mg, 0.67 mmol) and a catalytic amount of 4-dimethylaminopyridine. The mixture was stirred for 2.5 h. After the reaction, the mixture was extracted from methylene chloride in an aqueous solution of sodium bicarbonate. The organic layer was dried over anhydrous magnesium sulfate and then evaporated under reduced pressure. After evaporation, the residue was dissolved in methylene chloride and subjected to isolation by column chromatography using silica gel in an eluent of hexane:ethyl acetate. Concentration of the eluate resulted in the precipitation of a yellow solid of 2-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)butyl)-6-N,N'-bis(4-hydroxybutyl)-1,2,4,5-benzenetetracarboxylic diimide TLC (DCM/MeOH 9:1): Rf 0.81 (136.5 mg, 36.7%). ^{1}H NMR (400 MHz, $CDCl_3$, 25°C, TMS, ppm): δ = 8.25 (s, 2H, H-4. H-8), 7.35–7.12 (m, 9H, DMT), 6.80 (d, J = 8.8 Hz, 4H, $H_{arom.}$), 3.96 (s, 1H, OH), 3.78 (s, 6H, OMe), 3.76−3.66 (m, 4H, NCH_2CH_2), 3.08 (t, J = 6.2 Hz, 4H, $HOCH_2CH_2$), 1.86–1.75 (m, 8H, CH_2CH_2). ^{13}C NMR (101 MHz, CDCl3) δ 166.2, 158.1, 148.8, 137, 136.3, 129.8, 127.9, 127.5, 123.9, 112.8, 85.6, 63.8, 54.9, 38.5, 29.4, 27.2, 25.2.

5.2.3.11. Synthesis of 2-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-N6-(3-hydroxypropyl)naphthalene-2,6-dicarboxamide (2d)

250 mg of N2,N6-bis(3-hydroxypropyl)naphthalene-2,6-dicarboxamide **(1d)** (0.76 mmol) was suspended in pyridine. The mixture was heated to 60°C and stirred until the precipitate dissolved. To the stirring solution was added 1.2 eq of dimethoxytrityl chloride (346.8 mg, 1.02 mmol) and a catalytic amount of 4-dimethylaminopyridine. The mixture was stirred for 2h. After the reaction, the mixture was extracted from methylene chloride in an aqueous solution of sodium bicarbonate. The organic layer was dried over anhydrous magnesium sulfate and then evaporated under reduced pressure. After evaporation, the residue was dissolved in methylene chloride and subjected to column chromatography using silica gel in an eluent of hexane:ethyl acetate. Concentration of the eluate resulted in the precipitation of yellow 2-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-N6-(3-hydroxypropyl)naphthalene-2,6-dicarboxamide TLC (DCM/MeOH 9:1): Rf 0.88 (224.3 mg, 47%). ^{1}H NMR (400 MHz, $CDCl_3$, 25°C, TMS, ppm): δ = 7.97 (s, 2H, NH), 7.38 (d, J = 8.0 Hz, 2H, H-4, H-$8_{arom.}$), 7.27 (d, J = 8.8 Hz, 2H H-3, H-7_{arom}), 7.23 (dd, J = 10.3 4.8 Hz, 9H, 1,8-$H_{arom.}$), 6.77 (d, J = 8.7 Hz, 4H, DMT_{arom}. H Ar–OMe), 4.05 (d, J = 7.1 Hz, 1H, OH), 3.70 (s, 6H, OMe), 3.33 (dd, J = 12.4, 6.3 Hz, 2H, CH_2OH), 3.13 (t, J = 5.8 Hz, 4H, NHC**H$_2$**), 1.78–1.72 (m, 4H, CH_2C**H$_2$**CH_2). ^{13}C NMR (101 MHz, $CDCl_3$, 25°C, TMS, ppm): δ = 162.1, 158.1, 144.7, 135.8, 129.6, 127.7, 112.7, 85.7, 61.0, 54.8, 36.0, 35.7, 30.9, 29.1.

5.2.3.12. Synthesis of 2-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)butyl)-N6-(4-hydroxybutyl)naphthalene-2,6-dicarboxamide (2e)

225 mg of N2,N6-bis(4-hydroxybutyl)naphthalene-2,6-dicarboxamide **(1e)** (0.63 mmol) was suspended in pyridine. The mixture was heated to 60°C and stirred until the precipitate dissolved. To the stirring solution was added 1eq of dimethoxytrityl chloride (415 mg, 1.22 mmol) and a catalytic amount of 4-dimethylaminopyridine. The mixture was stirred for 2 h. After the reaction, the mixture was extracted from methylene chloride in an aqueous solution of sodium bicarbonate. The organic layer was dried over anhydrous magnesium sulfate and then evaporated under reduced pressure. After evaporation, the residue was dissolved in methylene chloride and subjected to column chromatography using silica gel in an eluent of hexane:ethyl acetate. Concentration of the eluate resulted in the precipitation of yellow 2-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)butyl)-N6-(4-hydroxybutyl)naphthalene-2,6-dicarboxamide TLC (DCM/MeOH 9:1): Rf 0.83 (181 mg, 24%). ^{1}H NMR (400 MHz, $CDCl_3$, 25°C, TMS, ppm): δ = 7.99 (s, 2H, NH), 7.38 (d, J = 7.9 Hz, 2H, H-1, H-$5_{arom.}$), 7.27 (d, J = 8.8 Hz, 2H, H-4, H-$8_{arom.}$), 7.2 (t, J = 7.6 Hz. 1H, $DMT_{arom.}$), 7.12 (t, J = 7.3 Hz, 6H, DMT_{arom}), 6.76 (d, J = 8.9 Hz, 6H, DMT_{arom}) 4.04 (q, J = 7.1 Hz, 1H, OH), 3.68 (s, 6H, OMe), 3.17 (q, J = 6.2 Hz, 4H, $NHCH_2$), 3.03 (m, 4H, CH_2OH), 1.15 (dt, J = 18.4, 5.6 Hz, 8H, $CH_2CH_2CH_2CH_2$). ^{13}C NMR (101 MHz, $CDCl_3$, 25°C, TMS, ppm): δ = 162.1, 158.0, 144.9, 136.1, 127.7, 126.2, 112.6, 85.4, 62.5, 54.7, 37.6, 30.9, 27.1, 26.0.

5.2.3.13. Synthesis of 2-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-7-(3-hydroxypropyl)benzo[3,8]phenanthroline-1,3,6,8(2H,7H)-tetraone (2f)

200 mg (0.52 mmol) NDI diol 2,7-bis(3-hydroxypropyl)benzo[3,8]phenanthroline-1,3,6,8(2H,7H) (**1f**) was dissolved in 4 ml anhydrous pyridine, with the mixture heated to 60 °C to ensure complete dissolution. Subsequently, 1.25 equivalents of dimethoxytrityl chloride (221.5 mg, 0.654 mmol) were added, and the reaction was stirred for 2 hours. The solvent was then removed under reduced pressure, and the crude product was dissolved in dichloromethane, washed with saturated aqueous sodium bicarbonate, and dried over anhydrous magnesium sulfate. Purification was carried out by column chromatography using a solvent system of hexane:ethyl acetate supplemented with 1% triethylamine. TLC (DCM/MeOH 9:1): Rf 0.8 (174.8 mg, 49 % yield) ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.72 (d, J = 7.6 Hz, 4H), 7.32 (d, J = 7.3 Hz, 5H), 7.21 – 7.13 (m, 4H), 6.63 (d, J = 8.8 Hz, 4H), 4.36 (dt, J = 14.0, 6.6 Hz, 4H), 3.73 (s, 6H), 3.66 (t, J = 5.6 Hz, 4H), 3.22 (t, J = 5.6 Hz, 2H), 2.10 (t, J = 6.3 Hz, 2H)

5.2.3.14. Synthesis of 2-(4-(bis(4-methoxyphenyl)(phenyl)methoxy)butyl)-7-(4-hydroxybutyl)benzo[3,8]phenanthroline-1,3,6,8(2H,7H)-tetraone (2g)

90 mg (0.22 mmol) NDI diol 2,7-bis(4-hydroxybutyl)benzo[3,8]phenanthroline-1,3,6,8(2H,7H)-tetraone (**1g**) was dissolved in 4 ml anhydrous pyridine, with the mixture heated to 60 °C to ensure complete dissolution. Subsequently, 1.25 equivalents of dimethoxytrityl chloride (92.75 mg, 0.27 mmol) were added, and the reaction was stirred for 2 hours. The solvent was then removed under reduced pressure, and the crude product was dissolved in dichloromethane, washed with saturated aqueous sodium bicarbonate, and dried over anhydrous magnesium sulfate. Purification was carried out by column chromatography using a solvent system of hexane:ethyl acetate supplemented with 1% triethylamine. TLC (DCM/MeOH 9:1): Rf 0.6 (58.3 mg, 37.3 % yield) ^{1}H NMR (400 MHz, $CDCl_3$): 8.16 (d, *J* = 6.7 Hz, 4H), 7.43 (4H, 7.31 (tt, *J* = 7.4, 1.4 Hz), 7.35 (dddd, *J* = 8.0, 7.4, 1.6, 0.6 Hz, 5H), 6.91 (ddd, *J* = 8.4, 1.1, 0.5 Hz, 2H), 7.01 (ddd, *J* = 8.4, 1.3, 0.5 Hz, 2H), 6.84-7.07 (m, 4H), 3.86 (t, *J* = 6.8 Hz, 4H), 3.86 (t, *J* = 6.8 Hz, 4H), 3.76 (s, 6H), 3.48 (t, *J* = 7.5 Hz, 2H), 3.34 (2H, t, *J* = 7.0 Hz), 1.80 (tt, *J* = 7.1, 7.0 Hz, 2H), 1.78 (tt, *J* = 7.5, 7.5 Hz, 2H).

5.2.3.15. Synthesis of 2-(6-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)-1,3,5,7-tetraoxo-3,5,6,7-tetrahydropyrrolo[3,4-f]isoindol-2(1H)-yl)ethyl(2-cyanoethyl)diisopropylphosphoramidite (3a)

133 mg of 2-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)-6-N,N'-bis(2-hydroxyethyl)-1,2,4,5-benzenetetracarboxylic acid diimide **(2a)** (0.2 mmol) dried overnight under a vacuum pump was dissolved in methylene chloride in an inert gas environment (argon) and cooled to 0°C in an ice bath. To the mixture were added 3.5 eq of diisopropylethylamine (140 µl, 0.8mmol) and 1.2 eq of N,N-diisopropylaminocyanoethylphosphonamidochloride (61 µl, 0.27 mmol). The reaction was carried out for 2.5 h. After the reaction, the mixture was extracted from methylene chloride in an aqueous solution of sodium

hydrogen carbonate. The organic layer was dried over anhydrous magnesium sulfate and then evaporated under reduced pressure. Then the solution was subjected to isolation on a chromatographic column using silica gel in methylene chloride. Concentration of the eluate resulted in a colorless oil of 2-(6-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)-1,3,5,7-tetraoxo-3,5,6,7-tetrahydropyrrolo[3,4-f]isoindol-2(1H)-yl)ethyl(2-cyanoethyl)diisopropylphosphoramidite TLC (Hexane/AcOEt 8:2): Rf 0.84 (106.14 mg, 60%), which was dried under a vacuum pump. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.19 (s, 2H), 7.32 (d, J = 7.6 Hz, 9H), 7.20 (d, J = 8.8 Hz, 2H), 6.70 (dd, J = 8.9, 2.5 Hz, 2H), 3.90 (d, J = 4.7 Hz, 1H), 3.71 (s, 6H), 3.55 – 3.41 (m, 2H), 2.95 – 2.89 (m, 2H), 2.76 (t, J = 5.6 Hz, 1H), 1.35 (d, J = 6.4 Hz, 12H).

5.2.3.16. Synthesis of 2-(6-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-1,3,5,7-tetraoxo-3,5,6,7-tetrahydropyrrolo[3,4-f]isoindol-2(1H)-yl)propyl(2-cyanoethyl)diisopropylphosphoramidite (3b)

104 mg of 2-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-6-N,N'-bis(3-hydroxypropyl)-1,2,4,5-benzenetetracarboxylic acid diimide **(2b)** (0.16 mmol) dried overnight under a vacuum pump was dissolved in methylene chloride in an inert gas environment (argon) and cooled to 0°C in an ice bath. To the mixture were added 3.5 eq of diisopropylethylamine (100 µl, 0.57 mmol) and 1.2 eq of N,N-diisopropylaminocyanoethylphosphonamidochloride (45 µl, 0.2 mmol). The reaction was carried out for 2.5 h. After the reaction, the mixture was extracted from methylene chloride in an aqueous solution of sodium hydrogen carbonate. The organic layer was dried over anhydrous magnesium sulfate and then evaporated under reduced pressure. Then the solution was subjected to isolation on a chromatographic column using silica gel in methylene chloride. Concentration of the eluate resulted in colorless 2-(6-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-1,3,5,7-tetraoxo-3,5,6,7-tetrahydropyrrolo[3,4-f]isoindol-2(1H)-yl)propyl(2-cyanoethyl)diisopropylphosphoramidite TLC (Hexane/AcOEt 8:2): Rf 0.93 (80.2 mg, 58%), which was dried under a vacuum pump. ^{1}H NMR (400 MHz): δ 8.76 (2H, d, J = 0.5 Hz), 7.35 (dddd, J = 8.0, 7.4, 1.6, 0.6 Hz, 9H), 7.31 (tt, J = 7.4, 1.4 Hz, 2H), 6.91 (ddd, J = 8.4, 1.1, 0.5 Hz, 2H), 3.87 (t, J = 6.9 Hz, 2H), 3.76 (s, 6H), 3.67-3.94 (m, 2H), 2.98-3.18 (m, 2H), 2.01 (tt, J = 6.9, 5.0 Hz), 1.18 (12H, d, J = 6.9 Hz).

5.2.3.17. Synthesis of 2-(6-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)butyl)-1,3,5,7-tetraoxo-3,5,6,7-tetrahydropyrrolo[3,4-f]isoindol-2(1H)-yl)butyl(2-cyanoethyl)diisopropylphosphoramidite (3c)

113 mg of 2-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)butyl)-6-N,N'-bis(4-hydroxybutyl)-1,2,4,5-benzenetetracarboxylic acid diimide **(2c)** (0.17 mmol) dried overnight under a vacuum pump was dissolved in methylene chloride in an inert gas environment (argon) and cooled to 0°C in an ice bath. To the mixture were added 3.5 eq of diisopropylethylamine (104 µl, 0.6 mmol) and 1.2 eq of N,N-

diisopropylaminocyanoethylphosphonamidochloride (46 µl, 0.2 mmol). The reaction was carried out for 2.5 h. After the reaction, the mixture was extracted with methylene chloride in an aqueous solution of sodium hydrogen carbonate. The organic layer was dried over anhydrous magnesium sulfate and then evaporated under reduced pressure. Then the solution was subjected to isolation on a chromatographic column using silica gel in methylene chloride. Concentration of the eluate resulted in colorless 2-(6-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)butyl)-1,3,5,7-tetraoxo-3,5,6,7-tetrahydropyrrolo[3,4-f]isoindol-2(1H)-yl)butyl(2-cyanoethyl)diisopropylphosphoramidite TLC (Hexane/AcOEt 8:2): Rf 0.72 (88.2 mg, 60%), which was dried under a vacuum pump. ^{1}HNMR (400 MHz): δ 8.76 (d, J = 0.5 Hz, 2H), 7.35 (dddd, J = 8.0, 7.4, 1.6, 0.6 Hz, 9H), 7.24-7.43 (m, 2H), 6.93 (ddd, J = 8.4, 1.1, 0.5 Hz, 2H), 3.85 (t, J = 6.8 Hz, 2H), 3.84 (t, J = 6.8 Hz, 2H), 3.85 (t, J = 6.9 Hz, 2H), 3.76 (s, 6H), 3.64-3.92 (m, 2H), 2.98-3.18 (m, 2H), 1.18 (, d, J = 6.9 Hz, 12H)

5.2.3.18. Synthesis of 3-(6-((3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)carbamoyl)2-naphthamido)propyl(2-cyanoethyl)diisopropylphosphoramidite (3d)

175.7 mg of 2-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-N6-(3-hydroxypropyl)naphthalene-2,6-dicarboxamide **(2d)**(0.28 mmol) dried overnight under a vacuum pump was dissolved in methylene chloride in an inert gas environment (argon) and cooled to 0^{o}C in an ice bath. To the mixture was added 5eq of diisopropylethylamine (170 µl, 0.98 mmol) and 2 eq of N,N-diisopropylaminocyanoethylphosphonamido chloride (74 µl, 0.33 mmol). The reaction was carried out for 2.5 h. After the reaction, the mixture was extracted with methylene chloride in an aqueous sodium hydrogen carbonate solution. The organic layer was dried over anhydrous magnesium sulfate and then evaporated under reduced pressure. Then, the solution was subjected to column chromatography using silica gel in methylene chloride. Concentration of the eluate resulted in colorless 3-(6-((3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)carbamoyl)2-naphthamido)propyl(2-cyanoethyl)diisopropylphosphoramidite TLC (Hexane/AcOEt 8:2): Rf 0.79 (138.8 mg, 60%), which was dried under a vacuum pump. ^{1}H NMR (400 MHz, $CDCl_3$, 25°C, TMS, ppm): δ = 8.09 (s, 1H), 7.35 (d, J = 7.5 Hz, 3H), 7.24 (d, J = 8.7 Hz, 9H), 6.77 (d, J = 9.0 Hz, 6H), 3.73 (s, 6H), 3.34 (q, J = 6.2 Hz, 2H), 3.15 (t, J = 5.7 Hz, 2H), 2.54 (t, J = 6.3 Hz, 1H), 1.78−1.71 (m, 2H), 1.20 (dd, J = 12.5, 6.1 Hz, 6H), 1.09 (d, J = 6.8 Hz, 3H), 0.98 (t, J = 7.1 Hz, 4H). ^{31}P NMR (162 MHz, $CDCl_3$, 25°C, TMS, ppm): δ = 150,81 (s).

5.2.3.19. Synthesis of 4-(6-((4-(bis(4-methoxyphenyl)(phenyl)methoxy)butyl)carbamoyl)-2-naphthamido)butyl(2-cyanoethyl)diisopropylamidophosphite (3e)

175.8 mg of 2-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-N6-(3-hydroxypropyl)naphthalene-2,6-dicarboxamide **(2e)** (0.27 mmol) dried overnight under a vacuum pump was dissolved in methylene chloride in

an inert gas environment (argon) and cooled to 0°C in an ice bath. 5 eq of diisopropylethylamine (160 µl, 0.92 mmol) and 2 eq of N,N-diisopropylaminocyanoethylphosphonamide chloride (70µl, 0.31mmol) were added to the mixture. The reaction was carried out for 2.5 h. After the reaction, the mixture was extracted with methylene chloride in an aqueous sodium hydrogen carbonate solution. The organic layer was dried over anhydrous magnesium sulfate and then evaporated under reduced pressure. Then, the solution was subjected to column chromatography using silica gel in methylene chloride. Concentration of the eluate resulted in colorless 3-(6-((3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)carbamoyl)2-naphthamido)propyl(2-cyanoethyl)diisopropylphosphoramidite TLC (Hexane/AcOEt 8:2): Rf 0.76 (150.4 mg, 65%), which was dried under a vacuum pump. ^{1}HNMR (400 MHz): δ 8.03 (dddd, J = 7.4, 1.7, 0.5, 0.4 Hz, 2H), 7.35 (dddd, J = 8.0, 7.4, 1.6, 0.6 Hz, 3H), 7.24-7.43 (m, 9H), 6.93 (ddd, J = 8.4, 1.1, 0.5 Hz, 2H) 6.87-7.07 (m, 2H), 3.84 (t, J = 6.8 Hz, 2H), 3.76 (s, 6H), 3.51 (t, J = 7.4 Hz, 2H), 3.16 (t, J = 7.5 Hz, 2H), 3.08 (p, J = 6.9 Hz, 2H), 1.75-1.93 (m, 2H), 1.63 (tt, J = 7.5, 7.4 Hz, 3H), 1.18 (d, J = 6.9 Hz, 12H)

5.2.3.20. Synthesis of 3-(7-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-1,3,6,8-tetraoxo-3,6,7,8-tetrahydrobenzo[3,8]phenanthrolin-2(1H)-yl)propyl (2-cyanoethyl) diisopropylphosphoramidite (3f)

70 mg (0.07 mmol) NDI-DMT derivative 2-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-7-(3-hydroxypropyl)benzo[3,8]phenanthroline-1,3,6,8(2H,7H)-tetraone (2f) was dissolved in anhydrous dichloromethane and placed under an argon atmosphere. The reaction mixture was cooled in an ice bath, followed by the addition of 5 equivalents of diisopropylethylamine (90 µl, 0.36 mmol) and 2 equivalents of 2-cyanoethyl N,N-diisopropylchlorophosphoramidite (32 µl, 0.14 mmol). The reaction was allowed to gradually warm to room temperature over the course of 2 hours, after which it was quenched by pouring into a solution of 2% triethylamine in dichloromethane. The resulting mixture was washed with saturated aqueous sodium bicarbonate and dried over anhydrous magnesium sulfate. Purification was carried out by column chromatography using 2% triethylamine in ethyl acetate as the eluent. TLC (DCM/MeOH 9:1): Rf 0.9 (51.1 mg, 56 % yield) ^{1}H NMR (400 MHz): 8.16 (d, *J* = 6.8 Hz), 7.35 (dddd, *J* = 8.0, 7.4, 1.6, 0.6 Hz), 7.31 (tt, *J* = 7.4, 1.4 Hz), 7.24-7.43 (m, 3H), 6.84-7.07 (m, 6H), 3.88 (t, *J* = 2.7 Hz), 3.8 (t, *J* = 7.0 Hz), 3.76 (s, 6H), 3.73 (t, *J* = 7.0 Hz), 2.98-3.18 (m, 4H), 2.00 (tt, *J* = 2.7, 2.7 Hz), 1.93-2.10 (m, 4H), 1.18 (12H, d, *J* = 6.9 Hz).

5.2.3.21. Synthesis of 4-(7-(4-(bis(4-methoxyphenyl)(phenyl)methoxy)butyl)-1,3,6,8-tetraoxo-3,6,7,8-tetrahydrobenzo[3,8]phenanthrolin-2(1H)-yl)butyl (2-cyanoethyl) diisopropylphosphoramidite (3g)

50 mg (0.1 mmol) NDI-DMT derivative 2-(4-(bis(4-methoxyphenyl)(phenyl)methoxy)butyl)-7-(4-hydroxybutyl)benzo[3,8]phenanthroline-1,3,6,8(2H,7H)-tetraone (2g) was dissolved in anhydrous dichloromethane and placed under an argon atmosphere. The reaction mixture was cooled in an ice bath, followed

by the addition of 5 equivalents of diisopropylethylamine (62 µl, 0.52 mmol) and 2 equivalents of 2-cyanoethyl N,N-diisopropylchlorophosphoramidite (50 µl, 0.22 mmol). The reaction was allowed to gradually warm to room temperature over the course of 2 hours, after which it was quenched by pouring into a solution of 2% triethylamine in dichloromethane. The resulting mixture was washed with saturated aqueous sodium bicarbonate and dried over anhydrous magnesium sulfate. Purification was carried out by column chromatography using 2% triethylamine in ethyl acetate as the eluent. TLC (DCM/MeOH 9:1): Rf 0.93 (38.3 mg, 60 % yield). ^{1}HNMR (400 MHz): δ 8.22 (d, J = 6.8 Hz, 4H), 7.22-7.4 (m, 4H), 7.01 (ddd, J = 8.4, 1.3, 0.5 Hz), 6.91 (ddd, J = 8.4, 1.1, 0.5 Hz, 6H), 3.86 (t, J = 6.8 Hz), 3.84 (t, J = 6.8 Hz), 3.73 (s, 6H), 3.62-3.9 (m, 4H), 3.48 (t, J = 7.5 Hz, 2H), 3.08 (t, J = 6.8 Hz, 2H), 1.85 (tt, J = 7.3, 7.2 Hz), 1.54-1.94 (m, 12H).

5.2.4. Oligonucleotide synthesis and purifying

All oligonucleotides were synthesized using standard synthesis protocol. RNA oligomers were synthesized on an Applied Biosystems 392 DNA/RNA synthesizer using the solid-phase phosphoramidite method with controlled pore glass (CPG) support. All syntheses were performed on a 1-µmol scale using commercially available 2'-O-TOM-protected phosphoramidites. Base-labile protecting groups were removed, and the oligomers were cleaved from the solid support by incubating the CPG resin with an ammonia/methylamine solution (1:1, v/v) for up to 4 h. DMT-ON oligomers were subsequently detritylated and purified using Glen-Pak DNA or RNA cartridges (Glen Research) according to the manufacturer's instructions. For RNA oligomers synthesized in the DMT-OFF format, the 2'-O-TOM protecting group was removed using 1 M triethylamine trihydrofluoride (TEA·3HF) in DMSO at 65°C for 2.5 h. All oligomers were desalted using Illustra NAP-25 columns (GE Healthcare). The purity of the obtained RNA was assessed using 10–20% denaturing polyacrylamide gel electrophoresis. Synthesis and purification process were described in publication: *Overview of Methods for Large-Scale RNA Synthesis* (187).

5.2.5. UV melting measurement

Ultraviolet (UV) thermal melting studies were performed on a Jasco V-700 spectrometer with a thermoprogrammer. The RNA oligomers were dissolved in a buffer containing 100 mM sodium chloride and 40 mM sodium cacodylate (pH 7.0). Oligonucleotides were prepared in nine different concentrations. The UV absorption versus temperature was measured at 260 nm at a heating rate of 0.5 °C / min in the range 5–90 °C (188). The melting curves were

analyzed using Microsoft Excel software to calculate the first derivative of the UV-melting curves and melting temperature.

UV melting procedure for RNA oligonucleotides with NCD ligand was described in publication: *Naphthyridine carbamate dimer ligand induces formation of Z-RNA-like fold of disease-related RNA and exhibits a molecular glue characteristics in crystal lattice formation* (159).

5.2.6. Circular dichroism

Circular dichroism (CD) spectra were recorded using a JASCO J-815 spectropolarimeter (Cremella (LC), Italy). Each oligonucleotide sample was evaporated to dryness and subsequently dissolved in 1 ml of buffer containing 100 mM NaCl and 40 mM sodium cacodylate, yielding a final concentration of 6.0 µM. Spectra were acquired five times over the 210–410 nm wavelength range at 20 °C, and the buffer baseline was subtracted from each sample spectrum.

5.2.7. RNA concentration measurement

The concentration of the prepared oligonucleotides was determined by measuring absorbance at 260 nm using a Implen Nanophotometer®. The absorbance values were subsequently converted to concentrations using the Lambert-Beer law, applying extinction coefficients specific to each oligoribonucleotide. These coefficients were calculated using the RNA Molecular Weight Calculator (https://www.aatbio.com/tools/calculate-RNA-molecular-weight-mw). Any introduced chemical modifications were excluded from the calculations.

5.2.8. RNA purification on polyacrylamide gel

Preparative-scale electrophoresis was performed on 15–20% polyacrylamide gels, with the gel concentration selected based on the length of the RNA. Each run was conducted in 1× TBE buffer and preceded by a pre-electrophoresis step at a constant current of 10 mA. Prior to loading, gel was prepared by increasing constant current to 50 mA for better sample penetration into the gel. Samples were mixed in a 1:1 ratio with an 8 M urea solution to ensure denaturation. Electrophoresis was carried out under gradually increasing current until a maximum of 50 mA was reached.

5.2.9. RNA elution

Following separation on a polyacrylamide gel, bands corresponding to the desired RNA fragments were visualized under UV light, excised from the gel, and eluted using two different methods: electroelution and the "crush and soak" technique. For electroelution, the gel slice containing the RNA was placed into a ThermoFisher SnakeSkin dialysis bag, which was then filled with 1× TBE buffer. The bag was submerged in the same buffer, and a voltage of 120 V was applied. Electroelution was carried out for 1 hour. Afterward, the eluate containing the RNA was transferred from the dialysis bag to an Eppendorf tube, and the RNA was precipitated using isopropanol. In the second method, RNA was recovered by incubating the gel slice in deionized water overnight at 4 °C with gentle shaking.

5.2.10. Nucleic acid precipitation

Oligomers were precipitated in the presence of 1 eq. volume of 96% ethanol, 100 mM sodium acetate buffer (pH 5.3) and 10 µl of glycogen. Samples were precipitated for 1 hour at −20 °C and then centrifuged for 30 minutes at 14,000 rpm at 4 °C. After removing the supernatant, the pellet was dried and resuspended in deionized water. Nucleic acid concentrations were determined by measuring absorbance at λ = 260 nm using a NanoDrop spectrophotometer.

5.2.11. Nucleic acids desalting

Salt-contaminated or low-molecular-weight impurity-contaminated nucleic acid samples were purified using commercially available NAP-25 columns packed with Sephadex G-25 resin. A 0.8–1 mL aliquot of the sample was applied to the column, which was then washed with deionized water. The first two fractions were collected in 1 ml volumes, followed by subsequent fractions collected in 500 µl volumes, totaling 6 ml. Fractions containing nucleic acids were identified by measuring absorbance at 260 nm.

5.2.12. *In Vitro* Transcription

All RNA molecules analyzed in this study were synthesized using the EurX in vitro transcription kit and T7 RNA polymerase produced in-house. Transcription reactions were performed using 1 µg of chemically synthesized DNA as the template. Each reaction mixture contained 30 µg of T7 RNA polymerase, 3.75 µM of each NTP, 9mM GTP and transcription buffer. Prior to transcription, DNA templates were denatured at 95 °C for 5 minutes, followed by renaturation at room temperature for 10 minutes. The transcription reactions were incubated overnight in

37 °C. Subsequently, residual DNA templates were digested with 10 units of DNase I at 37 °C for 15 minutes. To dissolve precipitated inorganic pyrophosphate, 50 mM EDTA was added to the reaction mixture. The resulting RNA products were purified using 8–12% denaturing polyacrylamide gel electrophoresis, eluted by electroelution, and precipitated with isopropanol.

5.2.13. Dephosphorylation

FastAP™ alkaline phosphatase was used to remove phosphate groups from the 5’ end of the RNA oligomer. The reaction was performed in the manufacturer-supplied buffer, using 160 pmol of RNA in a total volume of 50 µl, and incubated at 37 °C for 10 minutes. Enzyme inactivation was achieved by heating the reaction mixture at 75 °C for 10 minutes. Subsequently, RNA oligomers were desalted using Zeba™ Spin Desalting Columns following the manufacturer’s protocol.

5.2.14. Phosphorylation 5’ end of RNA and labeling of RNA at the 5' end with the radioactive isotope ^{32}P

RNA at concentration of 10 pmol/µl were labeled using T4 polynucleotide kinase. Prior to the reaction, the RNA was denatured for 5 minutes at 90 °C and subsequently incubated on ice for 10 minutes. The reaction mixture contained the supplied enzyme buffer, 1 µl of [γ-^{32}P]ATP with a specific activity of 4000–5000 Ci/mmol (0.02 mCi) or 0.015 mM ATP, and 10 U of the enzyme. The labeling reaction was carried out for 30 minutes at 37 °C.

5.2.15. RNA circularization

The reaction mixture contained RNA at a concentration of 10 pmol/µl, and was incubated in 50 mM KCl for 5 min at 95°C, 10 min at 0°C, and 10 min at room temperature. The prepared RNA was then circularized using 20u of T4 RNA ligase in the presence of 1 mM ATP in 1x ligation buffer for 4 hours. Subsequently, RNA oligomers were desalted using Zeba™ Spin Desalting Columns following the manufacturer’s protocol. Circularization efficiency was evaluated on a denaturing polyacrylamide gel using a control sample of unligated RNA for comparison. The product of ligation was purified using 12-20% polyacrylamide gel. For radiolabeled RNA, the circularization products were separated in a polyacrylamide gel with 8M urea and then visualized by autoradiography using a Fluorescent Image Analyzer FLA-5100 radioactivity scanner - FujiFilm.

5.2.16. Formamide hydrolysate

^{32}P-labeled RNA was mixed with formamide at a 1:5 ratio in the presence of 10 mM $MgCl_2$. The mixture was incubated at 100 °C for 10 minutes to denature the RNA. Following incubation, an equal volume of formamide loading buffer containing electrophoretic tracking dyes was added, and the sample was immediately transferred to dry ice. The reaction products were subsequently separated on a 15% polyacrylamide gel in 1× TBE buffer. Radiolabeled RNA fragments were visualized using either a FLA-5100 scanner (Fujifilm) or an Amersham Typhoon 5 Biomolecular Imager, and image analysis was performed with ImageQuant TL v10.1 software (Cytiva).

5.2.17. RNA cleavage by Dicer

Hydrolysis reactions of the radiolabeled RNA substrate were performed in Dicer RNA cleavage reaction buffer. Dicer prepared in human HEK293T cell lines (final concentration: 15 nM), was used to catalyze the reaction. Incubation was carried out at 37 °C for 30 minutes. To terminate the reaction, an equal volume of EDTA and urea solution was added. The reaction products were subsequently separated on a 15% polyacrylamide gel in 1× TBE buffer. Radiolabeled RNA fragments were visualized using either a FLA-5100 scanner (Fujifilm) or an Amersham Typhoon 5 Biomolecular Imager, and image analysis was performed with ImageQuant TL v10.1 software (Cytiva).

5.2.18. Crystallization

Lyophilized RNA oligomers were dissolved in deionized water, followed by the addition of sodium cacodylate to a final concentration of 10 mM. Final concentration of RNA was set to 1mM. The RNA solution was denatured by heating at 95 °C for 2–5 minutes, then rapidly cooled on ice for 10 minutes and subsequently incubated at room temperature for an additional 10 minutes to facilitate proper folding.

Crystals were obtained using the vapor diffusion method, which relies on the principle that a closed system will naturally progress toward equilibrium. In this approach, a small drop containing the RNA sample is suspended above a reservoir solution that contains a higher concentration of a precipitating agent. The drop is composed of a mixture of RNA solution and precipitant at a defined ratio, resulting in an initial precipitant concentration that is lower than that of the reservoir. Over time, the volatile components—typically water and buffer—diffuse

between the drop and the reservoir. This gradual equilibration increases the concentration of the precipitating agent within the drop, leading the RNA to reach a supersaturated state, which promotes nucleation and subsequent crystal growth or, in some cases, precipitation (189–191).

Crystallization conditions were screened using commercially available kits, including Natrix 1, Natrix 2, and Crystal Screen (Hampton Research), at incubation temperatures of 19 °C and 35 °C. RNA samples were mixed with crystallization solutions at ratios of 1:1, 1:2, and 2:1 (v/v). Crystallization was carried out either manually or using the Gryphon crystallization robot (Art Robbins Instruments) using sitting drop method.

5.2.19. Diffraction measurements

Synchrotron radiation refers to the high-intensity, continuous-spectrum X-ray light produced by electron storage rings and is widely utilized in biocrystallography. Compared to conventional X-ray sources, it offers significantly brighter and more tightly collimated beams, leading to enhanced signal-to-noise ratios and reduced radiation damage. Its tunable wavelength capability allows for precise energy selection and more effective absorption corrections, making it particularly well-suited for investigating large unit cells, such as those found in viral structures. Moreover, synchrotron radiation facilitates time-resolved data collection, enabling the study of dynamic processes in macromolecular reactions.

Diffraction data were collected using synchrotron radiation in EMBL on beamline P13 PETRA III storage ring (DESY, Hamburg, Germany) (192, 193). Prior to collecting data, crystals were frozen in liquid nitrogen and transported in dryshipper maintaining temperature under 100K. Collecting data was also performed in temperature of 100K. The lower temperature limits radiation damage that can occur to the crystal during measurement.

Radiation damage in crystals is mainly caused by direct interactions between the beam and molecules, leading to heat generation and bond breakage. This primary damage depends on the radiation dose. Two main mechanisms create reactive radicals: directly (e.g., damaging the polypeptide) and indirectly (*via* H* or OH* from water breakdown). These reactive species can diffuse through the crystal, causing secondary damage, which is influenced by time and temperature. At cryo-temperatures (~100 K), diffusion is minimal, limiting damage to the

exposed area. At room temperature, radicals spread, leading to widespread destruction. Therefore, cooling samples during data collection greatly reduces radiation damage (194).

Data collection in X-ray diffraction experiments involves capturing the scattering of X-rays by crystalline samples to reveal their atomic structures. The primary goal is to measure a large number of reflection intensities (195). During data collection, several hundred diffraction images are typically captured as the crystal rotates around an axis perpendicular to the X-ray beam. Each image corresponds to a small rotation (0.2–2.0 °) and involves exposing the crystal for a specified duration, which can range from a fraction of a second to several tens of seconds, depending on the beam intensity and diffraction quality of the crystal. To acquire a complete dataset, reflections must be measured over an appropriate angular range determined by the symmetry of the space group and the orientation of the crystal. To minimize measurement errors, it is advisable to record each reflection multiple times by measuring the intensities of the symmetry-related reflections.

5.2.20. Data processing and scaling

The diffraction images are processed into structure factors using programs such as HKL (196) or XDS (197). In this work XDS program was used. During this stage, the space group and unit cell parameters were determined, and reflections with the highest intensities (I) were identified. The integrated intensities of these reflections were then scaled and merged to produce a complete set of structure factors essential for calculating the electron density maps (198). In programs such as XDS, indexing and integration typically rely on default settings, whereas scaling and merging are performed with modules such as XSCALE and XDSCONV, using standard parameters for non-anomalous diffraction data (199). The XDS output log file, particularly the XDS.INP tab, provides detailed information on the processing steps, including data quality metrics from the COLSPOT, IDXREF, DEFPIX, INTEGRATE, and CORRECT stages. These outputs enable researchers to mask untrusted detector regions, refine the integration parameters, and correct for absorption effects (200).

XSCALE standardizes the datasets obtained from XDS processing by placing them on a common scale and, when required, merging them into one or several sets of unique reflections. The program offers the option of merging these datasets into one or several sets of unique reflections, while also providing key quality indicators such as data completeness

and the reliability of the integrated intensities. XSCALE also applies corrections for absorption effects, variations in detector sensitivity across the detector plane, and radiation damage. For cases in which radiation damage is significant, reflections can be individually corrected by extrapolation to their initial intensities at zero dose. In contrast, XDSCONV is a tool for converting reflection data files generated by XDS or XSCALE into the specific formats required by various software packages used in crystal structure determination. Beyond mere format conversion, XDSCONV can generate test reflections or import previously defined ones, which are then used to calculate the free R factor, a crucial metric for assessing the accuracy and progress of the structure refinement (197).

Finally, a statistical analysis is conducted to evaluate the quality of the obtained reflection set based on parameters such as completeness, I/sigma, and $Rmerge$. Completeness refers to the ratio of recorded reflections to the theoretical or expected number of reflections at a specific resolution, which should be at least 85%. I/σ represents the average signal-to-noise ratio and should exceed 2, indicating that the reflection intensity is twice as large as the measurement error. $Rmerge$ provides information about the intensity difference between symmetrical reflections, which should be identical; thus, the value of this parameter should be as low as possible. This parameter is derived from the ratio of the total number of reflections to the number of independent reflections and is known as the multiplicity or redundancy (201). $Rmerge$ serves as a quality indicator of the diffraction data, as given by the following equation:

$$R_{merge}(I) = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$$

5.2.21. Definition of the phase problem and methods for its solution

When electromagnetic radiation with a wavelength comparable to the interatomic distances interacts with a crystalline sample, each atom behaves as a scattering center. The constructive and destructive interference of scattered waves gives rise to diffraction. A crystal can be described in terms of lattice planes indexed by the Miller indices (hkl). The condition for constructive interference is given by Bragg's law:

$$\lambda = 2d_{hkl} sin\theta$$

where λ is the radiation wavelength, d_{hkl} the interplanar spacing, and 2θ the diffraction angle relative to the incident beam. For any crystal system, d_{hkl} can be expressed in terms of the unit-cell parameters $a, b, c, \alpha, \beta, \gamma$ through trigonometric relations (202).

In a diffraction experiment, the intensity of the observed spots varies according to the electron density distribution of the unit cell. These intensities are connected to the underlying electron density through a mathematical operation known as Fourier transform (Bragg, 1915) (203). Each diffraction spot corresponds to scattering from planes in the crystal, indexed by Miller indices (hkl), with the amplitude of the scattered wave $|F_{hkl}|$ being proportional to the square root of the measured intensity. The electron density at any point (x, y, z) within the unit cell can then be reconstructed by summing the contributions of all such scattered waves, each weighted by its amplitude and combined with the correct phase relationship.

$$I(hkl) = |F_{hkl}|^2$$

$$\rho(xyz) = \frac{1}{V}\sum_{hkl} |F_{hkl}| e^{-2\pi i(hx+hy+hz)-\varphi(hkl)}$$

where V is the volume of the unit cell and hkl is the phase associated with the structure-factor amplitude $|F_{hkl}|$. X-ray diffraction measurements of crystals cannot experimentally determine the phase angle ϕ(hkl), a limitation referred to as the **phase problem**.

Molecular replacement (MR) is the most widely used and straightforward method for overcoming the crystallographic phase problem (204–206). In MR, the initial phases are derived from a known homologous structure that serves as a phase model. The success of this approach depends on the structural similarity between the model and target molecule, and ~30% sequence identity is typically sufficient (207). The orientation and translation of the model in the new unit cell are determined using Patterson-based methods, which employ a map calculated from squared structure factor magnitudes with all phases set to zero. Peaks in the Patterson map correspond to interatomic vectors rather than atomic positions, allowing the identification of rotational and translational relationships. The molecular replacement procedure thus involves solving the rotation function to establish the model orientation, followed by the translation function to position the model correctly within the unit cell. Once

both are defined, the model provides approximate phases for all reflections, enabling the calculation of the electron density maps. These maps, derived from model phases and observed amplitudes, can be interpreted and refined to reveal the unknown structure (208–210).

All structures presented in this work were solved using the program PHASER from the CCP4 suite. Phaser enables the phasing of macromolecular crystal structures through both molecular replacement and experimental phasing approaches. The program incorporates advanced algorithms based on maximum likelihood and multivariate statistics, which significantly improve the reliability of the phase determination (211). In crystallography, maximum likelihood methods provide a powerful framework for statistical inference because they identify the model that explains the observed diffraction data with the highest probability.

5.2.22. Model refinement

The goal of crystallographic refinement is to optimize the agreement between the observed diffraction data and the calculated data derived from an atomic model, while simultaneously maintaining chemically reasonable local geometry and non-bonded interactions (212). A well-refined model is essential for the meaningful interpretation of the structure, including its biological function and interactions with other molecules. Following real-space model building, numerous small stereochemical and conformational inaccuracies typically remain, which are corrected through reciprocal-space restrained refinement. In this process, structural parameters, such as atomic coordinates, atomic displacement parameters (B-factors), and global corrections (e.g., scale factors, bulk solvent contributions, or anisotropy corrections), are iteratively adjusted against all available experimental data (213). The quality of the refinement was quantitatively assessed using residual indices, most notably the R-value, which measures the agreement between the observed F_{obs} and calculated F_{calc} structure factor amplitudes:

$$R = \frac{\sum_{hkl} |F_{obs}(hkl) - F_{calc}(hkl)|}{\sum_{hkl} F_{obs}(hkl)}$$

To prevent overfitting, cross-validation was conducted by calculating a similar statistic, R_{free} on a randomly chosen subset of reflections that were excluded from refinement:

$$R_{free} = \frac{\sum_{hkl \epsilon T} |F_{obs}(hkl) - F_{calc}(hkl)|}{\sum_{hkl \epsilon T} F_{obs}(hkl)}$$

where T denotes the test set of reflections. Together, these indicators provide a rigorous assessment of the global fit between model and data, while safeguarding against model bias and over-parameterization (214).

Refinement was performed in successive cycles, each consisting of manual model adjustments followed by computational optimization to minimize the differences between the observed and calculated structure factor amplitudes. Model building was performed using the molecular graphics program Coot, which allows corrections to be guided by electron density maps. The $2F_o$–Fc maps provide an overall representation of the electron density and assist in accurate atom placement, whereas the F_o–Fc difference maps highlight discrepancies between the model and experimental data, indicating incorrectly positioned atoms or unmodeled features. This iterative process, which combines manual rebuilding in Coot (215) with refinement calculations performed using REFMAC5 (216), aims to optimize both the stereochemical quality of the model and its fit to the diffraction data. The progress of refinement and accuracy of the resulting structure were assessed using statistical indicators R and R_{free}. For well-refined structures, the R index should typically not exceed 20%, and the difference between $Rfree$ and R should generally remain within approximately 3-7 percentage points (195). Following refinement, a new electron density map is generated from the improved structure, serving as the basis for subsequent model reconstruction cycles (195, 217–219).

Effective refinement of model parameters requires the incorporation of additional restraints, often referred to as soft constraints. These restraints provide prior information about the expected geometry of nucleic acids, typically derived from high-resolution small-molecule structures, and help to maintain chemically realistic conformations (220). The strength of the applied restraints depends on the resolution and overall quality of the diffraction data: higher-

resolution datasets allow weaker restraints, while lower-resolution data require stronger ones. In addition to stereochemical restraints, refinement can be improved by the use of TLS (translation, libration, screw) parameters, which describe anisotropic motions of groups of atoms. This approach assumes that subsets of atoms behave approximately as rigid bodies, allowing complex collective vibrations to be modeled with only a limited number of parameters.

5.2.23. Other programs used in the work

- ChimeraX - programs for visualizing crystal structures.
- 3DNA - to analyze nucleic acid structures (221). Calculates the number of helical parameters.
- ImageJ - software used for viewing, processing, and quantitatively analyzing scientific images.
- ChemDraw - software tool used for drawing chemical structures and reaction schemes.

6. Results

6.1. Design and Synthesis of Aromatic Linkers

The initial phase of my doctoral thesis focused on the chemical synthesis of a series of non-nucleosidic aromatic linkers to covalently connect the 5' and 3' ends of an RNA hairpin structure. The linkers were designed based on pyromellitic, naphthalene, and their derivatives. The inspiration was taken from phenanthrene- and phenanthroline-derived linkers reported in the literature as potential DNA/RNA structural stabilizers by introducing rigidity and π–π stacking potential with nucleobases (105, 106, 222, 223). Although these linkers have attractive structural properties and are fully compatible with standard solid-phase RNA synthesis protocols, their synthesis is difficult and expensive. Therefore, the selection of pyromellitic and naphthalene moieties of the linkers was based on the low price of chemicals and the high yield of synthesis.

The synthesis of naphthalene- and pyromellitic-derived phosphoramidite building blocks with linkers of different lengths is shown in Figure 12.

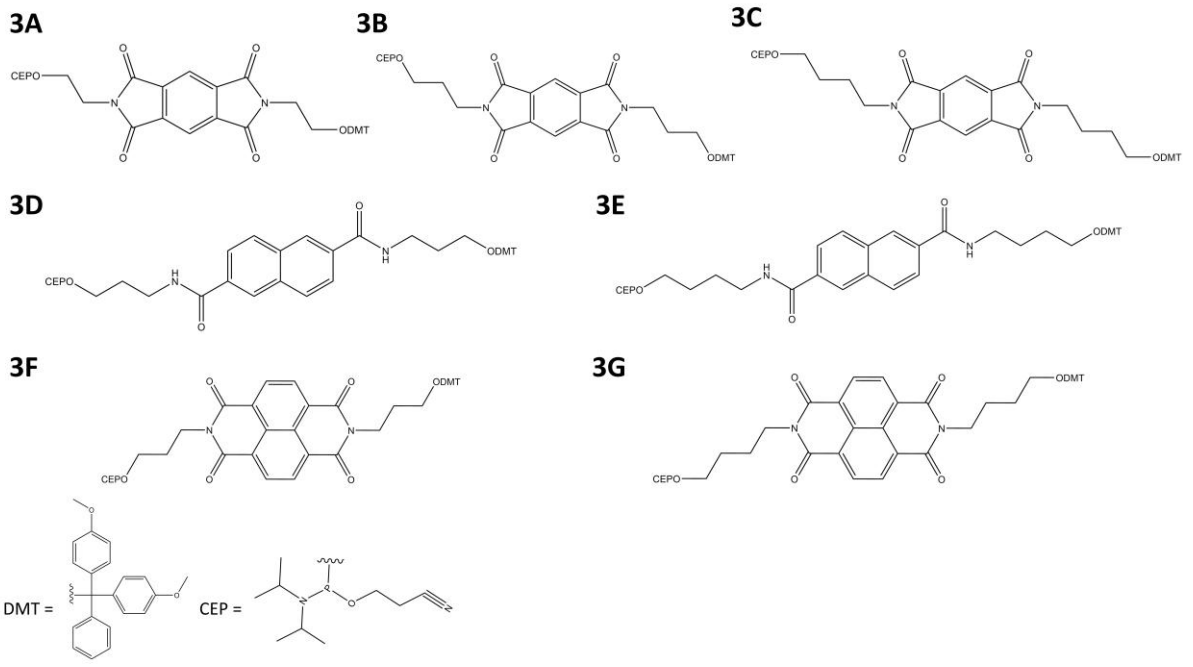
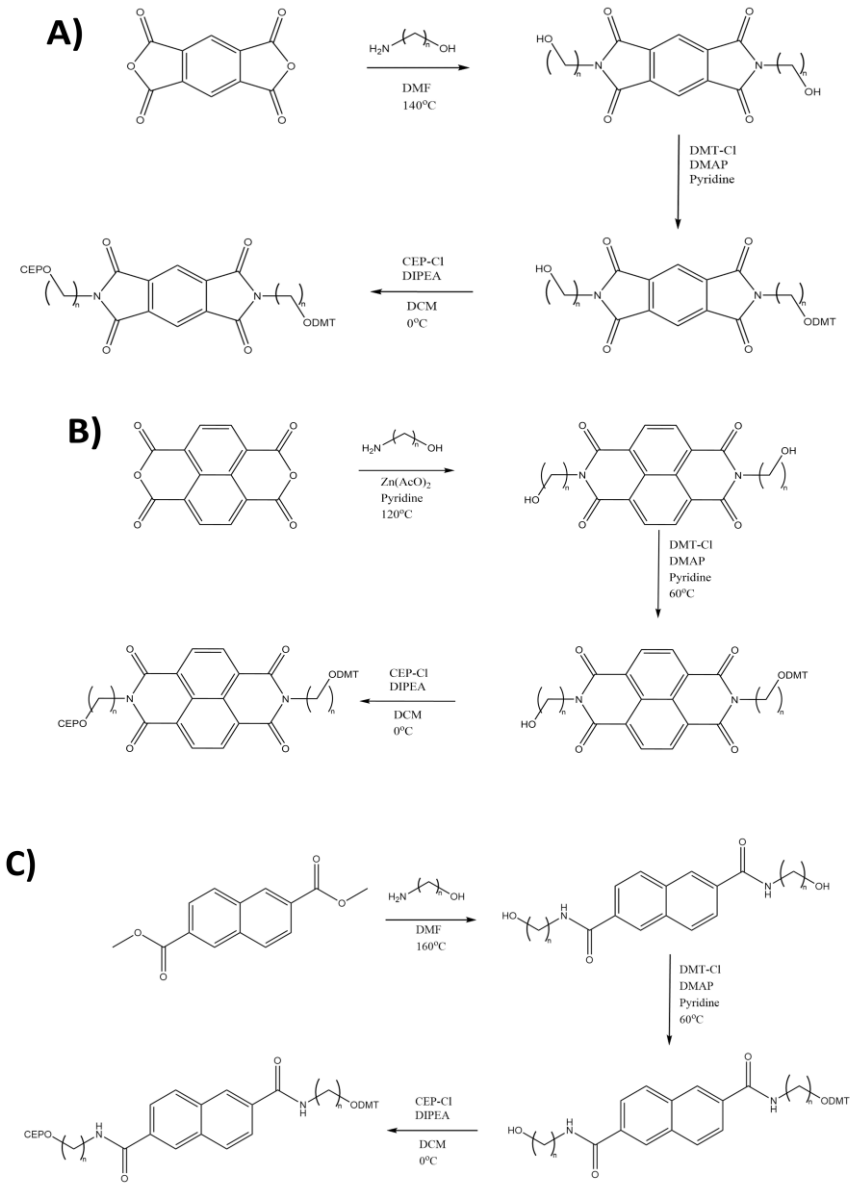


Figure 12. Chemical Structures of Naphthalene- and Pyromellitic-Based Aromatic Linkers.

Naphthalene and pyromellitic derivatives were prepared from benzene-1,2,4,5-tetracarboxylic (Scheme 1A), naphthalene-1,4,5,8-tetracarboxylic acid dianhydride (Scheme 1B) and dimethyl 2,6-naphthalene dicarboxylate (Scheme 1C). Derivatization with the

corresponding α,ω-aminoalcohols yielded amides 1a–e. These amides were subsequently monotritylated using 4,4’-dimethoxytrityl chloride under controlled conditions. The final step involved phosphitylation, yielding phosphoramidites 3a–g (Scheme 1). Pyromellitic (3a-c) and naphthalene (3d-e) derivatives were proposed as new compounds, and their synthesis was described in patent application no. 13P53836PL00.



Scheme 1. Synthesis pathways of phosphoramidites derived from (a) pyromellitic, (b) naphthalene-1,4,5,8-tetracarboxylic, and (c) 2,6-naphthalene precursors.

6.2. Preparation of RNA sequences for structural research

The next step in the preparation of the stabilized RNA hairpin was the incorporation of phenanthrene-derived phosphoramidite building blocks 3a–g during standard solid-phase synthesis into the oligonucleotide chain. First, I designed and synthesized one set of oligonucleotides (class I) composed of a 5'-CUGCUGCUG-3' chain linked via an appropriate linker to a second chain of a 5'-CUGCUGCUG-3' sequence. The reference 5'-CUGCUGCUG-3' RNA molecule formed a self-complementary homoduplex, whereas the tethered derivatives contained an appropriate linker spanning two RNA strands, forming a pseudo-hairpin structure (Figure 13). Class I contained oligonucleotides consisting of only ribonucleotides within the sequence and were used to check whether the incorporation of synthesized linkers was feasible. After chemical synthesis, the oligomers were deblocked and purified according to the developed procedure described in the publication: *Overview of Methods for Large-Scale RNA Synthesis* (187).

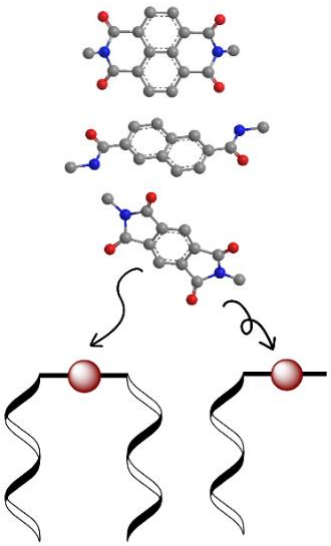


Figure 13. Schematic representation of pseudo hairpin and capped structures. Ribbon represents the chain of nucleotides, while red ball represents incorporated linker.

The purity and length of all synthesized oligonucleotides were assessed using 20% denaturing polyacrylamide gel electrophoresis and MALDI-TOF analyses. All results consistently showed that most class I tethered oligonucleotides were half the expected length (Figure 14 and Figure 16A), indicating cleavage of the RNA chain. Intramolecular hydrolysis and subsequent bond

cleavage most likely occurred in the linker vicinity due to the spatial proximity of the 2′-OH group to the covalent bond between the aromatic linker and adjacent nucleotide. The sole exception was the oligomer containing linker 3d, which retained its full expected length. Nevertheless, truncated oligonucleotides preserved valuable functional features, such as the aromatic linkers that behave as non-nucleotide caps, contributing to the structural stabilization of the molecule (Figure 13).

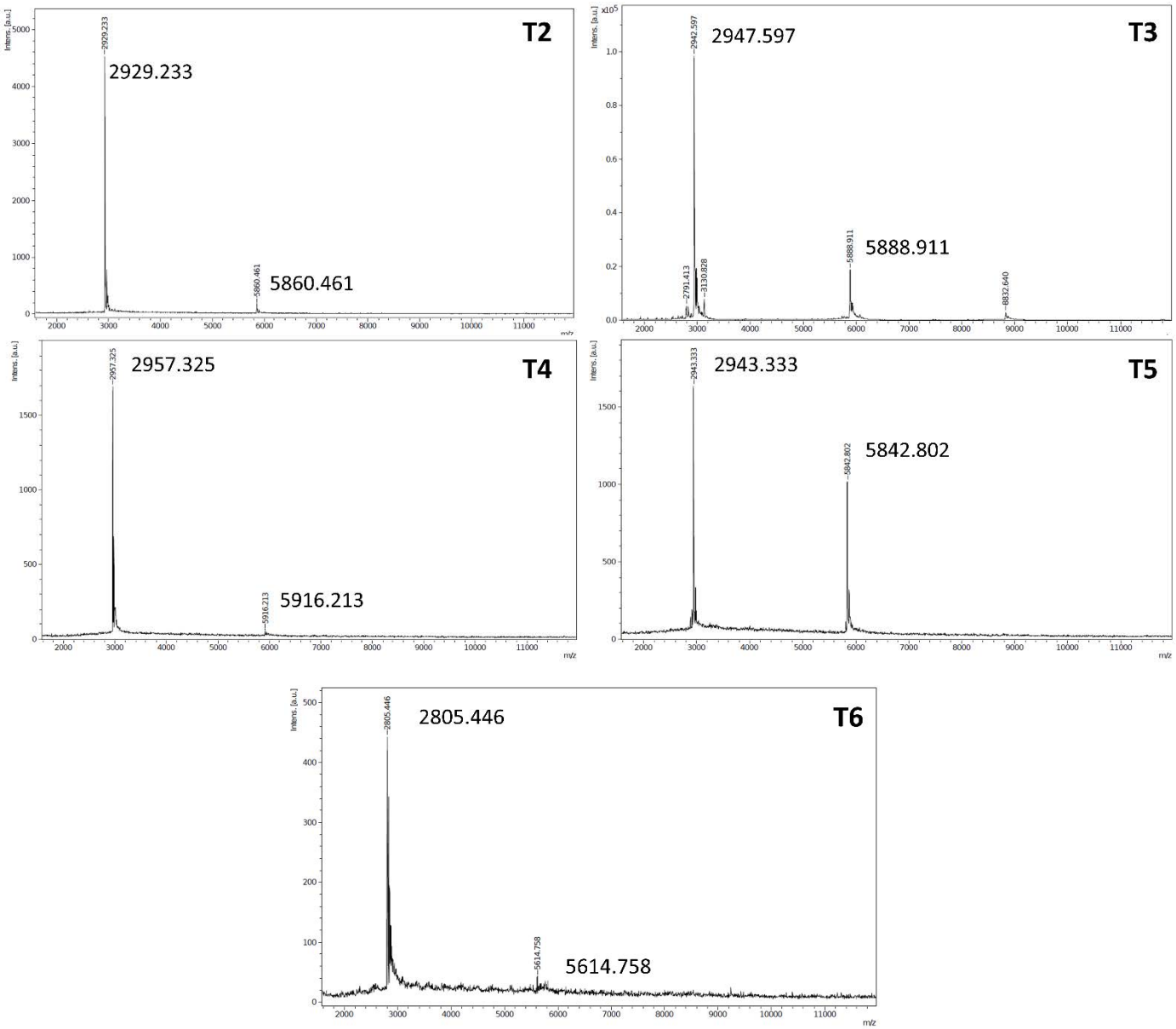


Figure 14. MALDI-TOF mass spectra of the chemically modified oligonucleotides.

Initial experiments showed that the modified RNA oligonucleotides were unstable and underwent cleavage. I found that this issue can be overcome by replacing the ribonucleotide with deoxyribonucleotides in close proximity to the linker. I designed a second class of oligomers (class II) containing hybrid RNA-DNA oligonucleotides. This time, the oligomers were composed of the shorter sequence 5’-GCUGdC-linker-dGCUGC-3’ (where dC and dG correspond to deoxycytosine and deoxyguanosine residues, respectively) to achieve the highest yield of tethered RNA. The lengths of class II oligomers were confirmed by 15% denaturing polyacrylamide gel electrophoresis (PAGE) and ESI-MS analyses (Figure 15 and 16B). The separation of oligomers using PAGE showed that the tethered oligomers had an expected migration rate relative to the control line.

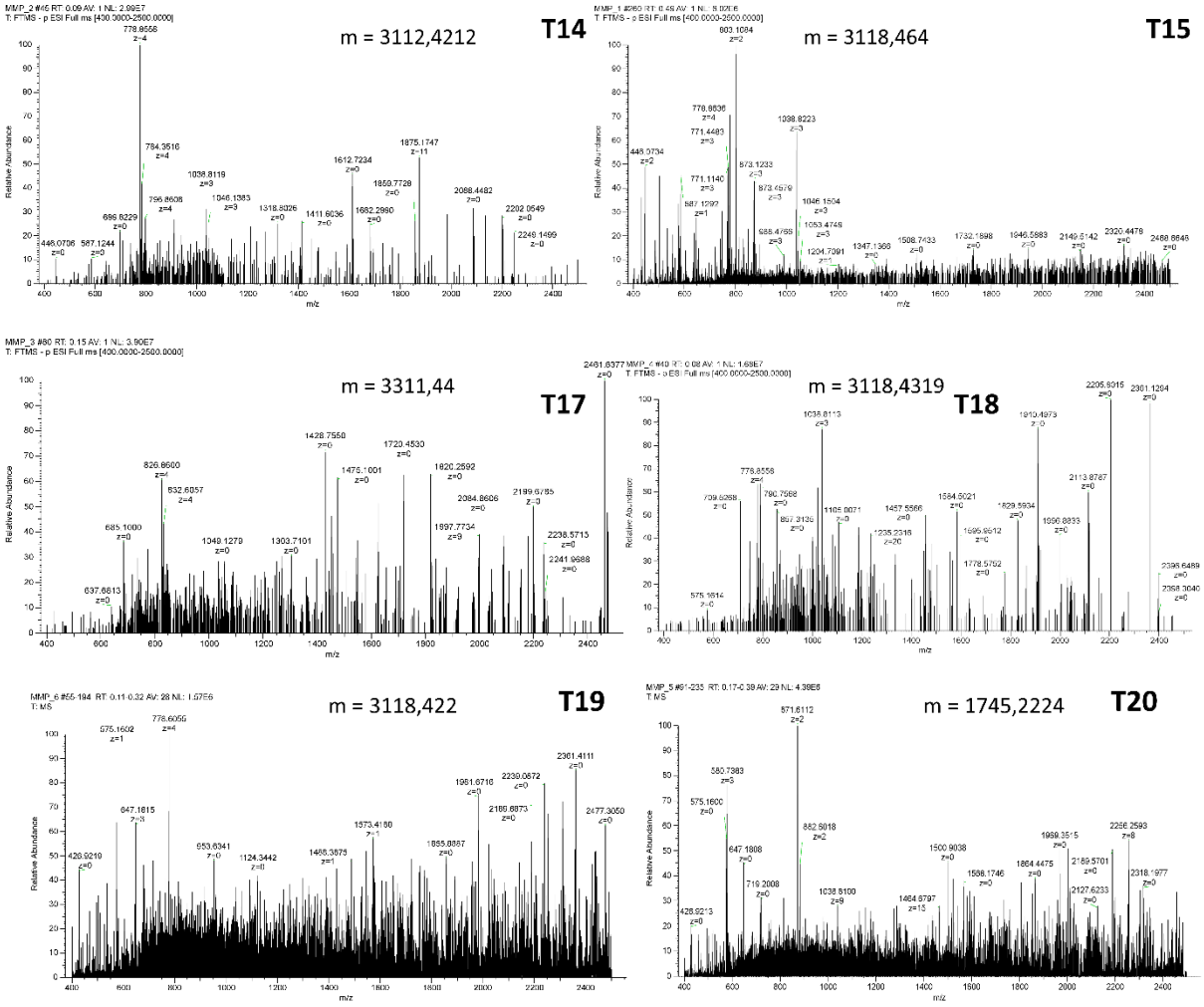


Figure 15. ESI-MS mass spectra of the chemically modified oligonucleotides.

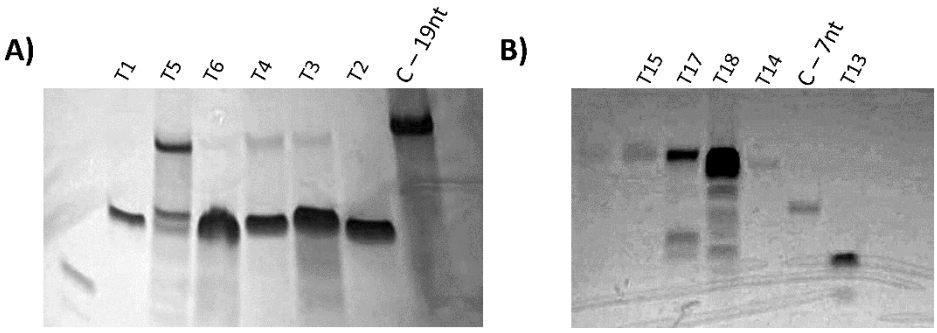


Figure 16. 20% Denaturating poliacrylamide gel of B) Class I oligomers and C) class II representative oligomers.

6.2.1. Physicochemical Evaluation of Modified Oligonucleotides

Following the design phase, the modified oligonucleotides were subjected to physicochemical evaluation to verify their folding behavior and thermodynamic stability. These analyses provided critical insights into which linker geometries and chemical compositions were most compatible with the RNA scaffold.

	Oligonucleotide	dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37ºC)	$-\Delta G_{298}$ (kcal/mol, 37ºC)	Tm (ºC)
			Class I			
T1	CUGCUGCUG	-57.67 ± 6.719	-173.90 ± 22.148	-3.74 ± 0.273	-	26.9
T2	CUGCUGCUG-3a-CUGCUGCUG	-57.33 ± 7.116	-171.40 ± 22.992	-4.17 ± 0.285	0.43	29.1
T3	CUGCUGCUG-3b-CUGCUGCUG	-56.00 ± 1.967	-165.71 ± 6.374	-4.61 ± 0.117	0.87	31.2
T4	CUGCUGCUG-3c-CUGCUGCUG	-52.49 ± 3.747	-156.08 ± 12.160	-4.08 ± 0.300	0.34	27.8
T5	CUGCUGCUG-3d-CUGCUGCUG	-60.76 ± 7.631	-165.67 ± 28.104	-9.38 ± 1.787	5.64	57.1
T6	CUGCUGCUG-3e-CUGCUGCUG	-56.72 ± 2.335	-167.60 ± 7.577	-4.73 ± 0.058	0.99	31.9
			Class II			
T13	GCUGC	-29.96 ± 3.360	-87.48 ± 10.921	-2.83 ± 0.421	-	10.1
T14	GCUGdC-3a-dGCUGC	-59.62 ± 6.199	-176.68 ± 20.065	-4.82 ± 0.192	1.99	32.6
T15	GCUGdC-3b-dGCUGC	-46.12 ± 6.905	-134.72 ± 21.913	-4.34 ± 0.361	1.51	28.3
T17	GCUGdC-3d-dGCUGC	-53.81 ± 8.886	-159.13 ± 29.451	-4.46 ± 0.282	1.63	30.1
T18	GCUGdC-3e-dGCUGC	-64.36 ± 14.555	-193.34 ± 48.049	-4.40 ± 0.439	1.57	31.0
T19	GCUGdC-3f-dGCUGC	-59.97 ± 5.595	179.66 ± 22.305	-4.25 ± 0.542	1.42	29.8
T20	GCUGdC-3g-dGCUGC	-30.07 ± 3.224	-87.16 ± 11.698	-3.04 ± 0.462	0.21	12

Table 1. Influence of pyromellitic and naphthalene linkers on the thermal stability of RNA oligonucleotides.

As part of the physicochemical evaluation, I performed thermal denaturation monitored by UV spectroscopy of both classes of tethered oligomers to evaluate the influence of

pyromellitic and naphthalene building blocks on RNA structural stability. The thermodynamic parameters, enthalpy (ΔH), entropy (ΔS), and free energy at 37 °C (ΔG_{37}), were extracted from the fits to the melting profiles (224). All tethered oligonucleotides displayed enhanced thermodynamic stability compared to their native counterparts, as reflected by increases in ΔG_{37}, ranging from 0.34 to 5.64 kcal/mol. Within class I, the **T5** oligomer exhibited the greatest stabilization and highest melting temperature (Tm). This pronounced effect is attributed to its predominant pseudohairpin conformation, in contrast to the other class I constructs, which mainly adopted a capped form of the oligomer (Figure 13). Among these capped variants, **T6** provided the strongest stabilizing contribution (ΔG = 0.99 kcal/mol). For class II, the variations in stability were more subtle and less pronounced. Nevertheless, the highest stabilization was observed for oligonucleotides with linkers having shorter aliphatic segments, specifically **T14** and **T17**. A comparison of the ΔG values obtained for **T5** of class I and **T17** of class II molecules shows that the increase in stability is greater for the longer RNA. These observations collectively indicate that both linker and RNA length and sequence play critical roles in modulating the thermodynamic stability of RNA structures (Table 1 and Figure 17).

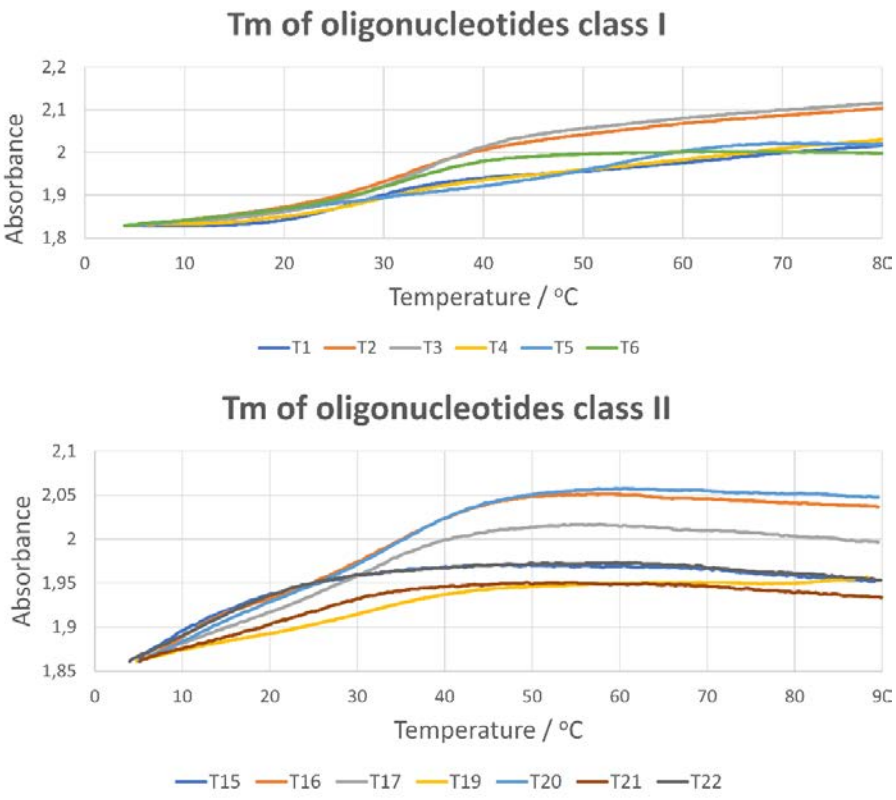


Figure 17. Thermal melting curve of oligonucleotides containing aromatic linkers. A) Class I and B) class II oligonucleotides.

6.2.2. Crystallization of RNA constructs

Based on the physicochemical evaluation of linker-modified pseudohairpins, additional RNA constructs were designed for crystallization trials. Although these constructs were not further characterized thermodynamically, they were intended to explore the structural compatibility of selected linkers under crystallization conditions.

I designed and synthesized five sets of oligonucleotides: 5'-CUGCUGCUG-L-CUGCUGCUG-3' (**T2-T6**); 5'-GCAGCUGCUG-L-CUGCUGCUG-3' (**T7-T9**); 5'-CGCUG-L-CUGCG-3' (**T10-T12**); 5'-GCUGdC-L-dGCUGC-3' (**T14-T20**); 5'-CCGUGdG-L-dCCGUGG-3' (**T21-T24**). Two sets were previously evaluated (**T2-T6** and **T14-T20**). Each construct was engineered to form a stable double stranded region, with an aromatic linker replacing the natural hairpin loop (Figure 18). The sequences were composed primarily of G-C and C-G base pairs to enhance duplex stability and promote crystal formation by reducing structural flexibility. All constructs contained exclusively ribonucleotides, except for two variants (**T14-T20** and **T21-T24**) in which two nucleotides within the aromatic linker region were substituted with deoxyribonucleotides to improve linker incorporation and overall RNA stability. The designed constructs were subsequently subjected to crystallization screening to assess their ability to form ordered RNA assemblies suitable for X-ray diffraction analysis.

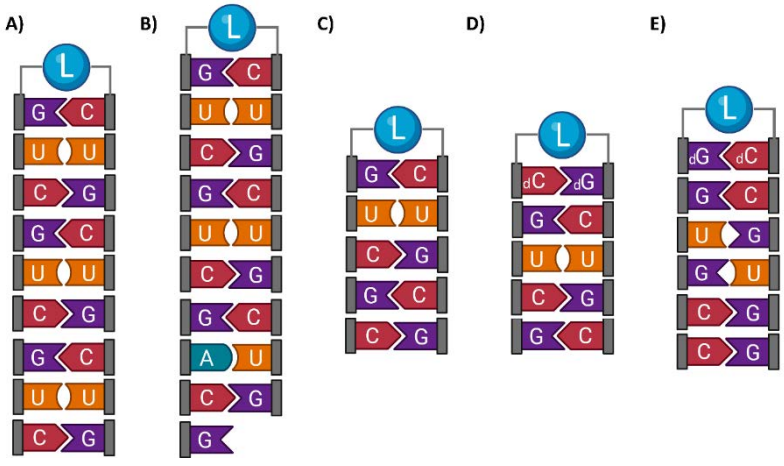


Figure 18. Constructs of linker-modified oligonucleotides engineered to form a stable duplex region with an aromatic linker replacing the natural hairpin loop. Letter **L** represents linker modification. Oligonucleotides A)**T2**-**T6** B) **T9**-**T9** C) **T10**-**T12** D) **T14**-**T20** E) **T21**-**T24**

Crystallization was carried out using the sitting-drop vapor diffusion method at 19 °C. Experiments were performed with two commercially available screening kits: Natrix™ and Crystal Screen™ (Hampton Research). Among the tested constructs, only one sequence, 5'-

CUGCUGCUG–3a–CUGCUGCUG-3’ (**T2**), produced diffraction-quality crystals (Figure 19). Crystals were obtained under the following conditions: 0.04 M lithium chloride, 0.08 M strontium chloride hexahydrate, 0.02 M magnesium chloride hexahydrate, 0.04 M sodium cacodylate trihydrate (pH 7.0), 30% (v/v) (±)-2-methyl-2,4-pentanediol, and 0.012 M spermine tetrahydrochloride.

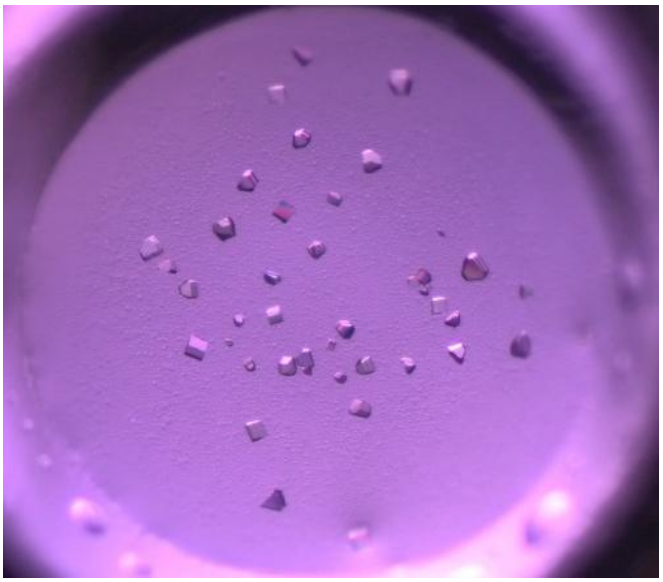


Figure 19. Crystals of the **T2** oligonucleotide.

6.2.3. X-ray diffraction experiments and data processing.

X-ray diffraction data were collected on the P13 beamline at the DESY synchrotron facility in Hamburg. The measurement was performed at 100 K using 30% (v/v) (±)-2-methyl-2,4-pentanediol as the cryoprotectant. 1800 diffraction patterns were recorded with a maximum resolution of 1.5 Å (Table 2).

Data Collection	
X-ray source	Dectris EIGER 16M,
Space group	R32
Cell parameters (Å)	a = 40.15, b = 40.15, c = 139.91
Resolution (Å)	46.64-1.5 (1.59-1.5)
R_{merge}	0.088 (1.46)
I/σ	12.7 (1.46)
$CC_{1/2}$	0.999 (0.955)
Completeness (%)	99.8 (99.0)
Redundancy	12.57 (12.23)
Number of unique reflections	7274
Refinement	
Software	Refmac 5.8.0425
Number of reflections: work/test	6902/355
Overall mean B value ($Å^2$)	48.67
R_{work}/R_{free}	0.2459/0.3632
RNA atoms	410
Water molecules	29
RMSD in bonds (Å)	0.012
RMSD in angles (°)	1.772

Table 2. Summary of X-ray data and model refinement statistics of **T2** oligonucleotide.

6.2.4. Solving and refinement of the structure

Analysis of the asymmetric unit content using the Matthews coefficient (225) indicated the presence of one duplex at a solvent content of 58.7 %. For molecular replacement, an extracted part of the structure containing 6 CUG repeats (pdb code: 3GM7) (226) was used as a search model. The structure was solved using the PHASER program from the CCP4 program suite.

Despite the high resolution of the collected dataset (1.5 Å), the refinement process did not yield satisfactory R_{work}/R_{free} statistics. The final values were 0.2459/0.3632 (Table 2). This discrepancy is likely a consequence of the poor quality of the diffraction images, which exhibited streaking and partially split, diffuse reflections (Figure 20). These features are consistent with the presence of an order–disorder effect within the crystal lattice, leading to a superposition of slightly shifted lattices or heterogeneous conformations. These effects reduce the internal consistency of the measured intensities and limit the achievable accuracy

during refinement, ultimately preventing the convergence of the model to the values expected for data of this resolution.

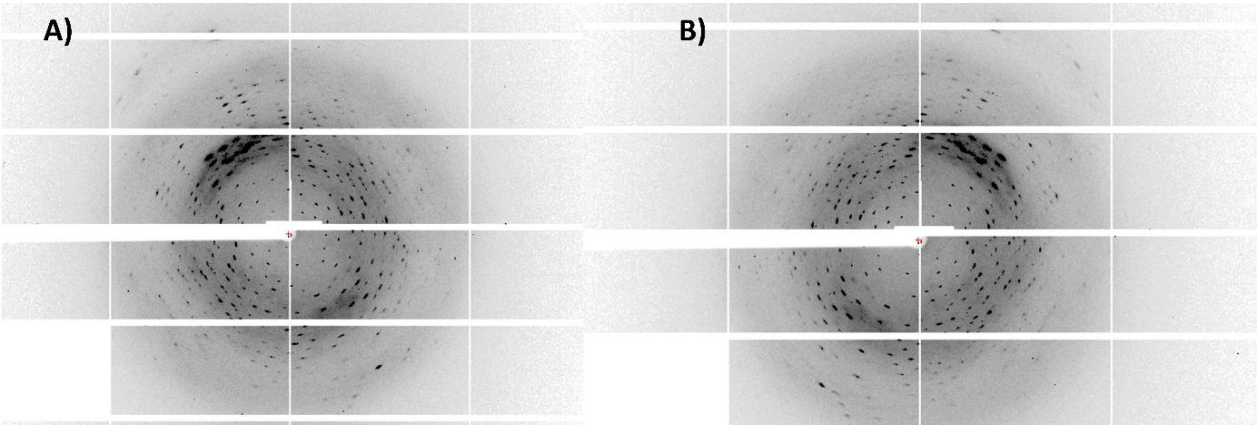


Figure 20. Diffraction images of the **T2** oligonucleotide crystal: (A) 1st frame, (B) 900th frame.

6.2.5. Structural analysis

Structural analysis of the RNA duplex revealed that the helical region adopts a canonical A-form geometry (Figure 21A). The sugar residues predominantly occupy the **C3'-endo** conformation, with occasional occurrences of **C2'-endo**, which is fully consistent with the typical A-RNA dynamics. Only one nucleotide adopts an **O4'-endo** pucker; however, this deviation appears at the terminal base pair, where substantial structural disorder is present. This local disorder is likely a consequence of the disrupted architecture near the site where the non-nucleosidic linker was expected to reside, resulting in increased flexibility and poor definition of the helix end. Helical parameters calculated using sequence-independent vectors defined by the positions of the C1' atoms of paired nucleotides further confirmed the A-RNA character of the duplex with **Zp > 1.5°**, **twist = 30.6°**, and **rise = 3.13 Å**. Deviations from the ideal A-form values were observed only for the terminal base pairs (**shear = 2.27 Å**, **propeller = -21.6 °** and **opening = -29.8 °**), reflecting the conformational disorder described above (Table 3).

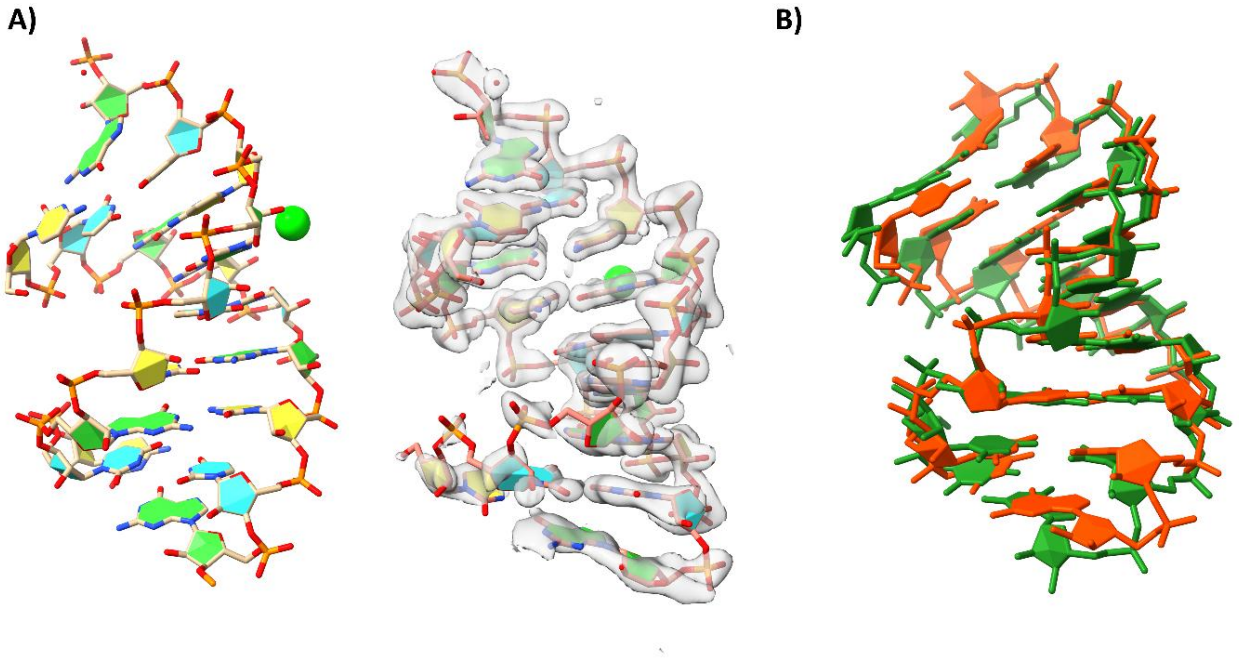


Figure 21. Crystal structure of **T2** oligonucleotide. A) Tertiary structure models of the duplex (left) and the Fo – Fc omit map of **T2.** contoured at the σ 1.0 level. B) Superposition of **T2** crystal structure and **3GM7** model.

The obtained model was compared with a reference structure containing three CUG repeats derived from the crystal structure of **3GM7**. For this purpose, an equivalent duplex fragment comprising three CUG units was extracted from 3GM7 and subjected to structural alignment with an experimentally determined model. The helical parameters of both duplexes were highly comparable, confirming that the presence of the disordered terminal region in our construct did not perturb the overall A-form geometry of the helix (Table 3). Superposition of the two structures yielded an RMSD of **1.095 Å** calculated over the main-chain atoms (Figure 21B), demonstrating a strong structural similarity between the models and supporting the conclusion that the ordered portion of the duplex is well defined and adopts a canonical A-form conformation.

T2 RNA

Base pairs	Shift	Slide	Rise	Tilt	Roll	Twist	h-Rise	Incl.	Tip	h-Twist	Zp
1 CU/UG	-0.79	-2.50	2.96	-4.77	17.99	15.90	0.26	48.16	12.77	24.44	2.52
2 UG/CU	0.70	-1.58	3.23	-0.08	9.15	35.13	2.74	14.85	0.12	36.27	2.71
3 GC/GC	-0.07	-1.68	3.10	2.32	3.83	35.00	2.90	6.33	-3.84	35.28	2.64
4 CU/UG	-0.25	-2.24	3.02	1.03	8.15	26.39	2.23	17.33	-2.18	27.62	2.74
5 UG/CU	-0.13	-1.75	3.20	-2.29	11.50	32.35	2.46	19.84	3.95	34.36	2.60
6 GC/GC	-0.03	-1.41	3.21	0.82	8.78	32.37	2.74	15.39	-1.44	33.52	2.30
7 CU/UG	1.01	-0.99	3.49	3.29	14.21	43.16	3.10	18.69	-4.32	45.44	2.28
8 UG/CU	-0.18	-1.19	2.84	9.13	15.45	24.17	1.64	31.88	-18.83	30.03	2.12
ave.	0.03	-1.67	3.13	1.18	11.13	30.56	2.26	21.56	-1.72	33.37	2.5

Base pairs	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	2.27	-0.25	-0.07	2.26	-21.63	-29.84
2 U-U	-0.05	-0.93	0.23	-1.32	-18.21	-11.32
3 G-C	-0.59	-0.30	0.28	-7.77	-15.69	2.61
4 C-G	0.19	-0.12	0.13	-3.65	-6.62	-0.56
5 U-U	-0.38	-0.26	-0.05	1.90	-13.67	-3.30
6 G-C	-0.30	-0.16	-0.01	1.65	-8.42	-0.12
7 C-G	0.41	-0.27	-0.20	5.76	-12.48	-1.05
8 U-U	2.29	-1.84	0.36	-2.14	-18.67	9.90
9 G-C	-0.12	-0.17	0.04	4.39	-11.59	3.33
ave.	0.41	-0.48	0.08	0.12	-14.11	-3.37

3GM7

Base pairs	Shift	Slide	Rise	Tilt	Roll	Twist	h-Rise	Incl.	Tip	h-Twist	Zp
1 CU/UG	-0.11	-2.33	3.56	-5.65	15.95	30.00	2.08	28.17	9.98	34.35	2.71
2 UG/CU	0.42	-1.60	3.40	-5.23	5.82	50.44	3.16	6.78	6.10	51.00	2.83
3 GC/GC	0.08	-2.14	3.17	3.88	8.88	27.22	2.36	18.15	-7.94	28.86	2.85
4 CU/UG	-1.13	-1.71	3.42	3.63	14.80	42.43	2.62	19.70	-4.83	44.97	2.56
5 UG/CU	0.83	-2.30	3.06	-1.78	10.87	21.06	1.62	27.46	4.50	23.73	2.79
6 GC/GC	0.30	-1.83	3.03	2.27	3.27	34.06	2.86	5.55	-3.85	34.29	2.86
7 CU/UG	-0.56	-1.80	3.10	-1.44	3.56	19.73	2.77	10.27	4.14	20.10	2.67
8 UG/CU	1.34	-1.92	3.20	2.22	12.88	47.43	2.69	15.67	-2.70	49.10	3.08
9 GC/GC	-0.01	-1.64	3.41	-1.03	5.61	33.98	3.11	9.52	1.75	34.44	2.58
10 CU/UG	-1.13	-1.34	3.25	4.76	13.81	40.24	2.54	19.32	-6.66	42.70	2.38
11 UG/CU	0.77	-2.12	3.46	0.90	10.44	20.62	2.17	27.03	-2.33	23.10	2.30
12 GC/GC	0.73	-2.36	3.08	0.09	6.39	27.38	2.48	13.27	-0.19	28.10	2.83
13 CU/UG	-1.99	-1.80	3.10	-0.44	-0.53	44.60	3.14	-0.70	0.58	44.61	2.71
14 UG/CU	0.74	-2.43	3.35	0.34	15.43	17.74	0.96	41.30	-0.92	23.47	2.78
15 GC/GC	0.23	-1.63	3.21	0.04	7.96	32.58	2.74	13.94	-0.06	33.51	2.56
16 CU/UG	-0.82	-1.65	3.47	4.69	14.20	42.54	2.71	18.89	-6.23	44.97	2.74
17 UG/CU	0.46	-1.77	2.88	-1.69	6.39	25.07	2.33	14.41	3.80	25.91	2.55
ave.	0.01	-1.90	3.25	0.33	9.16	32.77	2.49	16.98	-0.29	34.54	2.7

Base pairs	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-1.44	-0.20	-0.35	11.13	-19.16	-13.19
2 U-U	-2.32	-1.38	-0.18	0.89	-9.98	-21.23
3 G-C	1.16	-0.45	0.22	-0.68	-7.19	-6.77
4 C-G	0.22	-0.35	0.02	-0.94	-7.69	0.77
5 U-U	2.27	-1.17	-0.36	-6.03	-15.03	-22.06
6 G-C	0.08	-0.35	0.36	-3.33	-10.68	-3.51
7 C-G	0.14	-0.43	0.30	2.58	-9.01	-1.46
8 U-U	-2.50	-1.48	0.45	2.36	-11.27	-21.02
9 G-C	-0.14	-0.43	0.24	1.62	-10.12	-0.72
10 C-G	0.31	-0.38	0.17	-5.35	-14.66	1.65
11 U-U	1.71	-1.29	-0.44	-1.86	-12.20	-11.93
12 G-C	-0.30	-0.44	-0.17	-7.69	-6.45	-5.13
13 C-G	-0.22	-0.13	0.12	-2.45	-2.51	1.85
14 U-U	2.65	-1.29	-0.23	6.92	-11.44	-28.01
15 G-C	-0.09	-0.47	0.19	2.55	-9.31	-2.21
16 C-G	0.13	-0.03	0.16	2.26	-9.96	2.03
17 U-U	2.04	-1.11	-0.15	-4.69	-14.81	-14.80
18 G-C	-0.20	-0.32	0.17	2.01	-7.03	1.25
ave.	0.19	-0.65	0.03	-0.04	-10.47	-8.03

Table 3. Helical parameters calculated using 3DNA, based on C1'- C1' vectors for **T2** oligonucleotide and crystal structure of CUG repeats (pdb code: **3GM7**).

Although the refinement statistics were unacceptable and did not support a fully reliable structural solution, the electron density maps revealed a well-defined and continuous helical region corresponding to the ordered portion of the RNA duplex. Within the crystal lattice, this duplex participates in pseudocontinuous stacking interactions, extending the helical axis across symmetry-related molecules. This packing mode effectively occupies the spatial region where the non-nucleosidic linker would normally be accommodated, leaving no available space for its ordered incorporation. The geometry of the helix termini showed no classical stacking interactions; instead, the bases were displaced and misaligned, relatively to nucleotides in the reference model, a feature also reflected in the deviations of the helical parameters at the duplex ends (Table 3). Such disrupted terminal geometry is consistent with the presence of an order–disorder effect, which is further supported by the poor quality of the diffraction images, indicating structural heterogeneity that likely prevented proper refinement.

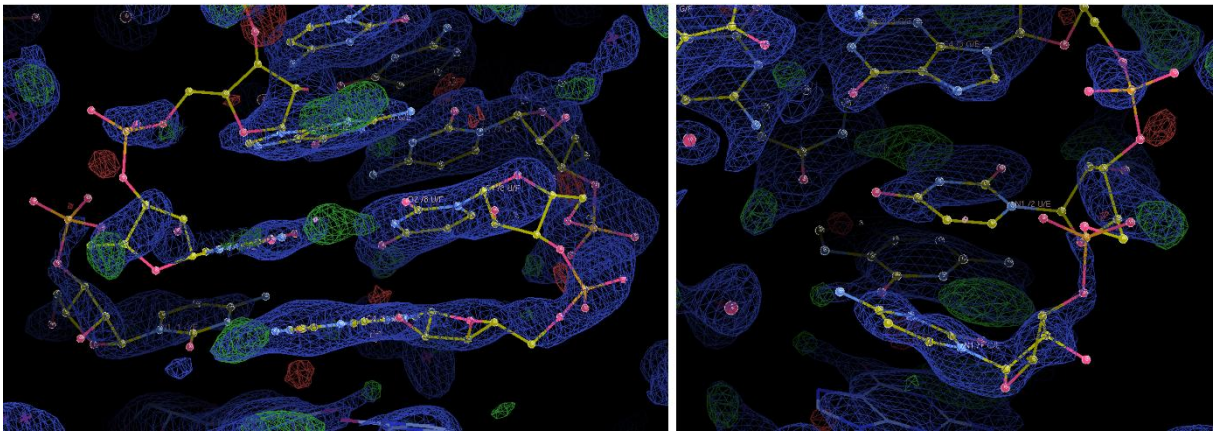


Figure 22. Unmodeled electron density attributed to nucleotide order–disorder for the terminal base pairs 1C-9G, 2U-8U and 3G-7C.

In the area where the linker was expected, electron density was entirely absent, preventing reliable model building (Figure 22). Only the phosphate group of the linker was well resolved (Figure 21A). To validate the presence of this terminal phosphate, refinement was performed using two alternative occupancy models: full occupancy (1.0) and zero occupancy (0.0). When phosphate was omitted (occupancy 0.0), a clear positive peak appeared in the Fo–Fc omit map, indicating missing electron density at this position. Upon refining the phosphate group at full occupancy, the omit density disappeared, and the phosphate fit cleanly and consistently into the electron density, confirming its presence (Figure 23).

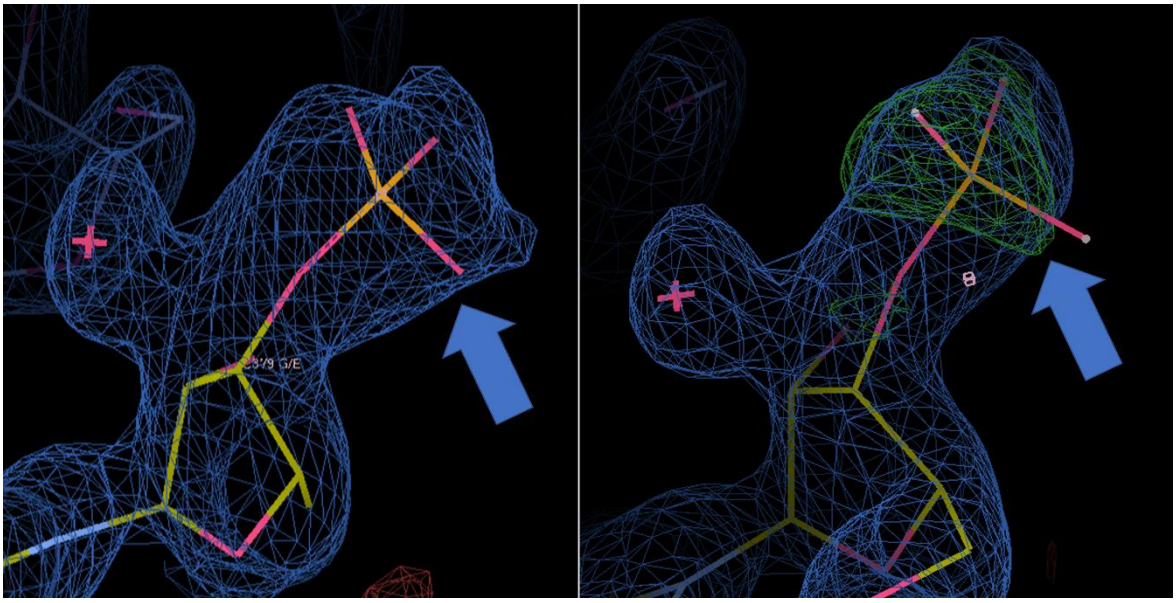


Figure 23. Refinement of the 3'-terminal phosphate group: (left) full-occupancy model, (right) zero-occupancy model. P Phosphate groups are indicated by the blue arrows.

Although the presence of the phosphate group confirmed the introduction of the linker during chemical synthesis, its presence in the crystal structure could not be confirmed by inspection of the electron density map. Therefore, it is possible that either the linker was cleaved during RNA deprotection after chemical synthesis or remained disordered. The elevated B-factors in adjacent nucleotides to the linker indicate increased local flexibility, which decreases progressively with increasing distance from the putative linker (Table 4). This increased flexibility can be attributed to the presence of a disordered linker.

Base pair	Average B-factor in chain A (Å^2)	Average B-factor in chain B (Å^2)
1C-9G	80.8	72.5
2U-8U	88	59
3G-7C	73	48.4

Table 4. Average values of B-factors for terminal base pairs of **T2** model.

6.3. Development of a method for stabilizing RNA hairpin structures under crystallization conditions

To overcome the tendency of RNA hairpin constructs to unwind or form semicomplementary duplexes at the high concentrations required for crystallization, I used a strategy, according to a publication by Kiliszek et. al. (98), to span the 5’ and 3’ end of a stem by linker resulting in circular hairpin. In this approach, the native loop conformation is preserved, while additional covalent stabilization and circularization are provided to prevent RNA from unfolding during crystallization. The linker is introduced into the hairpin sequence during RNA chemical synthesis. The resulting oligomer should fold into a pseudo-hairpin. In the next step, the apical loop is reconstructed by ligation of the 5’ and 3’ ends of the pseudo-hairpin using T4 RNA ligase. For the circularization reaction, T4 RNA ligase requires the RNA molecule to possess a 5’-monophosphate and 3’-hydroxyl group.

6.3.1. Designing the sequence of RNA oligomers forming a hairpin structure

To test a method and develop protocol for stabilizing the form of the hairpin created by RNA molecules, I used biologically significant RNA hairpin structures: sarcin-ricin loop (SRL) (one of the longest conserved sequences in the 23S rRNA crucial for the activity of the ribosome because it is targeted by cytotoxins such as α-sarcin and ricin that completely abolish translation) (227); TAR loop (the trans-activating responsive RNA element, located at the 5' untranslated end of all viral transcripts) (228, 229) and R06 (aptamer specific for the TAR RNA element of HIV-1 forming kissing complex) (Figure 24) (229–232).

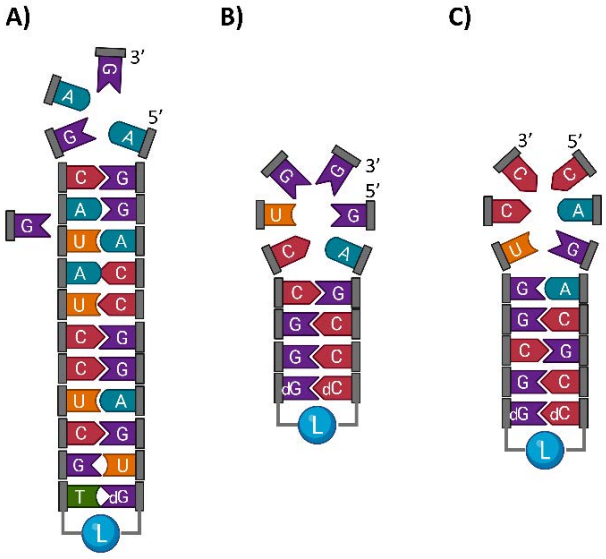


Figure 24. Constructs of linker-modified oligonucleotides engineered to form hairpin. Letter **L** in blue circle represents linker. Oligonucleotides A) **T26**-**T27** B) **T29**-**T32** C) **T34**-**T37**

6.3.2. Oligomer circularization reaction

In all oligonucleotides, the ligation site is located within the apical loop. The following linkers, which exhibited the best stabilization properties, were incorporated into the RNA sequences: 3a and 3d. Hexaethylene glycol linker (commercially known as spacer 18) was used as a reference molecule. After chemical synthesis, the oligomers were deblocked and purified according to the standard procedure described in the paper: *Overview of Methods for Large-Scale RNA Synthesis* (187). Oligomers were used to perform enzymatic circularization of RNA employing T4 RNA ligase. The circularization reaction was carried out following the protocol described in the *Handbook of RNA Biochemistry* and according to the publication by Kiliszek et al. (98). Sample preparation involved several steps:

- chemical synthesis of modified oligonucleotides,
- deprotection and desalting of RNA,
- purification of crude RNA on a 20% polyacrylamide gel, and subsequent electroelution of the RNA from the gel and then precipitated with isopropanol.
- introducing a phosphate group at the 5′ end using polynucleotide kinase,
- RNA circularization with T4 RNA ligase,
- Purification of the circularized RNA hairpin on a 20% polyacrylamide gel and precipitation with isopropanol.

Among all the tested constructs, successful circularization was achieved only for one RNA sequence (**T35** – R06 sequence) containing the 3d aromatic linker and for the control construct (**T37**) with the flexible spacer 18 linker (Figure 25). Both circularized and non-circularized RNA oligomers were verified by MALDI TOF mass spectrometry to confirm their integrity and expected molecular weight (Figure 26).

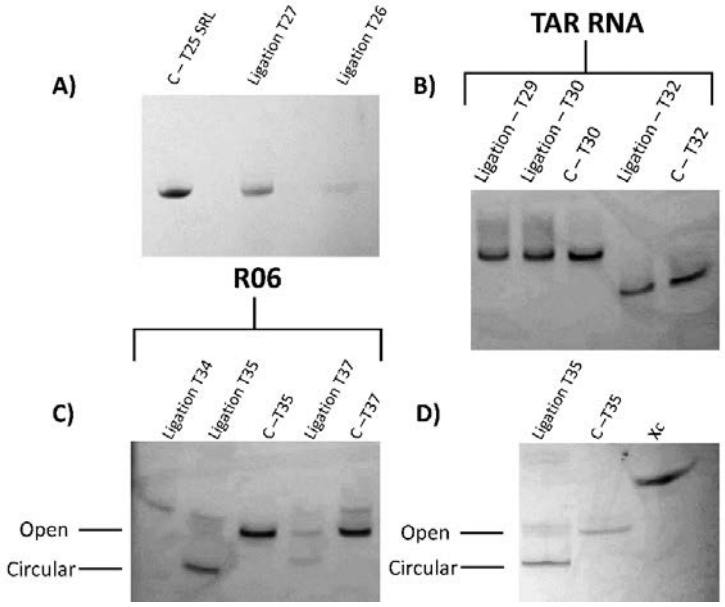


Figure 25. Circularization of RNA using T4 RNA ligase. Ligation reactions resolved on an 20% polyacrylamide gel in denaturing conditions. Open and circular forms of the RNA oligomers are indicated. Ligation reaction for A) Sarcin-ricin loop B) TAR-RNA C) R06. D) Ligation reaction on large scale for crystallographic studies

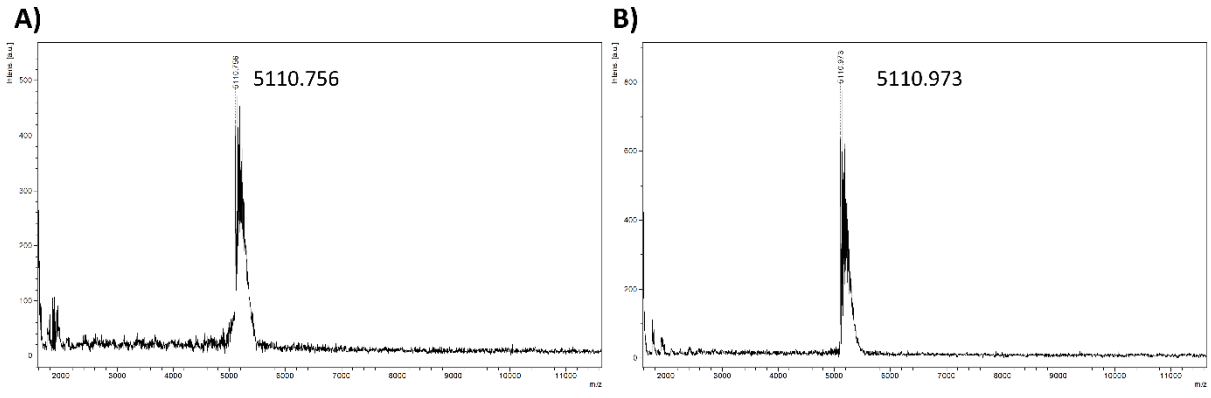


Figure 26. MALDI-TOF mass spectra of the chemically modified oligonucleotide **T35**. A) Circularized form and B) Open form.

The ligation reactions were analyzed on 20% denaturing polyacrylamide gels, which clearly showed the presence of circularized products in the reaction mixture (Figure 25C and Figure 25D). In these cases, approximately 50% of the open RNA oligomers were converted into the

circular form using T4 RNA ligase. Moreover, the circularized species migrated faster than the open substrates, consistent with the formation of a more compact, covalently closed structure (Figure 25C). Crystals were obtained in 0.2 M Potassium chloride, 0.01 M Calcium chloride dihydrate, 0.05 M Sodium cacodylate trihydrate pH 6.0, 10 % w/v Polyethylene glycol 4,000 for the complex of the modified R06 oligonucleotide with the TAR-HIV RNA **T28** (Figure 27). However, diffraction measurements revealed that the crystals did not diffract X-rays, indicating that the internal order of the lattice was insufficient for structural analysis.

Figure 27. Crystals of complex of circularized **T35** form of and non-ligated **T28** oligonucleotides.

6.3.3. Physicochemical evaluation of circularized RNA structure

To assess the structural integrity and thermodynamic properties of the circularized RNA hairpin compared to its linear (open) form, thermal denaturation monitored by UV spectrometry and circular dichroism (CD) spectroscopy were performed. These analyses aimed to evaluate the influence of covalent circularization on the overall stability (UV melting) and conformational behavior (CD) of the RNA.

6.3.3.1. Ultraviolet melting temperature measurements

UV melting experiments were performed for two RNA oligomers: the circularized and open forms of the R06 construct containing the 3D linker (**T35**). Thermodynamic parameters, including enthalpy (ΔH), entropy (ΔS), and free energy at 37 °C (ΔG_{37}), were derived by fitting the individual melting curves using the MARVIN program. The results revealed that circularization led to an increase in the thermodynamic stability of the RNA constructs, as reflected by a change in ΔG_{37} of 1.22 kcal/mol. Consequently, the melting temperature (T_m) of the circularized oligomers was approximately 2.1 °C higher than that of their open

counterparts, confirming the stabilizing effect of covalent closure on the RNA hairpin structure (Table 5 and Figure 28).

Oligonucleotide	dH (kcal/mol)	dS (cal/K·mol)	dG_{37} (kcal/mol, 37°C)	$-\Delta G_{298}$ (kcal/mol, 37°C)	T_m (°C)
Open T35	- 36.01 ± 0.778	-105.26 ± 2.59	-3.37 ± 0.168	-	69
Circularized T35	-47.16 ± 1.272	-137.35 ± 3.766	-4.59 ± 0.167	1.22	71.1

Table 5. Thermal stability of open and circularized form of **T35** determined by UV melting method

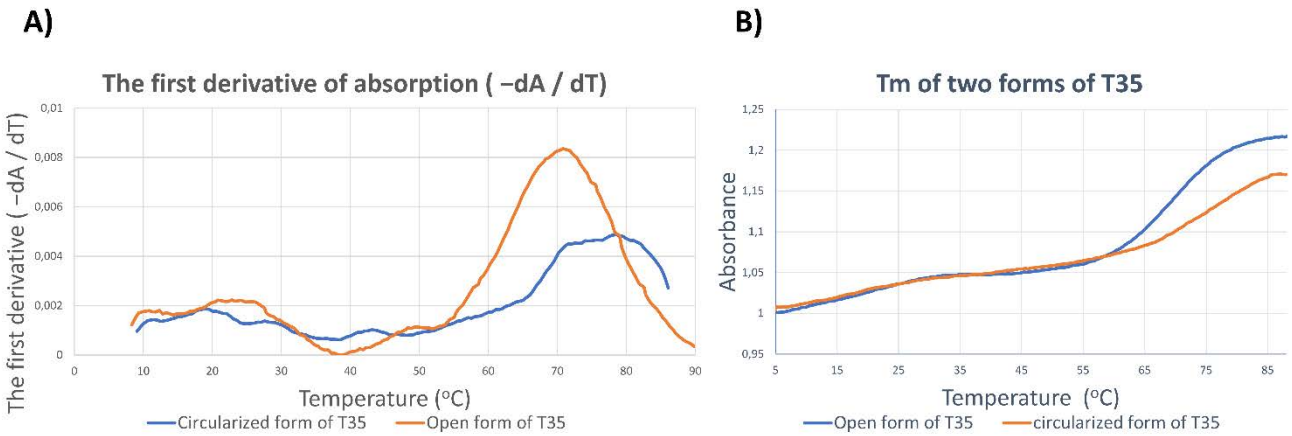


Figure 28. Physicochemical evaluation of open and circularized form of **T35**. A) thermal denaturation monitored by UV spectroscopy. The first derivative of absorption (−dA / dT) was plotted against temperature. B) The normalized UV melting curves of the open (orange) and circular (blue) form of the **T35** RNA oligomer.

6.3.3.2. CD spectra measurements

Circular dichroism (CD) spectra were recorded to compare the secondary structure of the native RNA sequence, open R06 construct containing the 3D linker, and its circularized form. The CD spectra of the native and open R06 oligomers were nearly identical, indicating that the introduction of a non-nucleosidic linker did not alter the overall RNA conformation. In contrast, circularized R06 exhibited slight spectral changes, including a decrease in the molar ellipticity values in the 220-230 nm, 250–270 nm regions and a minor shift in the characteristic bands (Figure 29). These observations suggest subtle conformational rearrangements upon circularization, while the overall RNA secondary structure remains preserved. The observed decrease in ellipticity can be attributed to restricted conformational dynamics and reduced flexibility of the backbone introduced by the circularization process. Such changes are consistent with the formation of a more rigid and compact molecule.

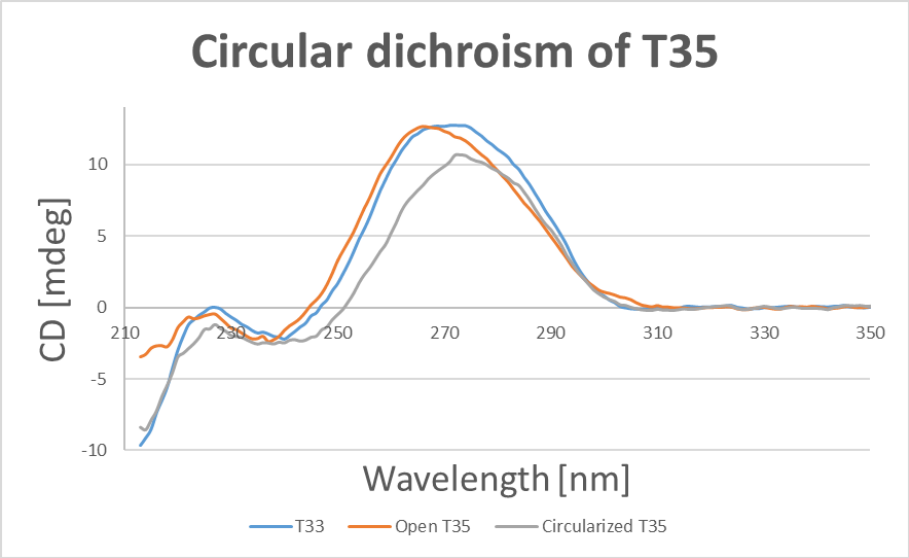


Figure 29. Physicochemical evaluation of unmodified and modified R06 oligonucleotide using circular dichroism. Three forms were measured: **T33** (blue), open form of **T35** (orange) and circularized form of **T35** (grey).

6.3.4. Structural characterization of circularized hairpin with linker S18

Despite many attempts, none of the aromatic-linker-modified constructs (R06, TAR, or SRL) yielded high-quality diffraction crystals. However, the reference construct, a sarcin–ricin loop (SRL) hairpin stabilized with the linear Spacer-18 linker, produced suitable crystals for X-ray diffraction experiments. This construct, unlike its aromatic counterparts, crystallized reproducibly and provided interpretable diffraction data, enabling direct insights into how a flexible non-nucleotide linker influences the architecture of an RNA hairpin prepared for crystallographic studies. Therefore, the obtained structure serves as a reference model for assessing whether linear linkers can support native-like RNA folding while maintaining the closed architecture necessary to prevent duplex formation during crystallization.

6.3.4.1. Crystallization of hairpin T42

To obtain RNA suitable for crystallographic analysis, a hairpin construct based on the sarcin–ricin loop motif was designed with the sequence 5'-AGGACCGGAGCG–S18–UGCUCCUAGUACGAG-3' **T42**, in which the Spacer-18 linker was positioned at the base of the stem (Figure 34). Prior to crystallization, the RNA was circularized using T4 RNA ligase to ensure a topologically constrained and structurally homogeneous sample. In the engineered SRL sequence, the 5'-terminal uridine was removed to facilitate optimal positioning and geometric accommodation of the linker between the two guanosines at the 5’ and 3’ ends.

Crystallization of the SRL hairpin stabilized with the Spacer-18 linker was performed using the sitting-drop vapor-diffusion method at 19 °C. The crystals appeared reproducibly under conditions containing 0.02 M Magnesium chloride hexahydrate, 0.05 M Sodium cacodylate trihydrate pH 7.0, 15% v/v 2-Propanol, 0.001 M Hexammine cobalt(III) chloride, 0.001 M Spermine. Drops were prepared by mixing equal volumes of RNA solution with crystallization buffer and equilibrated against the reservoir. These conditions yielded well-formed crystals that were suitable for subsequent data collection and structural analysis.

6.3.4.2. Data processing and structure solving

3600 diffraction patterns for the SRL hairpin stabilized with the Spacer-18 linker were collected using a wavelength of 91.7 Å and temperature of 100 K. The crystals belonged to the C121 space group with unit-cell parameters of a = 106.15 Å, b = 25.54 Å, c = 25.54 Å. The data were processed using XDS. The final dataset was extended to 1.27 Å resolution with an overall completeness of 96.6%, and an average ⟨I/σ(I)⟩ of 7.45. The merging statistics (R_merge = 10.4%, CC1/2 = 99.7) indicated good internal consistency of the measurements and were suitable for determination of x-ray structure (Table 6).

Data Collection	
X-ray source	Dectris PILATUS 6M-F,
Space group	C121
Cell parameters (Å)	a = 106.15, b = 25.54, c = 25.046
Resolution (Å)	26.53-1.27 (1.35-1.27)
R_{merge}	0.104 (0.937)
I/σ	7.45 (1.36)
$CC_{1/2}$	0.997 (0.515)
Completeness (%)	96.6 (92.3)
Redundancy	3.37 (3.3)
Number of unique reflections	33264
Refinement	
Software	Refmac 5.8.0158
Number of reflections: work/test	16419/1049
Overall mean B value (Å^2)	19.798
R_{work}/R_{free}	0.14988/0.19825
RNA atoms	746
Water molecules	102
RMSD in bonds (Å)	0.015
RMSD in angles (°)	2.146

Table 6. Summary of X-ray data and model refinement statistics of **T42** oligonucleotide.

Analysis of the asymmetric unit content using Matthews coefficient (225) indicated the presence of one RNA duplex per asymmetric unit, corresponding to an estimated solvent content of 57.4%. Molecular replacement was performed using the native sarcin–ricin loop structure (PDB ID: 1Q9A) (234) as the search model. The structure was solved using the *PHASER* program in the CCP4 suite. The final refinement statistics are summarized in Table 6.

6.3.4.3. Structural analysis

The X-ray RNA model exhibited the characteristic architecture of a sarcin–ricin loop–based hairpin (Figure 30). The molecule consists of a well-defined double-helical stem consisting of five canonical Watson–Crick base pairs and three non-canonical pairs (A•G, A•C, and U•C) that contribute to subtle local asymmetry while preserving the global A-form topology. The apical loop of the hairpin consists of a conserved GAGA tetraloop, a motif known for its structural stability in RNA molecules (234–236). The stem contains a characteristic 9U–18A pair engaged in Hoogsteen-type hydrogen bonding, further stabilized by an interaction with a neighboring 8G guanosine, a feature also observed in the native SRL architecture (PDB: 1Q9A), indicating the conservation of this structural motif (Figure 30C). At the 3′ terminus of the hairpin, three unpaired nucleotides participate in crystal-packing interactions. The 1G residue is looped out due to the presence of the introduced spacer-18 linker, which results in its displacement relative to the native conformation. The linker is covalently attached via phosphodiester bonds to guanosine residues flanking the loop at the 5′ and 3′ ends of the construct, and its position is only partially ordered in the electron-density map. Moreover in crystal structure are present 9 of cobaltamine ions interacting with all guanosine residues through bases of nucleotides or their phosphonate groups (Figure 30B).

The superposition of the native (PDB ID: 1Q9A) and modified models of SRL showed high similarity, with an RMSD of **0.8 Å** calculated for the 5′-**CUC CUA GUA CGA GAG GAC CGG AGU**-3′ fragment (Figure 31). The helical parameters remained consistent with those expected for the A-form geometry, including preserved base pair inclinations, helical rise and twist. No significant deviations from the reference architecture were observed, indicating that the introduced modification did not perturb the global SRL motif conformation.

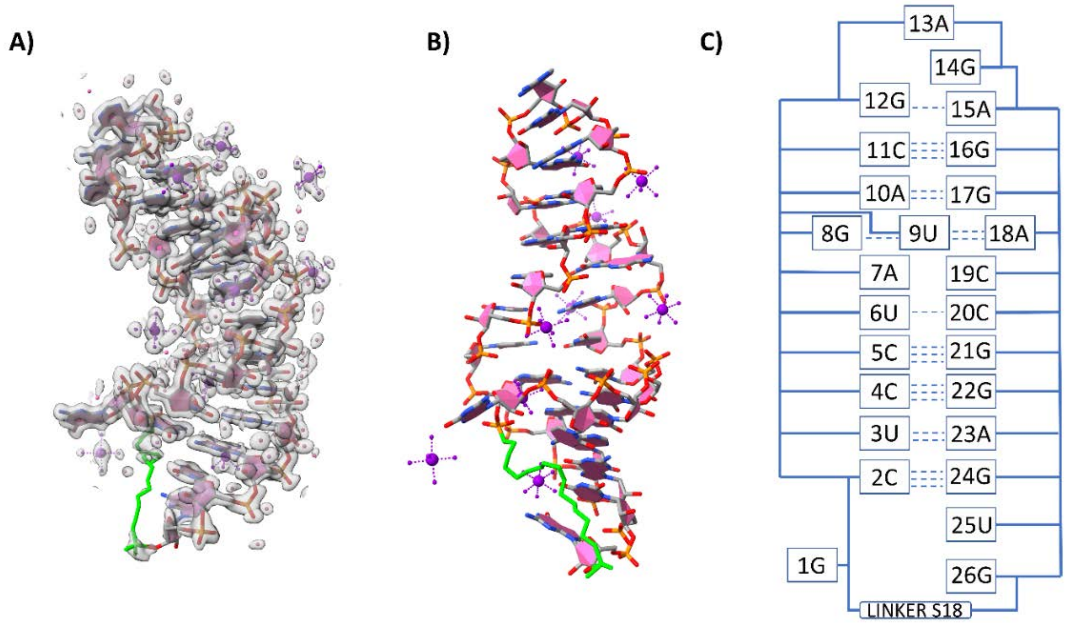


Figure 30. X-ray model of modified Sarcin-ricin loop containing linker spacer 18 as a covalent linkage of 3' and 5' ends. A) The Fo – Fc omit map of **T42** model contoured at the σ 1.0 level. B) Tertiary structure model of **T42** C) Secondary structure diagram.

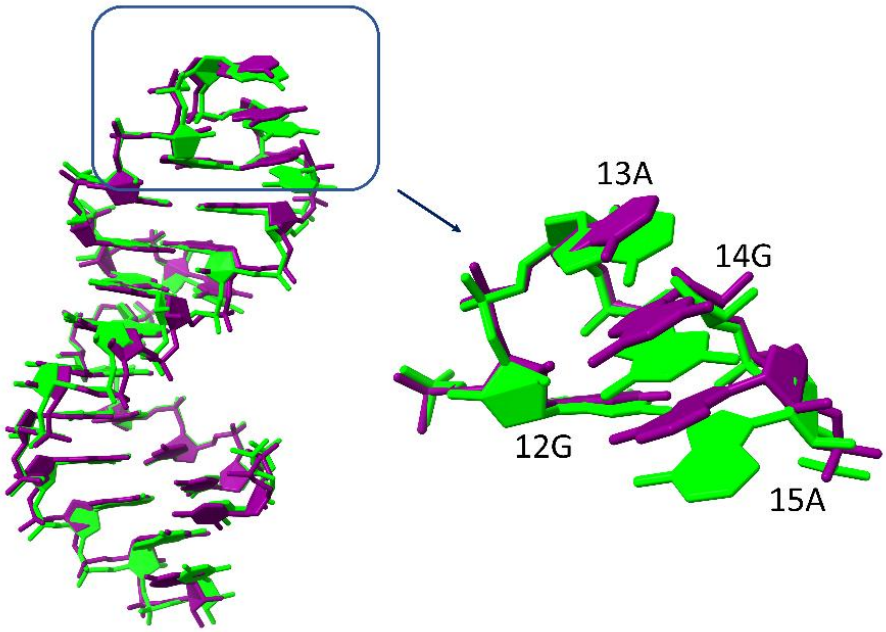


Figure 31. Superposition of **T42** crystal structure (green) and 1Q9A model (purple).

T42

Base pairs	Shift	Slide	Rise	Tilt	Roll	Twist	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CU/AG	-0.26	-1.50	3.25	-3.13	6.33	32.21	-3.68	-0.06	2.92	11.25	5.57	32.96
2 UC/GA	0.47	-1.49	3.29	-0.85	5.46	34.34	-3.30	-0.91	3.02	9.18	1.42	34.76
3 CC/GG	-0.11	-2.21	3.46	-0.14	6.98	29.63	-5.58	0.18	2.88	13.41	0.27	30.43
4 CU/CG	-0.95	-1.39	3.54	9.37	9.04	53.21	-2.07	1.61	3.09	9.92	-10.29	54.66
5 UU/AC	0.82	-0.65	6.44	8.94	0.20	31.22	-1.22	1.69	6.41	0.35	-16.19	32.44
6 UA/GA	4.98	-1.41	3.48	-1.74	2.46	-10.42	2.60	22.90	4.46	-13.18	-9.35	-10.85
7 AC/GG	0.05	-1.08	3.31	-2.97	1.00	62.64	-1.08	-0.18	3.29	0.96	2.85	62.71
8 CG/AG	-3.94	-1.78	2.71	-5.86	-8.22	55.14	-1.46	3.88	3.29	-8.79	6.27	55.99
ave.	0.13	-1.44	3.69	0.45	2.91	36.00	-1.98	3.64	3.67	2.89	-2.43	36.64

Base pairs	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	0.18	-0.22	-0.12	8.95	-12.23	-2.93
2 U-A	0.01	-0.13	-0.15	8.96	-10.06	-0.13
3 C-G	0.25	-0.14	-0.09	6.66	-12.99	1.74
4 C-G	0.27	-0.14	-0.03	-1.11	-4.84	0.14
5 U-C	5.99	-2.32	-0.60	-2.42	-10.12	-13.88
6 U-A	4.12	-1.85	-0.64	7.25	-16.49	-101.89
7 A-G	-6.85	-4.31	-0.16	-1.22	-2.23	-4.10
8 C-G	-0.03	-0.11	0.05	3.40	-1.81	-0.65
9 G-A	7.80	-5.40	-0.02	26.06	9.94	-47.88
ave.	1.30	-1.62	-0.20	6.28	-6.76	-18.84

	Strand I								Strand II						
base	α	β	γ	δ	ε	ζ	χ	base	α	β	γ	δ	ε	ζ	χ
1 C	-78.2	-158.6	51.2	81.1	-150.2	-68.7	-161.6	1 G	-65.1	170.5	59.0	78.7	-157.6	-63.5	-165.2
2 U	-70.2	179.7	49.9	78.1	-147.1	-73.0	-159.4	2 A	-69.9	176.5	51.2	79.1	-150.7	-70.0	-167.2
3 C	-66.2	166.9	58.8	80.8	-150.6	-73.5	-166.7	3 G	-69.2	175.1	59.6	76.8	-152.5	-74.3	-172.1
4 C	-65.8	175.8	57.3	77.7	-145.3	-66.8	-162.6	4 G	-71.5	-175.4	52.7	76.6	-153.9	-68.1	-171.5
5 U	-58.8	170.8	49.6	80.5	-157.2	60.8	-149.0	5 C	-72.6	167.9	56.8	77.5	-141.5	-51.0	-168.3
6 U	-74.8	157.9	42.6	90.8	-139.8	-66.5	-175.5	6 A	-104.5	77.6	173.2	82.9	-146.1	-74.4	-178.2
7 A	-74.2	178.3	52.3	80.2	-145.1	-53.1	-172.7	7 G	-69.8	170.5	55.7	74.9	-155.8	-76.0	-168.2
8 C	-74.7	177.2	52.9	80.1	-147.1	-62.0	-168.9	8 G	142.9	-122.9	179.6	86.7	-128.1	-56.9	-179.0
9 G	-71.1	174.6	56.1	74.3	-122.1	-66.8	-166.6	9 A	-69.2	179.8	50.8	79.1	-147.6	-61.8	-156.6

1Q9A

Base pairs	Shift	Slide	Rise	Tilt	Roll	Twist	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 GC/GU	-0.54	-1.22	3.29	-4.27	3.56	44.61	-1.92	0.33	3.23	4.66	5.59	44.93
2 CU/AG	0.42	-2.42	2.84	3.11	5.29	23.12	-7.22	-0.21	2.27	12.90	-7.59	23.91
3 UC/GA	0.85	-1.46	3.17	1.07	5.13	38.25	-2.81	-1.16	2.98	7.78	-1.62	38.59
4 CC/GG	-0.71	-1.93	3.34	0.00	8.18	32.01	-4.71	1.25	2.78	14.54	-0.01	33.02
5 CU/CG	-0.53	-1.01	3.49	7.64	8.66	54.50	-1.60	1.03	3.21	9.33	-8.23	55.62
6 UU/AC	0.64	-1.03	6.28	5.56	-3.01	32.04	-0.82	0.74	6.36	-5.40	-9.95	32.64
7 UA/GA	5.03	-1.24	3.58	-1.62	-1.26	-10.94	8.67	23.02	4.11	6.52	-8.40	-11.13
8 AC/GG	-0.02	-1.13	3.13	-2.61	3.89	59.57	-1.32	-0.11	3.06	3.91	2.63	59.73
9 CG/AG	-2.69	-1.29	3.02	-6.88	6.58	49.98	-1.94	2.68	3.15	7.70	8.05	50.82
ave.	0.27	-1.42	3.57	0.22	4.11	35.90	-1.52	3.06	3.46	6.89	-2.17	36.46

Base pairs	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 G-U	-2.51	-0.76	0.29	-0.17	-15.91	-0.54
2 C-G	0.39	-0.04	0.36	1.42	-10.43	-0.20
3 U-A	-0.39	-0.18	0.04	6.50	-13.93	-7.00
4 C-G	0.29	-0.13	0.04	5.55	-17.36	1.89
5 C-G	0.30	-0.15	0.00	-2.57	-12.51	-1.96
6 U-C	5.88	-2.20	-0.40	-3.86	-14.60	-13.09
7 U-A	4.14	-1.87	-0.71	6.72	-18.94	-102.88
8 A-G	-6.80	-4.31	-0.19	-2.39	4.17	-2.82
9 C-G	0.17	-0.10	-0.19	7.46	-1.95	0.47
10 G-A	7.22	-5.23	0.72	18.21	-5.01	-19.00
ave.	0.87	-1.50	-0.00	3.69	-10.65	-14.51

	Strand I								Strand II						
base	α	β	γ	δ	ε	ζ	χ	base	α	β	γ	δ	ε	ζ	χ
1 G	-80.0	150.8	59.1	88.2	-147.4	-68.3	-169.1	1 U	-62.5	171.4	48.2	84.6	-157.7	-71.3	-157.2
2 C	-67.7	175.5	47.2	80.4	-162.6	-65.6	-157.6	2 G	-79.3	-173.8	61.8	74.2	-151.8	-61.9	-165.9
3 U	143.6	-168.8	-170.4	81.3	-131.2	-71.9	-171.8	3 A	-66.0	169.6	51.3	82.4	-168.7	-67.3	-162.7
4 C	-64.8	164.2	57.2	79.0	-156.7	-70.9	-161.6	4 G	-63.6	168.6	60.7	81.2	-146.0	-79.4	-169.2
5 C	-66.7	179.4	52.4	81.4	-148.8	-59.2	-157.1	5 G	-69.0	174.6	52.8	76.5	-148.8	-69.9	-167.6
6 U	-68.3	172.3	52.4	85.9	-101.7	-59.0	-147.5	6 C	-64.5	163.6	53.5	84.9	-142.6	-58.0	-165.0
7 U	-68.8	151.9	39.3	89.4	-144.4	-63.6	-177.3	7 A	-106.1	76.7	166.3	81.9	-150.3	-74.7	178.8
8 A	-76.2	178.7	53.1	80.5	-147.8	-53.7	-166.9	8 G	-64.7	172.8	57.1	79.3	-158.9	-79.0	-164.7
9 C	-70.9	173.3	53.2	79.4	-148.1	-63.3	-163.1	9 G	144.5	-136.8	179.3	84.7	-138.1	-59.6	-173.3
10 G	-73.2	175.6	52.2	77.9	-131.0	-64.3	-169.2	10 A	-71.6	178.6	50.0	82.9	-139.7	-68.4	-150.2

Table 7. Helical parameters calculated using 3DNA, based on C1'- C1' vectors for **T42** oligonucleotide and crystal structure of SRL (pdb code: 1Q9A).

The backbone torsion angles (α, β, γ, δ, ε, and ζ) remained highly comparable to those observed in the reference SRL structure, with only minor fluctuations within the expected range for A-form RNA. Likewise, the glycosidic torsion angles (χ) of all nucleosides adopted *anti*-conformations consistent with the canonical SRL fold (Table 7) instead of approximately 180° (237). The helical parameters, including helical rise, twist, inclination, and x-displacement, align closely with those characteristic of the A-form geometry.

Subtle deviations from the native SRL topology were observed within the GAGA tetraloop, primarily involving the nucleotides 13A and 14G. In the native SRL, the 13A residue forms a stabilizing interaction with the extruded uridine at the 5′ end of the hairpin. In **T42**, the uridine was removed to accommodate the Spacer-18 linker; therefore, the 13A residue adopts a slightly altered orientation and establishes new intermolecular contacts within the crystal lattice. Specifically, 13A forms a stacking interaction with the symmetry-related 1G′′ residue, which is flipped out from the stem. In addition, 13A engages in hydrogen bonding with 20C′ residue from a neighboring symmetry-related molecule, involving the exocyclic amino group of 13A and O2 carbonyl of 20C′.′ The second tetraloop nucleotide, 14G, also participates in a distinct interaction that is absent in the native SRL structure. Whereas in the wild-type SRL, 14G interacts exclusively with water molecules, in the modified construct, it forms a nearly canonical base-pair-like contact with 19C′ from a symmetry-related RNA molecule. This interaction involves hydrogen bonding between the exocyclic amine of 14G and the O2 carbonyl of 19C′ and between the N1-H of 14G and the N3 atom of 19C′ residues (Figure 32). Together, these newly formed lattice-stabilized contacts explain the minor structural rearrangements observed in the GAGA loop and reflect the influence of linker-induced displacement of terminal nucleotides on the local interaction landscape.

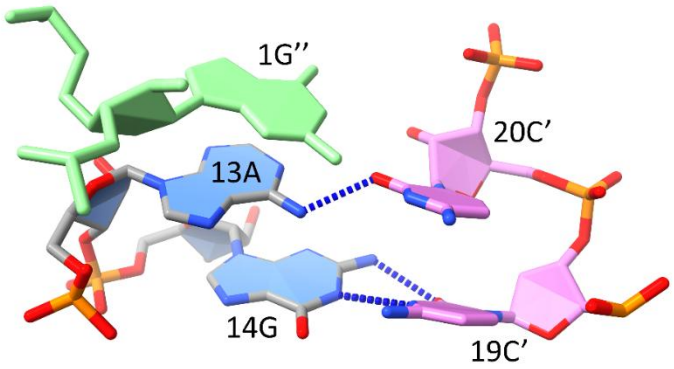


Figure 32. Apical loop GAGA in **T42** oligonucleotide crystal structure.

In addition to the alterations observed within the loop region, the modified SRL hairpin displayed a distinct set of interactions involving terminal guanosines of the stem. The 26G nucleotide engages in a stacking interaction with its symmetry-related counterpart, 26G'. Furthermore, 26G forms a hydrogen bond with one of the nine cobaltamine molecules in the asymmetric unit. The amino group of cobaltamine donates a hydrogen bond to the O6 carbonyl of 26G, while an analogous interaction is formed between the same cobaltamine molecule and the O6 atom of residue 24G, thereby stabilizing both the guanosines.

Hexaamminecobalt(III) $[Co(NH_3)_6]^{3+}$ is a highly charged, rigid octahedral complex whose ammine ligands establish extensive second-sphere hydrogen bonds of N–H···O6 and N–H···N7 with guanine bases and interactions with the phosphate backbone. These geometries are suited to the electron-rich surface of guanine, explaining the selectivity of cobaltamine for G-rich RNA regions (238–240). In the SRL hairpin, every guanosine residue participates in such interactions, resulting in the local stabilization of the RNA architecture. The nine cobaltamine molecules per asymmetric unit mediate contacts between the symmetry-related RNA molecules, reinforcing the lattice stability. Beyond intramolecular stabilization, these interactions promote intermolecular contacts within the crystal lattice, thereby facilitating crystal packing and formation (Figure 33).

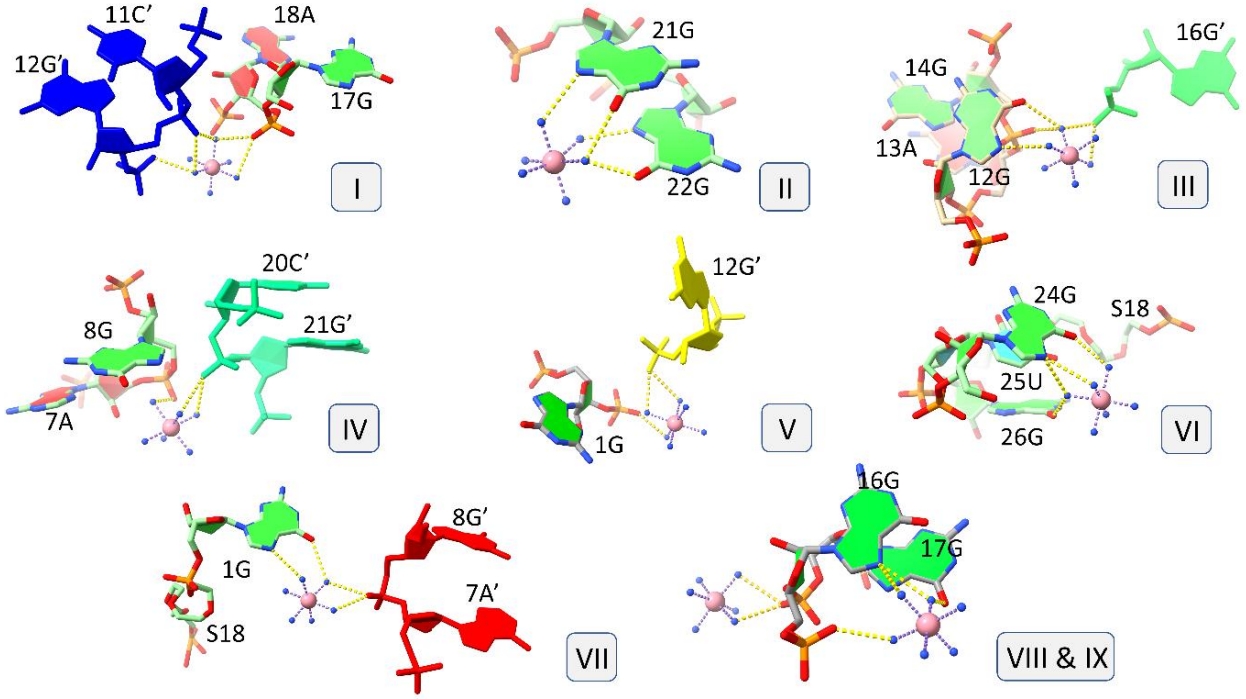


Figure 33. Interaction of cobaltamine ions with guanosine residues.

The hexaethylene glycol linker (Spacer 18) introduced into the modified SRL construct is an aliphatic –CH_2–O–CH_2–based connector characterized by a high conformational flexibility. Its sp^3-hybridized carbon atoms allow for the adoption of a broad range of torsional angles. This conformational plasticity enables it to adjust its geometry to the spatial arrangement of the RNA stem. Within the crystal structure, the observable distance between the 5′ and 3′ phosphates bridged by the linker was 16.13 Å. However, its flexibility resulted in poorly defined electron density for several of its segments (Figure 34). Only a single direct interaction with RNA was observed: a hydrogen bond between the linker O3 atom and the N3 atom of the residue 25U. Additionally, the linker (O6 atom) forms a hydrogen bond with the water molecule (Figure 34).

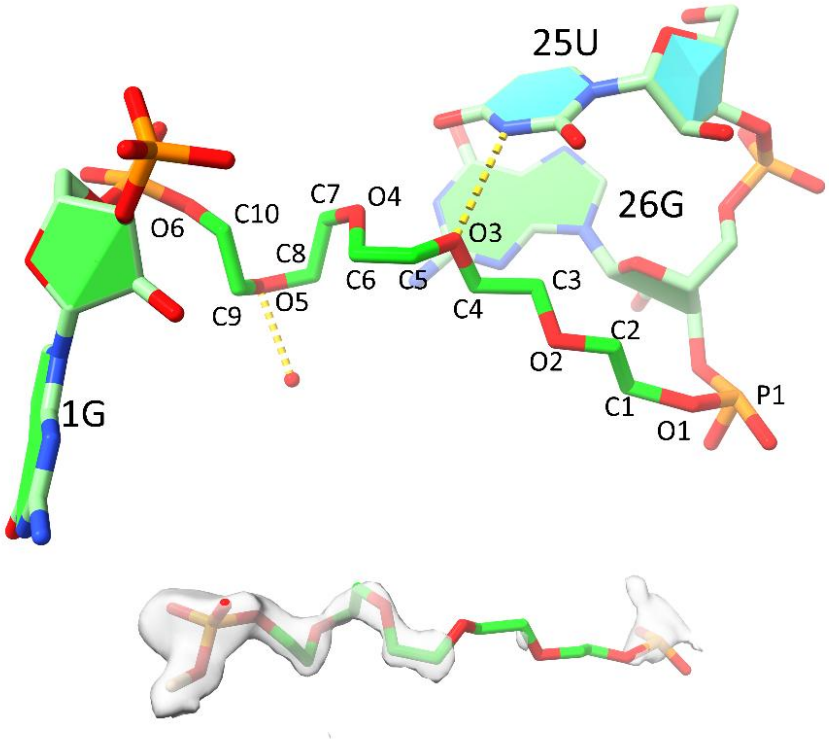


Figure 34. Interaction of linker spacer 18 with uridine residue and water molecule. The Fo – Fc omit map of linker contoured at the σ 1.0 level.

Analysis of the B-factor parameters further supports the intrinsic flexibility of the Spacer-18 linker. The B-factors ranged from 44 to 100 Å², increasing progressively along the linker as it extended away from residue 26G toward 1G. The portion of the linker that is supported by interpretable electron density—from atom C10 to C5—exhibits comparatively moderate B-factors of 43.94–52.54 Å², consistent with partial structural ordering near the RNA stem. Beyond this segment, the electron density becomes either discontinuous or absent, which correlates with a significant increase in B-factors, from 63.62 Å² for atom O3 to 101.68 Å² for atom O1 (Table 8). This gradient of increasing disorder aligns with the chemical nature of the linker and its lack of aromatic or other functional groups, which could form either stacking interactions with nucleobases or H-bonds with RNA or water molecules.

Atom	**C10**	**C9**	**O5**	**C8**	**C7**	**O4**	**C6**	**C5**
B-factor (Å²)	43.94	46.86	53.59	45.86	54.8	53.78	45.31	52.54
Atom	**O3**	**C4**	**C3**	**O2**	**C2**	**C1**	**O1**	**P1**
B-factor (Å²)	63.62	62.93	60.05	64.12	78.73	86.87	101.68	100.89

Table 8. B-factors of linker spacer 18 atoms.

6.3.5. Initial Biochemical Studies

Stabilized RNA hairpins were further evaluated for their biological activity. To assess whether the introduced modifications affected biological function, I performed in vitro Dicer cleavage assays using appropriate constructs of pre-miR-21 according to the protocol by Kurzyńska-Kokorniak, A. *et al.* (233). The goal of these experiments was to determine if the presence of non-nucleosidic linkers or additional recognition motifs interferes with the enzyme's ability to recognize and cleave the hairpin structure into mature miRNA products. Four variants derived from the pre-miR-21 hairpin were designed and synthesized. These included the native pre-miR-21 sequence (**T38**) (Figure 35A), a construct containing the non-nucleosidic linker spacer 18 (**T39**) (Figure 35B), two modified versions incorporating the U1A protein-binding site (**T41**) (Figure 35C), and another carrying distinct recognition motifs: one with the NCD ligand-binding loop (UGGAA) (**T40**) (Figure 35D). **T41** contains the U1A protein-binding site, originating from a stabilization strategy developed in a complementary project on RNA hairpin stabilization. This approach is described in greater detail in the doctoral dissertation of Dr. Marcin Ryczek (title: *Badania strukturalne RNA o znaczeniu w patogenezie chorób*

neurodegeneracyjnych). **T40** carries the UGGAA loop, which functions as a binding motif for the small-molecule ligand, NCD. The introduction of the UGGAA loop represents the second method explored in this doctoral study, which aims to achieve ligand-mediated stabilization of RNA structural motifs (publication: *Naphthyridine carbamate dimer ligand induces formation of Z-RNA-like fold of disease-related RNA and exhibits a molecular glue characteristics in crystal lattice formation* (159)). These constructs served as platforms for assessing the influence of loop modifications and linker insertion on the folding and potential ligand or protein interactions of pre-miR-21.

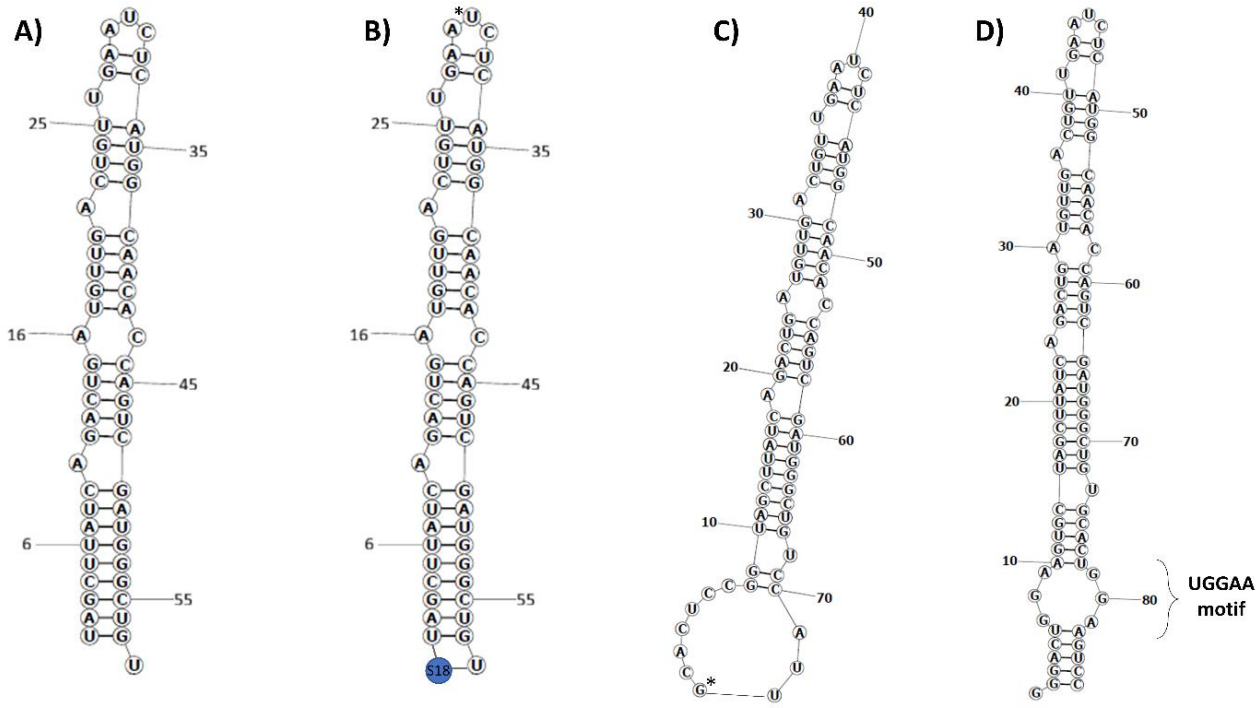


Figure 35. Constructs of engineered pre-miR-21 hairpins. A) Unmodified, B) Pre-miR-21 consisting of aliphatic linker spacer 18 (S18) as a covalent linkage between the 3' and 5' ends. Linker is assigned in blue circle. C) Hairpin consisting of the U1A protein-binding site, and D) Pre-miR-21 with UGGAA pentad motif. Ligation position is assigned with [*].

Dicer cleavage assays were performed using the purified human Dicer enzyme kindly provided by Prof. Anna Kurzyńska-Korkorniak from Institute of Bioorganic Chemistry PAN, and the four pre-miR-21 constructs labeled with the ^{32}P isotope at the 5' end described above. Each RNA was incubated under standard reaction conditions, and the

resulting cleavage products were analyzed on 20% denaturing polyacrylamide gel (Figure 36).

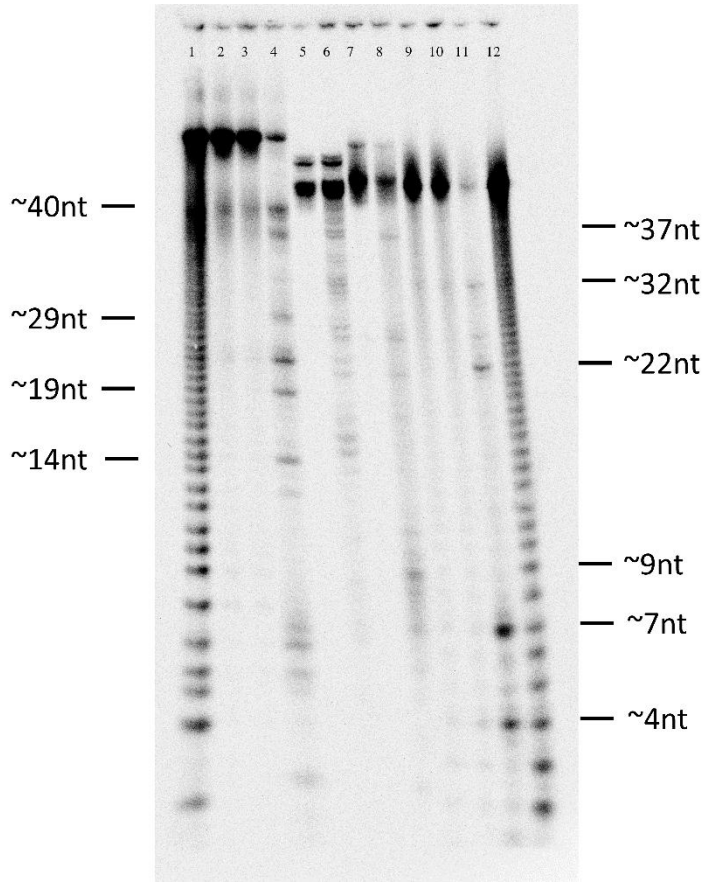


Figure 36. RNase activity of Dicer toward four pre-miR-21 constructs: unmodified hairpin, aliphatic linker (Spacer 18) variant, U1A hairpin construct, and UGGAA-motif construct. Lanes: (1) formamide hydrolysate ladder; (2) H_2O; (3) buffer only; (4) Dicer + buffer; (5) U1A construct + buffer; (6) U1A construct + Dicer; (7) Spacer-18 construct + buffer; (8) Spacer-18 construct + Dicer; (9) pre-miR-21 in H_2O; (10) pre-miR-21 + buffer; (11) pre-miR-21 + Dicer; (12) formamide hydrolysate ladder.

Native pre-miR-21 was efficiently processed into fragments corresponding to the expected lengths of mature miRNA and its complementary strand. The results showed that Dicer generated several cleavage products of approximately ~32, ~22, ~7, and ~4 nucleotides in length (Figure 36). Notably, the efficient production of the ~22 nt fragment—the characteristic length of mature microRNA–was clearly detected, confirming that the **T38** construct was properly recognized and processed along the canonical Dicer cleavage pathway.

For the pre-miR-21 construct **T39** containing the Spacer-18 linker and circularized via ligation of the 5′ end (radiolabeled with ^{32}**P**) at the apical loop, Dicer processing produced a distinct cleavage pattern, yielding fragments of approximately ~9 nt, ~24 nt, and ~37 nt in length

(Figure 37). The appearance of these products suggests that, despite the circular topology and presence of a flexible non-nucleotidic spacer, Dicer was still able to recognize the lower stem region and initiate cleavage at positions consistent with its known preference for measuring ~22 nt from a free or accessible terminus. The ~24 nt fragment likely corresponds to a shifted cleavage site, whereas the ~9 nt and ~37 nt products represent complementary paired fragments generated from asymmetric or sequential cleavage events within structured RNA (Figure 37). This pattern indicates that the circular construct remained accessible to Dicer. Other cleavage products characteristic of canonical Dicer processing may not be visible due to the localization of the **^{32}P** radiolabel within the apical loop, which allows the detection of only fragments containing labeled nucleotide residue.

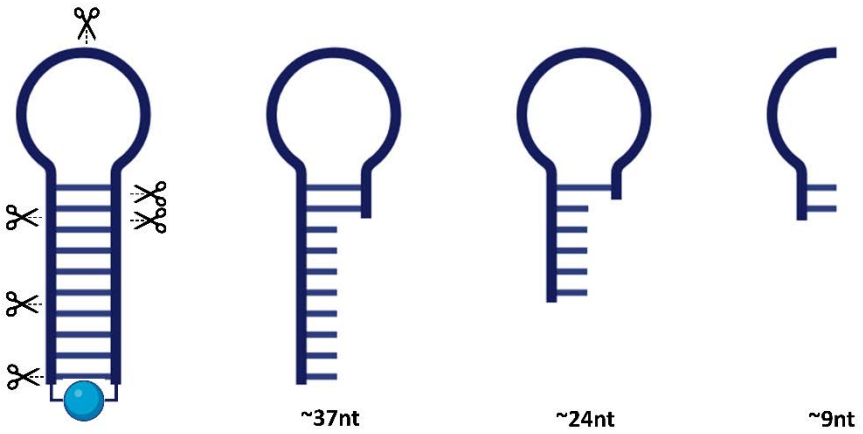


Figure 37. Predicted Dicer cleavage pattern of the circularized pre-miR-21 containing the Spacer-18 linker.

For the pre-miR-21 T41 construct containing the U1A protein-binding site, RNA was circularized to adopt a characteristic dumbbell-like architecture. The autoradiogram showed four major visible products of Dicer cleavage: fragments of approximately ~14 nt, ~21 nt, ~33 nt and ~37nt (Figure 38). The appearance of multiple poorly defined bands (a diffuse “ladder-like” pattern) suggests that Dicer cleavage in this construct proceeded in a largely nonspecific manner, likely due to the altered topology or increased rigidity introduced by circularization combined with a structured U1A-binding loop. Nevertheless, despite its closed dumbbell architecture, Dicer was still able to process and cut out RNA fragments.

The ~14 nt product corresponds to the isolated U1A-binding loop itself, consistent with the radiolabel (^{32}**P**) being positioned within this region (Figure 38). The ~33 nt fragment likely represents cleavage within both arms of the stem, retaining the radiolabeled loop and one extended segment of the hairpin. The ~21 nt product— not a canonical Dicer-generated length, likely arises from an alternative or shifted cleavage event, reflecting Dicer's struggle to establish proper distance measurement on a circular, topologically constrained substrate. Because the radioactive label is located exclusively within the U1A-binding loop, additional complementary fragments produced during processing remain undetected in this assay.

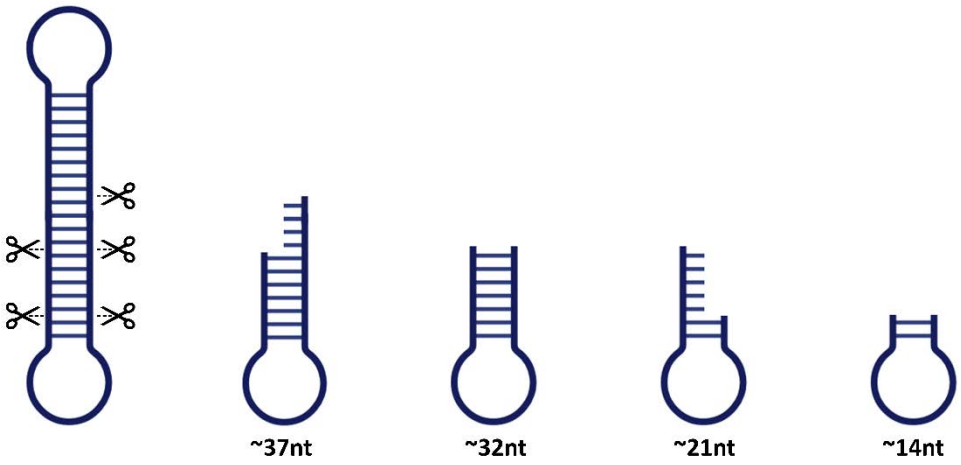


Figure 38. Predicted Dicer cleavage pattern of the dumbbell form pre-miR-21 containing the U1A binding loop.

For the pre-miR-21 **T40** construct containing the UGGAA loop, the Dicer cleavage pattern closely resembled the characteristic processing profile of native pre-miR-21. Prominent bands appeared at approximately ~14 nt, ~22 nt, ~29 nt, and ~40 nt, corresponding to the known Dicer cleavage pattern of the pre-miR-21 stem, including fragments derived from sequences **5'-UAGCUUAUC-3'**, **5'-GACUG-3'**, and **5'-UGUUG-3'**. These positions were consistent with the canonical cut established for pre-miR-21 processing, indicating that the introduction of the UGGAA loop maintained the recognition ability of Dicer for the structured stem of **T40**. Additionally, a band at ~19 nt was detected, which likely reflects nonspecific cleavage occurring within the double-stranded region, suggesting occasional off-register cuts but without affecting the overall processing efficiency.

The results of the Dicer cleavage assay demonstrated that, despite the introduced modifications, such as circularization with the Spacer-18 linker, circularization containing the U1A protein-binding site, and the UGGAA loop serving as a ligand-binding motif, all constructs were successfully processed by the Dicer enzyme. This indicates that the modified pre-miR-21 constructs retained their biological activity.

6.4. Ligand-mediated stabilization method induced by NCD molecule

In contrast to the covalent stabilization strategies explored in this study, the ligand-based approach demonstrated an alternative and effective mechanism for promoting RNA crystallization. The interaction of NCD with the UGGAA motif illustrates how ligand binding can promote local conformational rearrangements, including the induction of Z-RNA–like geometry. Moreover, the ligand acts as a molecular bridge, generating specific intermolecular contacts between symmetry-related RNA molecules that facilitate the formation of an ordered crystal lattice. This dual action, local structural stabilization combined with lattice-bridging interactions, highlights why small-molecule ligands can be a tool for the crystallization of RNA molecules. The methods and results are described in detail in the publication: *Naphthyridine carbamate dimer ligand induces formation of Z-RNA-like fold of disease-related RNA and exhibits a molecular glue characteristics in crystal lattice formation* (159).

7. Summary of publications included in the doctoral dissertation

7.1. Overview of Methods for Large-Scale RNA Synthesis

Marcin Ryczek, Martyna Pluta, Leszek Błaszczyk*, Agnieszka Kiliszek*

*Applied Sciences **2022**, 12, 1543.*

IF_{2022} = 2.7, $MNiSW_{2022}$ = 40

The publication *entitled "Overview of Methods for Large-Scale RNA Synthesis"* provides an essential methodological framework for the chemical synthesis of RNA constructs intended for high-resolution crystallographic analysis. The strategies reported are relevant to the synthetic approaches applied in this dissertation, particularly in the context of incorporating covalently bonded non-nucleosidic linkers and preparing RNA constructs of sufficient quantity and purity for crystallization. Although the systems investigated in the publication differ functionally from the linker-modified RNAs described in this thesis, both rely on rigorous chemical synthesis methods that enable precise structural manipulation and production of milligram-scale RNA samples suitable for structural biology.

This publication demonstrates the utility of solid-phase phosphoramidite chemistry as the primary platform for generating highly defined RNA oligonucleotides. Chemical synthesis is essential for introducing modifications that cannot be achieved through enzymatic routes, such as the incorporation of aromatic linkers bridging the 5' and 3' termini of RNA in the present work. The flexibility of solid-phase synthesis allows the controlled incorporation of modified monomers, linker phosphoramidites, and sequence alterations required to accommodate structural elements, such as Spacer-18 or pyromellitic and naphthalene aromatic linkers.

A critical strength highlighted in the publication is the ability to perform syntheses on a large scale (up to multi-milligram quantities). Crystallographic studies typically require significantly more material than biochemical or spectroscopic assays, often in the range of several milligrams of highly purified RNA. This methodology emphasizes the optimization of synthetic

cycles, reagent volumes, and coupling efficiencies to ensure high-quality RNA suitable for demanding applications.

Both approaches also use DMT-on and DMT-off purification strategies, which are chosen according to the length, hydrophobicity, and overall chemical properties of the synthesized RNA. The DMT-on approach, in which the 5’-dimethoxytrityl protecting group is retained during the initial purification step, is the most reliable for RNA constructs lacking hydrophobic modifications. This limitation arises because both the DMT group and hydrophobic linkers display a strong affinity for the reverse-phase media used in Glen-Pak cartridges, which can lead to retention issues, poor elution, or loss of product. Consequently, for aromatic-linker–modified RNAs, synthesis must be completed using the DMT-off strategy. In this case, the 5’-DMT group is removed from the solid support, followed by standard deprotection and cleavage from the resin, after which purification is performed by denaturing PAGE. This approach avoids any interaction of the hydrophobic linker with reverse-phase matrices and ensures the reliable isolation of full-length RNA suitable for crystallographic applications.

The publication further underscores the importance of controlled deprotection and cleavage conditions, as RNA is considerably more labile than DNA. The removal of protecting groups, particularly 2’-O-silyl protecting groups, requires mild conditions that minimize strand scission while preserving chemically delicate modifications. This observation is directly relevant to the present work, as certain aromatic linkers are sensitive to deprotection conditions, and purification steps can result in the cleavage of the linker from the RNA backbone. Insights from this publication include strategies for minimizing linker dissociation, including adjustments to deprotection times, temperature control, and careful selection of purification conditions to reduce exposure to harsh reagents. Furthermore, this publication highlights the necessity of multiple purification steps followed by denaturing polyacrylamide gel electrophoresis (PAGE) to ensure homogeneity required for crystallization.

7.2. Naphthyridine carbamate dimer ligand induces formation of Z-RNA-like fold of disease-related RNA and exhibits a molecular glue characteristics in crystal lattice formation

Martyna Mateja-Pluta, Leszek Błaszczyk, Magdalena Bejger, Kazuhiko Nakatani, Agnieszka Kiliszek*

Nucleic Acids Research, **2025**, 53, gkaf924

IF_{2024} = 13.1, $MNiSW_{2025}$ = 200

The publication entitled "*Naphthyridine carbamate dimer ligand induces formation of Z-RNA-like fold of disease-related RNA and exhibits a molecular glue characteristics in crystal lattice formation*" (included as part of this doctoral dissertation) presents a structural analysis of the complex formed between the small-molecule ligand NCD and RNA containing the UGGAA motif. This work is appended to the dissertation because it provides a detailed description of how NCD interacts with RNA, elucidating both the mode of binding and the mechanism by which the ligand stabilizes and orders the RNA structure. The crystal structure revealed that two NCD molecules bind directly to the UGGAA motif, each bridging two guanosine residues on the same RNA strand. A comparison with the structure of unliganded RNA demonstrated that the GGA region does not engage in stabilizing hydrogen bonds in its native form but instead forms only stacking interactions and occasional contacts with cobaltamine ions. These observations suggest that the GGA residues are likely highly flexible and only weakly pre-organized in solution. However, upon NCD binding, both guanosine residues engage in specific hydrogen-bonding interactions with the ligand, leading to pronounced stabilization of the UGGAA motif. This ligand-induced structural rearrangement resulted in increased thermal stability, as confirmed by differential scanning calorimetry (DSC) and UV melting experiments.

The structural rearrangement promoted by NCD resulted in the formation of a Z-RNA–like geometry, a structural state that is unfavorable in the absence of a ligand but becomes stabilized through the introduction of precise hydrogen-bonding and stacking interactions. This highlights a key principle: ligands can impose local order in otherwise dynamic or weakly structured RNA regions, thereby enabling the capture of conformations that would be difficult or impossible to observe without chemical assistance.

In addition to inducing localized structural stabilization, NCD exerts a second important effect. It acts as a molecular glue that promotes interactions within the crystal lattice. The planar heteroaromatic system of the ligand is suitable for mediating interactions between symmetry-related RNA molecules, effectively bridging the individual strands within the crystal lattice. These intermolecular contacts, absent in unliganded RNA, enable the assembly of an ordered three-dimensional lattice that supports high-resolution diffraction. The dual functionality of NCD, which combines local conformational stabilization with lattice-bridging capacity, is a particularly valuable insight for RNA crystallography, where obtaining well-ordered crystals is often the limiting step.

A comparison of ligand-mediated stabilization with the covalent strategies employed in this dissertation underscores the complementary nature of these approaches. In contrast, non-nucleosidic linkers enforce global architectural constraints by covalently bridging the RNA backbone. Ligands such as NCD exert their effects locally stabilizing specific motifs without restricting the global conformational landscape. Moreover, ligands enhance intermolecular interactions within the crystal lattice, a feature not inherently achievable by covalent linkers.

8. Discussion

The structural characterization of RNA is constrained by the intrinsic properties of RNA molecules, such as high flexibility, formation of alternative 3D structures and ease at structural rearrangements induced by experimental conditions (12, 26). Although X-ray crystallography continues to provide the most reliable high-resolution insights into RNA architecture, crystallization of RNA remains a major barrier, particularly for motifs such as hairpins that readily unfold and form thermodynamically favored duplexes at elevated RNA or salt concentrations typically used in crystallization. As a result, native RNA structures, despite their biological importance, often cannot be captured in a homogenous crystalline environment without additional stabilization strategies (41–43).

In this context, the present study explored chemical strategies to mitigate these limitations. A key premise of the project was to identify a linker capable of enhancing structural stability without disrupting the native RNA architecture. These approaches include the incorporation of non-nucleosidic linkers, circularization of RNA molecules, and ligand-mediated stabilization. These results provide insights into the impact of these modifications on RNA stability and folding behavior, as well as their overall potential to overcome crystallization challenges.

In this study, a series of non-nucleotidic linkers were designed to covalently connect the 5’ and 3’ termini of RNA strands. These compounds are inexpensive to obtain, and their synthesis is straightforward and high-yielding, making them practical tools for routine RNA modification. Among the proposed candidates, the strongest stabilizing effects were observed for two compounds: a naphthalene-derived linker containing three methyl groups on each side of the aromatic core (3d) and a pyromellitic derivative featuring aliphatic side chains with two methyl groups on each arm (3a). Although these modifications effectively thermodynamically stabilize RNA hairpins and duplexes, their incorporation requires a careful design of RNA constructs in terms of their secondary and tertiary structure. Developed aromatic linkers possess stiff aromatic interface which might require the specific geometry of neighboring residues and formation of canonical base pairing. Otherwise the linker can induce the local and global changes in secondary and three-dimensional structure. This was observed for low efficiency of ligation reactions performed on different RNA models. Among all tested constructs, only the **T35** hairpin containing the **3d** linker underwent successful circularization

by T4 RNA ligase which also require the proper geometry of apical loop. The lack of circular product suggest that the single stranded ligation site could be structurally rearranged and became inaccessible for enzymatic reaction. The example of SRL–Spacer-18 construct **T42** showed that the incorporation of the non-nucleosidic linker required sequence modification, specifically the removal of the 5' uridine of the native SRL, to allow proper geometric accommodation of the linker at the base of the stem. Despite this alteration and the presence of the linker, the global fold of the SRL hairpin remained preserved. However, it induced a distinct pattern of crystal lattice interactions compared to native SRL structures. In the unmodified SRL constructs (PDB IDs: 1Q9A and 3DVZ), crystal contacts are limited to stacking interactions between the loop adenosine in the GAGA tetraloop and the flipped-out 5'-terminal uridine, as well as interactions with solvent molecules or components of the crystallization buffer (Figure 39). In contrast, the modified SRL hairpin engages in additional and more extensive contacts with nucleobases from symmetry-related molecules, reflecting the altered molecular environment produced by the linker (Figure 32). The modification also resulted in a different space group: while native SRL crystals adopted $P4_3$ (54, 234), the current structure crystallized in C121 (Table 6), indicating that the linker reshaped the intermolecular organization within the lattice. The newly introduced interactions appear to support the lattice packing and may facilitate crystal formation. Moreover, the flexible Spacer-18 linker, due to its conformational lability, can adapt to the RNA fold without restricting the packing of nucleotides, thereby enabling the RNA to accommodate alternative crystal packing arrangements. Such adaption for aromatic linkers can be difficult to occur due to rigidity of aromatic compounds and tendency for π-π stacking interactions.

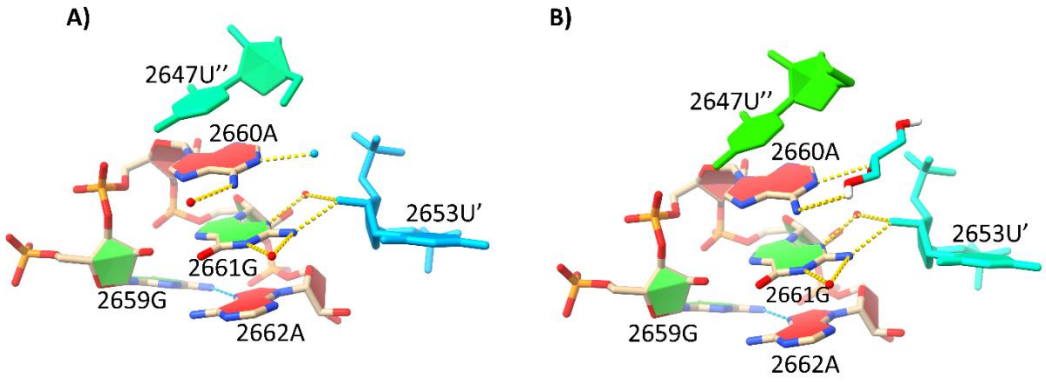


Figure 39. Interaction in crystal lattice with GAGA loop in two crystal structures pdb code: A) 1Q9A, B) 3DVZ

Several studies reported in the literature have successfully employed rigid aromatic linkers in structural investigations; however, these achievements have been confined to DNA-based systems. For example, Lewis et al. (103) demonstrated that stilbene-modified DNA hairpins could be efficiently stabilized and structurally characterized, benefiting from the inherent rigidity and predictable helical parameters of B-form DNA. Similarly, DNA assemblies incorporating aromatic chromophores have been resolved at the atomic level by co-crystallization with protein partners, such as the recombinant binding domain of the restriction enzyme BpuJI, which serves as a crystallization scaffold (107). Aromatic foldamer units such as 8-aminomethyl-2-quinolinecarboxylic acid or 8-amino-2-quinolinecarboxylic acid dimers have also been successfully integrated into B-DNA duplexes and hairpins (124), further illustrating that DNA can readily tolerate such non-natural moieties. All of these examples involve DNA, protein–DNA complexes, or racemic DNA mixtures, which are structurally more regular, less prone to alternative secondary structures, and more compatible with rigid aromatic scaffolds. In contrast, RNA adopts more diverse and dynamic secondary and tertiary structures, and crystallization typically requires a greater degree of conformational adaptability. This distinction suggests that the structural requirements for the successful incorporation of aromatic linkers are more restrictive for RNA than for DNA, offering a likely explanation for the limited crystallization success observed in the present study.

Although the introduction of non-nucleosidic linkers significantly increased the thermodynamic stability of the modified RNA constructs, as demonstrated by UV melting and CD spectra, these gains did not directly translate into successful crystallization. However, the Spacer-18–modified SRL hairpin and UGGAA repeat RNA in complex with the NCD ligand both yielded ordered crystalline lattices (159), despite exhibiting different stabilization profiles. The flexible aliphatic linker in Spacer-18 provided global stabilization while preserving the conformational adaptability required for crystal packing, whereas NCD did not globally rigidify the RNA but instead locally stabilized UGGAA loop, enabling the ligand to act as a structural organizer and promote lattice formation. These observations indicate that crystallization success depends not only on stabilizing the RNA fold but also on achieving a balance between structural rigidity and the capacity to engage in specific intermolecular interactions required for ordered packing of the RNA.

The interaction between NCD and the UGGAA motif illustrates how small-molecule ligands can stabilize RNA through mechanisms that are fundamentally different from those of covalent linkers. NCD induces a local conformational rearrangement within the loop, promoting the formation of a non-canonical Z-RNA–like geometry that would otherwise be thermodynamically unfavorable. Simultaneously, the ligand acts as a *molecular glue*, creating specific intermolecular contacts between symmetry-related RNA molecules that facilitate the assembly of an ordered crystal lattice. This dual role, local structural stabilization combined with lattice-bridging interactions, highlights why ligands such as NCD can be effective in enabling the crystallization of RNA motifs (*vide in publication: Naphthyridine carbamate dimer ligand induces formation of Z-RNA-like fold of disease-related RNA and exhibits a molecular glue characteristics in crystal lattice formation* (159)).

Beyond their application in crystallographic studies, the linkers developed in this work possess considerable potential for therapeutic and diagnostic RNA technologies, where structural stability in the cellular environment is more critical than the ability to crystallize. Although the aromatic linkers did not facilitate RNA crystallization, their ability to enhance thermodynamic stability and enforce defined secondary structures suggests that they may be highly valuable in biochemical or therapeutic contexts (241, 242). Biochemical assays using human Dicer demonstrated that the introduced modifications did not compromise the biological functions of the RNA constructs. Despite the incorporation of the Spacer-18 linker into the pre-miR-21 hairpin, introduction of the U1A-binding loop, or replacement of the native apical loop with the UGGAA motif for ligand recognition, all engineered pre-miR-21 variants were efficiently processed by Dicer. The enzyme correctly recognized the double-stranded stem region and generated cleavage products, indicating that the overall secondary structure features required for Dicer activity were preserved.

Chemical modifications of nucleotides, whether at the sugar moiety, nucleobase, or phosphate backbone, have long been explored to improve RNA stability, bioavailability, and resistance to nucleases, contributing significantly to advances in antisense and RNA-based therapeutics (243–245). However, the synthesis of nucleoside-modified oligonucleotides is technically challenging and often costly owing to their high polarity, multiple ionizable groups, and incompatibility with standard organic chemistry conditions (246–248). Non-nucleotidic modifications, such as the linkers examined in this study, are attractive alternatives. Their

modular chemical architecture allows extensive structural tuning without the constraints imposed by canonical nucleotide chemistry, opening new avenues for functional nucleic acid design. In contrast to known phenanthrene and pyrene derivatives, presented linkers Aromatic linkers, including phenanthrene- and pyrene-based systems, have been shown to act as “molecular prostheses” capable of substituting for missing bases in damaged DNA, thereby restoring the duplex character (114–118). This demonstrates that well-designed non-nucleotidic moieties can function as structural surrogates, supporting the integrity of nucleic acid frameworks. Thus, the RNA linkers developed in this study may find future applications in stabilizing therapeutic RNAs, protecting susceptible regions from degradation, and modulating folding landscapes in a controlled manner (119, 120). Such uses highlight the broader relevance of these chemical tools, extending well beyond crystallography to the field of RNA-based medicine.

9. Conclusions

The results presented in this doctoral dissertation allow for the formulation of the following conclusions:

1. **Aromatic linkers based on naphthalene and pyromellitic derivatives constitute a synthetically accessible class of RNA-stabilizing modification.** Their preparation is cost-efficient, provides high yields, and does not require specialized instrumentation, making them attractive candidates for broader applications in nucleic acid chemistry.
2. Aromatic linkers possess a rigid planar interface that necessitates a specific geometry of neighboring nucleotides and the preservation of canonical base pairing to accommodate their steric and electronic properties. Therefore, their incorporation can have local or global impacts on RNA flexibility, base pairing, and geometry.
3. Successful crystallization requires balancing RNA structural rigidity with the ability to form specific intermolecular contacts that are crucial for ordered packing. **The flexible Spacer-18 linker was more compatible with the crystallization of the Sarcin-Racin RNA hairpin.** Its conformational lability allows it to adapt to the natural RNA topology, without restricting nucleotide packing. This adaptability enables the formation of alternative crystal packing arrangements and was essential for solving the SRL–Spacer-18 structure.
4. Ligand-mediated stabilization, exemplified by the interaction between NCD and guanosine residues of RNA, represents an effective alternative strategy. Unlike covalent linkers, NCD does not rigidify the entire RNA structure but instead induces localized stabilization within the UGGAA loop while simultaneously acting as a molecular glue between symmetry-related molecules.
5. **Despite their limitations in crystallographic applications, aromatic linkers have significant potential in RNA-based therapies and diagnostics.** Their ability to enhance stability, resist degradation, and modulate local RNA structures suggests that they may serve as valuable stabilizing agents in medicinal chemistry, antisense technologies, and nucleic acid–based drug design.

10. References

1. Bose,R., Saleem,I. and Mustoe,A.M. (2024) Causes, functions, and therapeutic possibilities of RNA secondary structure ensembles and alternative states. *Cell Chemical Biology*, **31**, 17–35.

2. Leitão,A.L. and Enguita,F.J. (2025) The Unpaved Road of Non-Coding RNA Structure-Function Relationships: Current Knowledge, Available Methodologies and Future Trends. 10.20944/preprints202501.1068.v1.

3. Lin,J.-C., Hyeon,C. and Thirumalai,D. (2012) RNA under Tension: Folding Landscapes, Kinetic Partitioning Mechanism, and Molecular Tensegrity. *J. Phys. Chem. Lett.*, **3**, 3616–3625.

4. Mattick,J.S. (2011) The central role of RNA in human development and cognition. *FEBS Letters*, **585**, 1600–1616.

5. Goodall,G.J. and Wickramasinghe,V.O. (2021) RNA in cancer. *Nat Rev Cancer*, **21**, 22–36.

6. Spitale,R.C. and Incarnato,D. (2023) Probing the dynamic RNA structurome and its functions. *Nat Rev Genet*, **24**, 178–196.

7. Connelly,C.M., Moon,M.H. and Schneekloth,J.S. (2016) The Emerging Role of RNA as a Therapeutic Target for Small Molecules. *Cell Chemical Biology*, **23**, 1077–1090.

8. Cooper,T.A., Wan,L. and Dreyfuss,G. (2009) RNA and Disease. *Cell*, **136**, 777–793.

9. Mandal,M. and Breaker,R.R. (2004) Gene regulation by riboswitches. *Nat Rev Mol Cell Biol*, **5**, 451–463.

10. Breaker,R.R. (2008) Complex Riboswitches. *Science*, **319**, 1795–1797.

11. Serganov,A. and Patel,D.J. (2007) Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat Rev Genet*, **8**, 776–790.

12. Ke,A. (2004) Crystallization of RNA and RNA?protein complexes. *Methods*, **34**, 408–414.

13. Batey,R.T., Rambo,R.P. and Doudna,J.A. (1999) Tertiary Motifs in RNA Structure and Folding. *Angew. Chem. Int. Ed.*, **38**, 2326–2343.

14. Higgs,P.G. (2000) RNA secondary structure: physical and computational aspects. *Quart. Rev. Biophys.*, **33**, 199–253.

15. Simons,R.W. and Grunberg-Manago,M. (1998) RNA structure and function Cold Spring Harbor Laboratory press, Cold Spring Harbor.

16. Kim,S.H., Suddath,F.L., Quigley,G.J., McPherson,A., Sussman,J.L., Wang,A.H.J., Seeman,N.C. and Rich,A. (1974) Three-Dimensional Tertiary Structure of Yeast Phenylalanine Transfer RNA. *Science*, **185**, 435–440.

17. Dunham,C.M., Selmer,M., Phelps,S.S., Kelley,A.C., Suzuki,T., Joseph,S. and Ramakrishnan,V. (2007) Structures of tRNAs with an expanded anticodon loop in the decoding center of the 30S ribosomal subunit. *RNA*, **13**, 817–823.

18. Agirrezabala,X. and Valle,M. (2015) Structural Insights into tRNA Dynamics on the Ribosome. *IJMS*, **16**, 9866–9895.

19. Pley,H.W., Flaherty,K.M. and McKay,D.B. (1994) Three-dimensional structure of a hammerhead ribozyme. *Nature*, **372**, 68–74.

20. Scott,W.G., Finch,J.T. and Klug,A. (1995) The crystal structure of an AII-RNAhammerhead ribozyme: A proposed mechanism for RNA catalytic cleavage. *Cell*, **81**, 991–1002.

21. Anderson,M., Schultz,E.P., Martick,M. and Scott,W.G. (2013) Active-Site Monovalent Cations Revealed in a 1.55-Å-Resolution Hammerhead Ribozyme Structure. *Journal of Molecular Biology*, **425**, 3790–3798.

22. Zhan,X., Wilson,T.J., Li,Z., Zhang,J., Yang,Y., Lilley,D.M.J. and Liu,Y. (2024) The structure and catalytic mechanism of a pseudoknot-containing hammerhead ribozyme. *Nat Commun*, **15**, 6628.

23. Cate,J.H., Gooding,A.R., Podell,E., Zhou,K., Golden,B.L., Kundrot,C.E., Cech,T.R. and Doudna,J.A. (1996) Crystal Structure of a Group I Ribozyme Domain: Principles of RNA Packing. *Science*, **273**, 1678–1685.

24. Ferré-D'Amaré,A.R., Zhou,K. and Doudna,J.A. (1998) Crystal structure of a hepatitis delta virus ribozyme. *Nature*, **395**, 567–574.

25. Tinoco,I. and Bustamante,C. (1999) How RNA folds. *Journal of Molecular Biology*, **293**, 271–281.

26. Pujari,N., Saundh,S.L., Acquah,F.A., Mooers,B.H.M., Ferré-D'Amaré,A.R. and Leung,A.K.-W. (2021) Engineering Crystal Packing in RNA Structures I: Past and Future Strategies for Engineering RNA Packing in Crystals. *Crystals*, **11**, 952.

27. Schneider,B., Sweeney,B.A., Bateman,A., Cerny,J., Zok,T. and Szachniuk,M. (2023) When will RNA get its AlphaFold moment? *Nucleic Acids Research*, **51**, 9522–9532.

28. Zhang,S., Li,J. and Chen,S.-J. (2024) Machine learning in RNA structure prediction: Advances and challenges. *Biophysical Journal*, **123**, 2647–2657.

29. Jackson,R.W., Smathers,C.M. and Robart,A.R. (2023) General Strategies for RNA X-ray Crystallography. *Molecules*, **28**, 2111.

30. Ganser,L.R., Kelly,M.L., Herschlag,D. and Al-Hashimi,H.M. (2019) The roles of structural dynamics in the cellular functions of RNAs. *Nat Rev Mol Cell Biol*, **20**, 474–489.

31. Russell,R. (2008) RNA misfolding and the action of chaperones. *Front Biosci*, **13**, 1.

32. Ferré-D'Amaré,A.R. (2010) Use of the spliceosomal protein U1A to facilitate crystallization and structure determination of complex RNAs. *Methods*, **52**, 159–167.

33. Rosenbach,H., Victor,J., Borggräfe,J., Biehl,R., Steger,G., Etzkorn,M. and Span,I. (2020) Expanding crystallization tools for nucleic acid complexes using U1A protein variants. *Journal of Structural Biology*, **210**, 107480.

34. Ferré-D'Amaré,A.R., Zhou,K. and Doudna,J.A. (1998) A general module for RNA crystallization. *Journal of Molecular Biology*, **279**, 621–631.

35. Bonilla,S.L. and Jang,K. (2024) Challenges, advances, and opportunities in RNA structural biology by Cryo-EM. *Current Opinion in Structural Biology*, **88**, 102894.

36. Benseler,F., Fu,D.J., Ludwig,J. and McLaughlin,L.W. (1993) Hammerhead-like molecules containing non-nucleoside linkers are active RNA catalysts. *J. Am. Chem. Soc.*, **115**, 8483–8484.

37. Thomson,J.B., Tuschl,T. and Eckstein,F. (1993) Activity of hammerhead ribozymes containing nonnucleotidic linkers. *Nucl Acids Res*, **21**, 5600–5603.

38. Adams,P.D., Baker,D., Brunger,A.T., Das,R., DiMaio,F., Read,R.J., Richardson,D.C., Richardson,J.S. and Terwilliger,T.C. (2013) Advances, Interactions, and Future Developments in the CNS, Phenix, and Rosetta Structural Biology Software Systems. *Annu. Rev. Biophys.*, **42**, 265–287.

39. Conev,A., Rigo,M.M., Devaurs,D., Fonseca,A.F., Kalavadwala,H., De Freitas,M.V., Clementi,C., Zanatta,G., Antunes,D.A. and Kavraki,L.E. (2023) EnGens: a computational framework for generation and analysis of representative protein conformational ensembles. *Briefings in Bioinformatics*, **24**, bbad242.

40. Leman,J.K., Weitzner,B.D., Lewis,S.M., Adolf-Bryfogle,J., Alam,N., Alford,R.F., Aprahamian,M., Baker,D., Barlow,K.A., Barth,P., *et al.* (2020) Macromolecular modeling and design in Rosetta: recent methods and frameworks. *Nat Methods*, **17**, 665–680.

41. Holbrook,S.R., Cheong,C., Tinoco,I. and Kim,S.-H. (1991) Crystal structure of an RNA double helix incorporating a track of non-Watson–Crick base pairs. *Nature*, **353**, 579–581.

42. Rypniewski,W., Adamiak,D.A., Milecki,J. and Adamiak,R.W. (2008) Noncanonical G(*syn*)–G(*anti*) base pairs stabilized by sulphate anions in two X-ray structures of the (GUGGUCUGAUGAGGCC) RNA duplex. *RNA*, **14**, 1845–1851.

43. Bernacchi,S., Ennifar,E., Tóth,K., Walter,P., Langowski,J. and Dumas,P. (2005) Mechanism of Hairpin-Duplex Conversion for the HIV-1 Dimerization Initiation Site. *Journal of Biological Chemistry*, **280**, 40112–40121.

44. Le,B.T., Chen,S. and Veedu,R.N. (2023) Evaluation of Chemically Modified Nucleic Acid Analogues for Splice Switching Application. *ACS Omega*, **8**, 48650–48661.

45. Khvorova,A. and Watts,J.K. (2017) The chemical evolution of oligonucleotide therapies of clinical utility. *Nat Biotechnol*, **35**, 238–248.

46. Chernikov,I.V., Ponomareva,U.A. and Chernolovskaya,E.L. (2023) Structural Modifications of siRNA Improve Its Performance In Vivo. *IJMS*, **24**, 956.

47. Harp,J.M., Guenther,D.C., Bisbe,A., Perkins,L., Matsuda,S., Bommineni,G.R., Zlatev,I., Foster,D.J., Taneja,N., Charisse,K., *et al.* (2018) Structural basis for the synergy of 4′- and 2′-modifications on siRNA nuclease resistance, thermal stability and RNAi activity. *Nucleic Acids Research*, **46**, 8090–8104.

48. Lu,G., Shivalila,C., Monian,P., Yu,H., Harding,I., Briem,S., Byrne,M., Faraone,A., Friend,S., Huth,O., *et al.* (2024) Rational design of base, sugar and backbone modifications improves ADAR-mediated RNA editing. *Nucleic Acids Research*, **52**, 10068–10084.

49. Hassler,M.R., Turanov,A.A., Alterman,J.F., Haraszti,R.A., Coles,A.H., Osborn,M.F., Echeverria,D., Nikan,M., Salomon,W.E., Roux,L., *et al.* (2018) Comparison of partially and fully chemically-modified siRNA in conjugate-mediated delivery in vivo. *Nucleic Acids Research*, **46**, 2185–2196.

50. Kornienko,I.V., Aramova,O.Yu., Tishchenko,A.A., Rudoy,D.V. and Chikindas,M.L. (2024) RNA Stability: A Review of the Role of Structural Features and Environmental Conditions. *Molecules*, **29**, 5978.

51. Kierzek,E., Pasternak,A., Pasternak,K., Gdaniec,Z., Yildirim,I., Turner,D.H. and Kierzek,R. (2009) Contributions of Stacking, Preorganization, and Hydrogen Bonding to the Thermodynamic Stability of Duplexes between RNA and 2′- *O* -Methyl RNA with Locked Nucleic Acids. *Biochemistry*, **48**, 4377–4387.

52. Jiang,J., Sheng,J., Carrasco,N. and Huang,Z. (2006) Selenium derivatization of nucleic acids for crystallography. *Nucleic Acids Research*, **35**, 477–485.

53. Siegmund,V., Santner,T., Micura,R. and Marx,A. (2011) Enzymatic synthesis of 2′-methylseleno-modified RNA. *Chem. Sci.*, **2**, 2224.

54. Olieric,V., Rieder,U., Lang,K., Serganov,A., Schulze-Briese,C., Micura,R., Dumas,P. and Ennifar,E. (2009) A fast selenium derivatization strategy for crystallization and phasing of RNA structures. *RNA*, **15**, 707–715.

55. Sheng,J. and Huang,Z. (2008) Selenium Derivatization of Nucleic Acids for Phase and Structure Determination in Nucleic Acid X-ray Crystallography. *IJMS*, **9**, 258–271.

56. Salon,J., Jiang,J., Sheng,J., Gerlits,O.O. and Huang,Z. (2008) Derivatization of DNAs with selenium at 6-position of guanine for function and crystal structure studies. *Nucleic Acids Research*, **36**, 7009–7018.

57. Hostetter,A.A., Chapman,E.G. and DeRose,V.J. (2009) Rapid Cross-Linking of an RNA Internal Loop by the Anticancer Drug Cisplatin. *J. Am. Chem. Soc.*, **131**, 9250–9257.

58. Pallan,P.S., Von Matt,P., Wilds,C.J., Altmann,K.-H. and Egli,M. (2006) RNA-Binding Affinities and Crystal Structure of Oligonucleotides Containing Five-Atom Amide-Based Backbone Structures′. *Biochemistry*, **45**, 8048–8057.

59. Zhang,D. and Paukstelis,P.J. (2016) Enhancing DNA Crystal Durability through Chemical Crosslinking. *ChemBioChem*, **17**, 1163–1170.

60. Gyssels,E., De Laet,N., Lumley,E. and Madder,A. (2016) Interstrand Cross-Linking of Nucleic Acids: From History to Recent and Future Applications. In Jurga,S., Erdmann,V.A., Barciszewski,J. (eds), *Modified Nucleic Acids in Biology and Medicine*, RNA Technologies. Springer International Publishing, Cham, pp. 339–369.

61. Flamme,M., McKenzie,L.K., Sarac,I. and Hollenstein,M. (2019) Chemical methods for the modification of RNA. *Methods*, **161**, 64–82.

62. Bartosik,K., Debiec,K., Czarnecka,A., Sochacka,E. and Leszczynska,G. (2020) Synthesis of Nucleobase-Modified RNA Oligonucleotides by Post-Synthetic Approach. *Molecules*, **25**, 3344.

63. Somoza,Á. (2008) Protecting groups for RNA synthesis: an increasing need for selective preparative methods. *Chem. Soc. Rev.*, **37**, 2668.

64. Hu,W., Wang,Z., Xiao,Y., Zhang,S. and Wang,J. (2019) Advances in crosslinking strategies of biomedical hydrogels. *Biomater. Sci.*, **7**, 843–855.

65. Wahba,M.I. (2024) A comprehensive review on genipin: an efficient natural cross-linker for biopolymers. *Polym. Bull.*, **81**, 14251–14305.

66. Noll,D.M., Mason,T.M. and Miller,P.S. (2006) Formation and Repair of Interstrand Cross-Links in DNA. *Chem. Rev.*, **106**, 277–301.

67. Xu,Y., Suzuki,Y. and Komiyama,M. (2009) Click Chemistry for the Identification of G-Quadruplex Structures: Discovery of a DNA–RNA G-Quadruplex. *Angew Chem Int Ed*, **48**, 3281–3284.

68. Zheng,K., Xiao,S., Liu,J., Zhang,J., Hao,Y. and Tan,Z. (2013) Co-transcriptional formation of DNA:RNA hybrid G-quadruplex and potential function as constitutional cis element for transcription control. *Nucleic Acids Research*, **41**, 5533–5541.

69. Zhang,X., Spiegel,J., Martínez Cuesta,S., Adhikari,S. and Balasubramanian,S. (2021) Chemical profiling of DNA G-quadruplex-interacting proteins in live cells. *Nat. Chem.*, **13**, 626–633.

70. Siddik,Z.H. (2017) Alkylating Agents and Platinum Antitumor Compounds. In Bast,R.C., Hait,W.N., Kufe,D.W., Weichselbaum,R.R., Holland,J.F., Croce,C.M., Piccart-Gebart,M., Wang,H., Hong,W.K., Pollock,R.E. (eds), *Holland-Frei Cancer Medicine*. Wiley, pp. 1–20.

71. Zhang,C., Xu,C., Gao,X. and Yao,Q. (2022) Platinum-based drugs for cancer therapy and anti-tumor strategies. *Theranostics*, **12**, 2115–2132.

72. Chaney,S.G. and Sancar,A. (1996) DNA Repair: Enzymatic Mechanisms and Relevance to Drug Response. *JNCI: Journal of the National Cancer Institute*, **88**, 1346–1360.

73. Murakami,A., Yamamoto,Y., Namba,M., Iwase,R. and Yamaoka,T. (2001) Photo-Cross-Linked Oligonucleotide Duplex as a Decoy-DNA for Inhibition of Restriction Endonuclease Activity. *Bioorganic Chemistry*, **29**, 223–233.

74. Gao,S., Drouin,R. and Holmquist,G.P. (1994) DNA Repair Rates Mapped Along the Human *PGK1* Gene at Nucleotide Resolution. *Science*, **263**, 1438–1440.

75. Gao,H., Yang,M. and Cook,A.F. (1995) Stabilization of double-stranded oligonucleotides using backbone-linked disulfide bridges. *Nucl Acids Res*, **23**, 285–292.

76. Lee,I., Ahn,J., Kim,H., Park,J. and Lee,K. (2003) Advantages of the Circular Dumbbell Decoy in Gene Therapy and Studies of Gene Regulation. *CDT*, **4**, 619–623.

77. Amoah-Apraku,B., Fang,M.-Z. and Guzman,N.J. (2000) A non-nucleotide–bridged DNA decoy inhibits renal epithelial nitric oxide synthase expression. *Kidney International*, **57**, 83–91.

78. Mann,M.J. and Dzau,V.J. (2000) Therapeutic applications of transcription factor decoy oligonucleotides. *J. Clin. Invest.*, **106**, 1071–1075.

79. Matsuyama,Y., Yamayoshi,A., Kobori,A. and Murakami,A. (2014) Functional regulation of RNA-induced silencing complex by photoreactive oligonucleotides. *Bioorganic & Medicinal Chemistry*, **22**, 1003–1007.

80. Mikame,Y., Eshima,H., Toyama,H., Nakao,J., Matsuo,M., Yamamoto,T., Hari,Y., Komano,J.A. and Yamayoshi,A. (2023) Development and Crosslinking Properties of Psoralen-Conjugated Triplex-Forming Oligonucleotides as Antigene Tools Targeting Genome DNA. *ChemMedChem*, **18**, e202300348.

81. Leriche,G., Chisholm,L. and Wagner,A. (2012) Cleavable linkers in chemical biology. *Bioorganic & Medicinal Chemistry*, **20**, 571–582.

82. Durand,M., Chevrie,K., Chassignol,M., Thuong,N.T. and Maurizot,J.C. (1990) Circular dichroism studies of an oligodeoxyribonucleotide containing a hairpin loop made of a hexaethylene glycol chain: conformation and stability. *Nucl Acids Res*, **18**, 6353–6359.

83. Kotikam,V. and Rozners,E. (2020) Amide-Modified RNA: Using Protein Backbone to Modulate Function of Short Interfering RNAs. *Acc. Chem. Res.*, **53**, 1782–1790.

84. Doktycz,M.J., Paner,T.M. and Benight,A.S. (1993) Studies of DNA dumbbells. IV. Preparation and melting of a DNA dumbbell with the 16 base-pair sequence$^{5'}$ G-T-A-T-C-C-C-T-C-T-G-G-A-T-A-C$^{3'}$ linked on the ends by dodecyl chains. *Biopolymers*, **33**, 1765–1777.

85. Zhou,L., Myers,A.N., Vandersteen,J.G., Wang,L. and Wittwer,C.T. (2004) Closed-Tube Genotyping with Unlabeled Oligonucleotide Probes and a Saturating DNA Dye. *Clinical Chemistry*, **50**, 1328–1335.

86. Durand,M., Peloille,S., Thuong,N.T. and Maurizot,J.C. (1992) Triple-helix formation by an oligonucleotide containing one (dA)12 and two (dT)12 sequences bridged by two hexaethylene glycol chains. *Biochemistry*, **31**, 9197–9204.

87. Guzaev,A., Salo,H., Azhayev,A. and Lönnberg,H. (1996) Novel Non-Nucleosidic Building Blocks for the Preparation of Multilabeled Oligonucleotides. *Bioconjugate Chem.*, **7**, 240–248.

88. Salunkhe,M., Wu,T. and Letsinger,R.L. (1992) Control of folding and binding of oligonucleotides by use of a nonnucleotide linker. *J. Am. Chem. Soc.*, **114**, 8768–8772.

89. Rumney,S. and Kool,E.T. (1995) Structural Optimization of Non-Nucleotide Loop Replacements for Duplex and Triplex DNAs. *J. Am. Chem. Soc.*, **117**, 5635–5646.

90. Pils,W. and Micura,R. (2000) Flexible non-nucleotide linkers as loop replacements in short double helical RNAs. *Nucleic Acids Research*, **28**, 1859–1863.

91. Efimov,V.A., Buryakova,A.A., Kalinkina,A.L., Choob,M.V., Chakhmakhcheva,O.G., Ojwang,J.O. and Rando,R.F. (1998) Convenient Approaches to the Synthesis of Oligonucleotide Macrocycles Containing Non-nucleotide Linkers. *Nucleosides, Nucleotides & Nucleic Acids*, **17**, 379–396.

92. Seitz,O. (1999) Chemically Modified Antisense Oligonucleotides-Recent Improvements of RNA Binding and Ribonuclease H Recruitment. *Angewandte Chemie International Edition*, **38**, 3466–3469.

93. Letsinger,R.L. and Wu,T. (1995) Use of a Stilbenedicarboxamide Bridge in Stabilizing, Monitoring, and Photochemically Altering Folded Conformations of Oligonucleotides. *J. Am. Chem. Soc.*, **117**, 7323–7328.

94. Wu,X. (1998) Synthesis and pairing properties of oligoribonucleotide analogues containing a metal-binding site attached to beta-D-allofuranosyl cytosine. *Nucleic Acids Research*, **26**, 4315–4323.

95. Dempcy,R. (1999) Linkers designed to intercalate the double helix greatly facilitate DNA alkylation by triplex-forming oligonucleotides carrying a cyclopropapyrroloindole reactive moiety. *Nucleic Acids Research*, **27**, 2931–2937.

96. Wiederholt,K. (1999) A 2,2"-bipyridine ligand for incorporation into oligodeoxynucleotides: synthesis, stability and fluorescence properties of ruthenium-DNA complexes. *Nucleic Acids Research*, **27**, 2487–2493.

97. Rajur,S.B., Robles,J., Wiederholt,K., Kuimelis,R.G. and McLaughlin,L.W. (1997) Hoechst 33258 Tethered by a Hexa(ethylene glycol) Linker to the 5‘-Termini of Oligodeoxynucleotide 15-Mers: Duplex Stabilization and Fluorescence Properties. *J. Org. Chem.*, **62**, 523–529.

98. Kiliszek,A., Błaszczyk,L., Kierzek,R. and Rypniewski,W. (2017) Stabilization of RNA hairpins using non-nucleotide linkers and circularization. *Nucleic Acids Research*, **45**, e92–e92.

99. Siegmund,K., Hariharan,M. and Lewis,F.D. (2011) Conformation of a Dodecane DNA Hairpin Linker. Multiple Gauche Bonds Cover the Bases. *J. Phys. Chem. B*, **115**, 3740–3746.

100. Hunter,C.A., Lawson,K.R., Perkins,J. and Urch,C.J. (2001) Aromatic interactions. *J. Chem. Soc., Perkin Trans. 2*, 10.1039/b008495f.

101. Egli,M., Tereshko,V., Mushudov,G.N., Sanishvili,R., Liu,X. and Lewis,F.D. (2003) Face-to-Face and Edge-to-Face π−π Interactions in a Synthetic DNA Hairpin with a Stilbenediether Linker. *J. Am. Chem. Soc.*, **125**, 10842–10849.

102. Lewis,F.D., Liu,X., Miller,S.E., Hayes,R.T. and Wasielewski,M.R. (2002) Dynamics of Electron Injection in DNA Hairpins. *J. Am. Chem. Soc.*, **124**, 11280–11281.

103. Lewis,F.D., Wu,Y. and Liu,X. (2002) Synthesis, Structure, and Photochemistry of Exceptionally Stable Synthetic DNA Hairpins with Stilbene Diether Linkers. *J. Am. Chem. Soc.*, **124**, 12165–12173.

104. Trkulja,I. and Häner,R. (2007) Monomeric and Heterodimeric Triple Helical DNA Mimics. *J. Am. Chem. Soc.*, **129**, 7982–7989.

105. Langenegger,S.M. and Häner,R. (2004) A Simple, Non-Nucleosidic Base Surrogate Increases the Duplex Stability of DNA Containing an Abasic Site. *Chemistry & Biodiversity*, **1**, 259–264.

106. Langenegger,S.M. and Häner,R. (2004) DNA containing phenanthroline- and phenanthrene-derived, non-nucleosidic base surrogates. *Tetrahedron Letters*, **45**, 9273–9276.

107. Probst,M., Aeschimann,W., Chau,T.T.H., Langenegger,S.M., Stocker,A. and Häner,R. (2016) Structural insight into DNA-assembled oligochromophores: crystallographic analysis of pyrene- and phenanthrene-modified DNA in complex with BpuJI endonuclease. *Nucleic Acids Research*, **44**, 7079–7089.

108. Ackermann,D. and Häner,R. (2004) Nonnucleosidic Base Surrogates: The Effect of 1,2-Disubstituted Phenanthrenes on DNA Duplex Stability. *Helvetica Chimica Acta*, **87**, 2790–2804.

109. Stutz,A., Langenegger,S.M. and Häner,R. (2003) Phenanthrene-Derived DNA Hairpin Mimics. *Helvetica Chimica Acta*, **86**, 3156–3163.

110. Zhang,L., Long,H., Schatz,G.C. and Lewis,F.D. (2007) Synthesis and properties of nicked dumbbell and dumbbell DNA conjugates having stilbenedicarboxamide linkers. *Org. Biomol. Chem.*, **5**, 450–456.

111. Lewis,F.D., Liu,X., Wu,Y., Miller,S.E., Wasielewski,M.R., Letsinger,R.L., Sanishvili,R., Joachimiak,A., Tereshko,V. and Egli,M. (1999) Structure and Photoinduced Electron Transfer in Exceptionally Stable Synthetic DNA Hairpins with Stilbenediether Linkers. *J. Am. Chem. Soc.*, **121**, 9905–9906.

112. Lewis,F.D., Liu,X., Wu,Y. and Zuo,X. (2003) Stepwise Evolution of the Structure and Electronic Properties of DNA. *J. Am. Chem. Soc.*, **125**, 12729–12731.

113. Malinovskii,V.L., Wenger,D. and Häner,R. (2010) Nucleic acid-guided assembly of aromatic chromophores. *Chem. Soc. Rev.*, **39**, 410–422.

114. Lindahl,T. and Nyberg,B. (1972) Rate of depurination of native deoxyribonucleic acid. *Biochemistry*, **11**, 3610–3618.

115. Berthet,N., Constant,J.-F., Demeunynck,M., Michon,P. and Lhomme,J. (1997) Search for DNA Repair Inhibitors: Selective Binding of Nucleic Bases−Acridine Conjugates to a DNA Duplex Containing an Abasic Site. *J. Med. Chem.*, **40**, 3346–3352.

116. Berthet,N., Michon,J., Lhomme,J., Teulade-Fichou,M.-P., Vigneron,J.P. and Lehn,J.-M. (1999) Recognition of Abasic Sites in DNA by a Cyclobisacridine Molecule. *Chem. Eur. J.*, **5**, 3625–3630.

117. Yoshimoto,K., Nishizawa,S., Minagawa,M. and Teramae,N. (2003) Use of Abasic Site-Containing DNA Strands for Nucleobase Recognition in Water. *J. Am. Chem. Soc.*, **125**, 8982–8983.

118. Lhomme,J., Constant,J.-F. and Demeunynck,M. (1999) Abasic DNA structure, reactivity, and recognition. *Biopolymers*, **52**, 65–83.

119. Matray,T.J. and Kool,E.T. (1998) Selective and Stable DNA Base Pairing without Hydrogen Bonds. *J. Am. Chem. Soc.*, **120**, 6191–6192.

120. Smirnov,S. (2002) Integrity of duplex structures without hydrogen bonding: DNA with pyrene paired at abasic sites. *Nucleic Acids Research*, **30**, 5561–5569.

121. Malinovskii,V.L., Samain,F. and Häner,R. (2007) Helical Arrangement of Interstrand Stacked Pyrenes in a DNA Framework. *Angew Chem Int Ed*, **46**, 4464–4467.

122. Ziach,K., Chollet,C., Parissi,V., Prabhakaran,P., Marchivie,M., Corvaglia,V., Bose,P.P., Laxmi-Reddy,K., Godde,F., Schmitter,J.-M., *et al.* (2018) Single helically folded aromatic oligoamides that mimic the charge surface of double-stranded B-DNA. *Nature Chem*, **10**, 511–518.

123. Corvaglia,V., Wu,J., Deepak,D., Loos,M. and Huc,I. (2024) Enhancing the Features of DNA Mimic Foldamers for Structural Investigations. *Chemistry A European J*, **30**, e202303650.

124. Loos,M., Xu,F., Mandal,P.K., Chakrabortty,T., Douat,C., Konrad,D.B., Cabbar,M., Singer,J., Corvaglia,V., Carell,T., *et al.* (2025) Interfacing B-DNA and DNA Mimic Foldamers. *Angew Chem Int Ed*, **64**, e202505273.

125. Childs-Disney,J.L., Yang,X., Gibaut,Q.M.R., Tong,Y., Batey,R.T. and Disney,M.D. (2022) Targeting RNA structures with small molecules. *Nat Rev Drug Discov*, **21**, 736–762.

126. Ursu,A., Childs-Disney,J.L., Andrews,R.J., O'Leary,C.A., Meyer,S.M., Angelbello,A.J., Moss,W.N. and Disney,M.D. (2020) Design of small molecules targeting RNA structure from sequence. *Chem. Soc. Rev.*, **49**, 7252–7270.

127. Washietl,S., Hofacker,I.L., Lukasser,M., Hüttenhofer,A. and Stadler,P.F. (2005) Mapping of conserved RNA secondary structures predicts thousands of functional noncoding RNAs in the human genome. *Nat Biotechnol*, **23**, 1383–1390.

128. Zafferani,M. and Hargrove,A.E. (2021) Small molecule targeting of biologically relevant RNA tertiary and quaternary structures. *Cell Chemical Biology*, **28**, 594–609.

129. Bouton,L., Ecoutin,A., Malard,F. and Campagne,S. (2024) Small molecules modulating RNA splicing: a review of targets and future perspectives. *RSC Med. Chem.*, **15**, 1109–1126.

130. Mittal,A., Ali,S.E. and Mathews,D.H. (2024) Using the RNAstructure Software Package to Predict Conserved RNA Structures. *Current Protocols*, **4**, e70054.

131. Menichelli,E., Lam,B.J., Wang,Y., Wang,V.S., Shaffer,J., Tjhung,K.F., Bursulaya,B., Nguyen,T.N., Vo,T., Alper,P.B., *et al.* (2022) Discovery of small molecules that target a tertiary-structured RNA. *Proc. Natl. Acad. Sci. U.S.A.*, **119**, e2213117119.

132. Umuhire Juru,A., Patwardhan,N.N. and Hargrove,A.E. (2019) Understanding the Contributions of Conformational Changes, Thermodynamics, and Kinetics of RNA–Small Molecule Interactions. *ACS Chem. Biol.*, **14**, 824–838.

133. Chu,W.-T., Yan,Z., Chu,X., Zheng,X., Liu,Z., Xu,L., Zhang,K. and Wang,J. (2021) Physics of biomolecular recognition and conformational dynamics. *Rep. Prog. Phys.*, **84**, 126601.

134. Zhou,Y., Jiang,Y. and Chen,S. (2022) RNA –ligand molecular docking: Advances and challenges. *WIREs Comput Mol Sci*, **12**, e1571.

135. Zhuo,C., Zeng,C., Liu,H., Wang,H., Peng,Y. and Zhao,Y. (2025) Advances and Mechanisms of RNA–Ligand Interaction Predictions. *Life*, **15**, 104.

136. Silvestri,I., Manigrasso,J., Andreani,A., Brindani,N., Mas,C., Reiser,J.-B., Vidossich,P., Martino,G., McCarthy,A.A., De Vivo,M., *et al.* (2024) Targeting the conserved active site of splicing machines with specific and selective small molecule modulators. *Nat Commun*, **15**, 4980.

137. Roth,A. and Breaker,R.R. (2009) The Structural and Functional Diversity of Metabolite-Binding Riboswitches. *Annu. Rev. Biochem.*, **78**, 305–334.

138. Daems,E., Dewaele,D., Barylyuk,K., De Wael,K. and Sobott,F. (2021) Aptamer-ligand recognition studied by native ion mobility-mass spectrometry. *Talanta*, **224**, 121917.

139. Padroni,G., Patwardhan,N.N., Schapira,M. and Hargrove,A.E. (2020) Systematic analysis of the interactions driving small molecule–RNA recognition. *RSC Med. Chem.*, **11**, 802–813.

140. Hu,G., Li,H., Xu,S. and Wang,J. (2020) Ligand Binding Mechanism and Its Relationship with Conformational Changes in Adenine Riboswitch. *IJMS*, **21**, 1926.

141. Huang,L., Wang,J., Watkins,A.M., Das,R. and Lilley,D.M.J. (2019) Structure and ligand binding of the glutamine-II riboswitch. *Nucleic Acids Research*, **47**, 7666–7675.

142. Han,X., Gong,N., Xue,L., Billingsley,M.M., El-Mayta,R., Shepherd,S.J., Alameh,M.-G., Weissman,D. and Mitchell,M.J. (2023) Ligand-tethered lipid nanoparticles for targeted RNA delivery to treat liver fibrosis. *Nat Commun*, **14**, 75.

143. Orlovsky,N.I., Al-Hashimi,H.M. and Oas,T.G. (2020) Exposing Hidden High-Affinity RNA Conformational States. *J. Am. Chem. Soc.*, **142**, 907–921.

144. Taghavi,A., Springer,N.A., Zanon,P.R.A., Li,Y., Li,C., Childs-Disney,J.L. and Disney,M.D. (2025) The evolution and application of RNA-focused small molecule libraries. *RSC Chem. Biol.*, **6**, 510–527.

145. Lee,Y.-T. (2024) Nexus between RNA conformational dynamics and functional versatility. *Current Opinion in Structural Biology*, **89**, 102942.

146. Meyer,S.M., Williams,C.C., Akahori,Y., Tanaka,T., Aikawa,H., Tong,Y., Childs-Disney,J.L. and Disney,M.D. (2020) Small molecule recognition of disease-relevant RNA structures. *Chem. Soc. Rev.*, **49**, 7167–7199.

147. De Souza Degenhardt,M.F. and Wang,Y.-X. (2025) Challenges and opportunities in technologies and methods for lncRNA structure determination. *Cell Biosci*, **15**, 132.

148. Tessaro,F. and Scapozza,L. (2020) How 'Protein-Docking' Translates into the New Emerging Field of Docking Small Molecules to Nucleic Acids? *Molecules*, **25**, 2749.

149. Wang,M., Yu,Y., Liang,C., Lu,A. and Zhang,G. (2016) Recent Advances in Developing Small Molecules Targeting Nucleic Acid. *IJMS*, **17**, 779.

150. Luo,J., Wei,W., Waldispühl,J. and Moitessier,N. (2019) Challenges and current status of computational methods for docking small molecules to nucleic acids. *European Journal of Medicinal Chemistry*, **168**, 414–425.

151. Hargrove,A.E. (2020) Small molecule–RNA targeting: starting with the fundamentals. *Chem. Commun.*, **56**, 14744–14756.

152. Dickerhoff,J., Warnecke,K.R., Wang,K., Deng,N. and Yang,D. (2021) Evaluating Molecular Docking Software for Small Molecule Binding to G-Quadruplex DNA. *IJMS*, **22**, 10801.

153. Falese,J.P., Donlic,A. and Hargrove,A.E. (2021) Targeting RNA with small molecules: from fundamental principles towards the clinic. *Chem. Soc. Rev.*, **50**, 2224–2243.

154. Bai,Z. and Chen,A. (2023) Predicting Small Molecule Ligand – RNA Binding Pocket Binding Modes Using Metadynamics. 10.1101/2023.10.04.560960.

155. Johnson-Buck,A.E., McDowell,S.E. and Walter,N.G. (2011) Metal Ions: Supporting Actors in the Playbook of Small Ribozymes. In Sigel,A., Sigel,H., Sigel,R.K.O., Sigel,A., Sigel,H., Sigel,R.K.O., Sigel,A., Sigel,H., Sigel,R.K.O., Sigel,A., et al. (eds), *Structural and Catalytic Roles of Metal Ions in RNA*. The Royal Society of Chemistry, pp. 175–196.

156. Bernetti,M., Aguti,R., Bosio,S., Recanatini,M., Masetti,M. and Cavalli,A. (2022) Computational drug discovery under RNA times. *QRB Discovery*, **3**, e22.

157. Veenbaas,S.D., Koehn,J.T., Irving,P.S., Lama,N.N. and Weeks,K.M. (2025) Ligand-binding pockets in RNA, and where to find them. 10.1101/2025.03.13.643147.

158. Rhodes,C., Balaratnam,S., Yazdani,K., Seshadri,S. and Schneekloth,J.S. (2024) Targeting RNA-protein interactions with small molecules: promise and therapeutic potential. *Med Chem Res*, **33**, 2050–2065.

159. Mateja-Pluta,M., Błaszczyk,L., Bejger,M., Nakatani,K. and Kiliszek,A. (2025) Naphthyridine carbamate dimer ligand induces formation of Z-RNA-like fold of disease-related RNA and exhibits a molecular glue characteristics in crystal lattice formation. *Nucleic Acids Research*, **53**, gkaf924.

160. Warner,K.D., Hajdin,C.E. and Weeks,K.M. (2018) Principles for targeting RNA with drug-like small molecules. *Nat Rev Drug Discov*, **17**, 547–558.

161. Panei,F.P., Torchet,R., Ménager,H., Gkeka,P. and Bonomi,M. (2022) HARIBOSS: a curated database of RNA-small molecules structures to aid rational drug design. *Bioinformatics*, **38**, 4185–4193.

162. Murata,A., Nakamori,M. and Nakatani,K. (2019) Modulating RNA secondary and tertiary structures by mismatch binding ligands. *Methods*, **167**, 78–91.

163. Mukherjee,S., Błaszczyk,L., Rypniewski,W., Falschlunger,C., Micura,R., Murata,A., Dohno,C., Nakatani,K. and Kiliszek,A. (2019) Structural insights into synthetic ligands targeting A–A pairs in disease-related CAG RNA repeats. *Nucleic Acids Research*, **47**, 10906–10913.

164. Nakatani,K., Hagihara,S., Goto,Y., Kobori,A., Hagihara,M., Hayashi,G., Kyo,M., Nomura,M., Mishima,M. and Kojima,C. (2005) Small-molecule ligand induces nucleotide flipping in (CAG)n trinucleotide repeats. *Nat Chem Biol*, **1**, 39–43.

165. Błaszczyk,L., Ryczek,M., Das,B., Mateja-Pluta,M., Bejger,M., Śliwiak,J., Nakatani,K. and Kiliszek,A. (2024) Antisense RNA C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms a triplex-like structure and binds small synthetic ligand. *Nucleic Acids Research*, **52**, 6707–6717.

166. Nakamori,M., Panigrahi,G.B., Lanni,S., Gall-Duncan,T., Hayakawa,H., Tanaka,H., Luo,J., Otabe,T., Li,J., Sakata,A., *et al.* (2020) A slipped-CAG DNA-binding small molecule induces trinucleotide-repeat contractions in vivo. *Nat Genet*, **52**, 146–159.

167. Yamada,T., Furuita,K., Sakurabayashi,S., Nomura,M., Kojima,C. and Nakatani,K. (2022) NMR determination of the 2:1 binding complex of naphthyridine carbamate dimer (NCD) and CGG/CGG triad in double-stranded DNA. *Nucleic Acids Research*, **50**, 9621–9631.

168. Shibata,T., Nakamachi,A. and Nakatani,K. (2024) Enhancing Binding Affinity to CGG/CGG Triad: Optimizing Naphthyridine Carbamate Dimer Derivatives with Varied Linker Lengths. *ChemMedChem*, **19**, e202400351.

169. Simeth,N.A., Kobayashi,S., Kobauri,P., Crespi,S., Szymanski,W., Nakatani,K., Dohno,C. and Feringa,B.L. (2021) Rational design of a photoswitchable DNA glue enabling high regulatory function and supramolecular chirality transfer. *Chem. Sci.*, **12**, 9207–9220.

170. Yamada,T., Miki,S., Ul'Husna,A., Michikawa,A. and Nakatani,K. (2017) Synthesis of Naphthyridine Carbamate Dimer (NCD) Derivatives Modified with Alkanethiol and Binding Properties of G–G Mismatch DNA. *Org. Lett.*, **19**, 4163–4166.

171. Yagi,Y., Aikawa,H., Yamada,T. and Nakatani,K. (2018) Expanding chemical space of DNA-binding molecules with three base-binding units. *Bioorganic & Medicinal Chemistry Letters*, **28**, 2894–2898.

172. Konieczny,P., Mukherjee,S., Stepniak-Konieczna,E., Taylor,K., Niewiadomska,D., Piasecka,A., Walczak,A., Baud,A., Dohno,C., Nakatani,K., *et al.* (2021) Cyclic mismatch binding ligands interact with disease-associated CGG trinucleotide repeats in RNA and suppress their translation. *Nucleic Acids Research*, **49**, 9479–9495.

173. Yamada,T., Sakurabayashi,S., Sugiura,N., Haneoka,H. and Nakatani,K. (2024) NMR analysis of^{15} N-labeled naphthyridine carbamate dimer (NCD) to contiguous CGG/CGG units in DNA. *Chem. Commun.*, **60**, 3645–3648.

174. Nakatani,K. (2022) Possibilities and challenges of small molecule organic compounds for the treatment of repeat diseases. *Proc. Jpn. Acad., Ser. B*, **98**, 30–48.

175. Hong,C., Hagihara,M. and Nakatani,K. (2011) Ligand-Assisted Complex Formation of Two DNA Hairpin Loops. *Angew Chem Int Ed*, **50**, 4390–4393.

176. Hong,C., Otabe,T., Matsumoto,S., Dohno,C., Murata,A., Hagihara,M. and Nakatani,K. (2014) Formation of a Ligand-Assisted Complex of Two RNA Hairpin Loops. *Chemistry A European J*, **20**, 5282–5287.

177. Dohno,C., Atsumi,H. and Nakatani,K. (2011) Ligand inducible assembly of a DNA tetrahedron. *Chem. Commun.*, **47**, 3499.

178. Dohno,C. and Nakatani,K. (2011) Control of DNA hybridization by photoswitchable molecular glue. *Chem. Soc. Rev.*, **40**, 5718.

179. Ni,L., Yamada,T., Murata,A. and Nakatani,K. (2022) Mismatch binding ligand upregulated back-splicing reaction producing circular RNA in a cellular model. *Chem. Commun.*, **58**, 3629–3632.

180. Niimi,Y., Takahashi,M., Sugawara,E., Umeda,S., Obayashi,M., Sato,N., Ishiguro,T., Higashi,M., Eishi,Y., Mizusawa,H., *et al.* (2013) Abnormal RNA structures (RNA foci) containing a penta-nucleotide repeat (UGGAA)$_n$ in the P urkinje cell nucleus is associated with spinocerebellar ataxia type 31 pathogenesis. *Neuropathology*, **33**, 600–611.

181. Ishiguro,T., Nagai,Y. and Ishikawa,K. (2021) Insight Into Spinocerebellar Ataxia Type 31 (SCA31) From Drosophila Model. *Front. Neurosci.*, **15**, 648133.

182. Loureiro,J.R., Castro,A.F., Figueiredo,A.S. and Silveira,I. (2022) Molecular Mechanisms in Pentanucleotide Repeat Diseases. *Cells*, **11**, 205.

183. Ishikawa,K. (2023) Spinocerebellar ataxia type 31 (SCA31). *J Hum Genet*, **68**, 153–156.

184. Ishikawa,K. (2024) Recent topics of spinocerebellar ataxia type 31. *Neurology & Clinical Neurosc*, **12**, 24–28.

185. Shibata,T. and Nakatani,K. (2023) A small molecule binding to TGGAA pentanucleotide repeats that cause spinocerebellar ataxia type 31. *Bioorganic & Medicinal Chemistry Letters*, **79**, 129082.

186. Shibata,T., Nagano,K., Ueyama,M., Ninomiya,K., Hirose,T., Nagai,Y., Ishikawa,K., Kawai,G. and Nakatani,K. (2021) Small molecule targeting r(UGGAA)n disrupts RNA foci and alleviates disease phenotype in Drosophila model. *Nat Commun*, **12**, 236.

187. Ryczek,M., Pluta,M., Błaszczyk,L. and Kiliszek,A. (2022) Overview of Methods for Large-Scale RNA Synthesis. *Applied Sciences*, **12**, 1543.

188. Pasternak,A. and Wengel,J. (2010) Thermodynamics of RNA duplexes modified with unlocked nucleic acid nucleotides. *Nucleic Acids Research*, **38**, 6697–6706.

189. Dessau,M.A. and Modis,Y. (2011) Protein Crystallization for X-ray Crystallography. *JoVE*, 10.3791/2285.

190. Kondo,J., Sauter,C. and Masquida,B. (2014) RNA Crystallization. In Hartmann,R.K., Bindereif,A., Schön,A., Westhof,E. (eds), *Handbook of RNA Biochemistry*. Wiley, pp. 481–498.

191. McPherson,A. and Cudney,B. (2014) Optimization of crystallization conditions for biological macromolecules. *Acta Crystallogr F Struct Biol Commun*, **70**, 1445–1467.

192. Cianci,M., Bourenkov,G., Pompidor,G., Karpics,I., Kallio,J., Bento,I., Roessle,M., Cipriani,F., Fiedler,S. and Schneider,T.R. (2017) P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable beam focusing. *J Synchrotron Rad*, **24**, 323–332.

193. Burkhardt,A., Pakendorf,T., Reime,B., Meyer,J., Fischer,P., Stübe,N., Panneerselvam,S., Lorbeer,O., Stachnik,K., Warmer,M., *et al.* (2016) Status of the crystallography beamlines at PETRA III. *Eur. Phys. J. Plus*, **131**, 56.

194. Garman,E. (1999) Cool data: quantity AND quality. *Acta Crystallogr D Biol Crystallogr*, **55**, 1641–1653.

195. Wlodawer,A., Minor,W., Dauter,Z. and Jaskolski,M. (2008) Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *The FEBS Journal*, **275**, 1–21.

196. Otwinowski,Z. and Minor,W. (1997) [20] Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology*. Elsevier, Vol. 276, pp. 307–326.

197. Kabsch,W. (2010) *XDS*. *Acta Crystallogr D Biol Crystallogr*, **66**, 125–132.

198. Powell,H.R. (2017) X-ray data processing. *Bioscience Reports*, **37**, BSR20170227.

199. Gagnér,V.A., Lundholm,I., Garcia-Bonete,M.-J., Rodilla,H., Friedman,R., Zhaunerchyk,V., Bourenkov,G., Schneider,T., Stake,J. and Katona,G. (2019) Clustering of atomic displacement parameters in bovine trypsin reveals a distributed lattice of atoms with shared chemical properties. *Sci Rep*, **9**, 19281.

200. Brehm,W., Triviño,J., Krahn,J.M., Usón,I. and Diederichs,K. (2023) *XDSGUI* : a graphical user interface for *XDS* , *SHELX* and *ARCIMBOLDO*. *J Appl Crystallogr*, **56**, 1585–1594.

201. Karplus,P.A. and Diederichs,K. (2012) Linking Crystallographic Model and Data Quality. *Science*, **336**, 1030–1033.

202. Evans,J.S.O. and Evans,I.R. (2021) Structure Analysis from Powder Diffraction Data: Rietveld Refinement in Excel. *J. Chem. Educ.*, **98**, 495–505.

203. IX. Bakerian Lecture.— X-rays and crystal structure (1915) *Phil. Trans. R. Soc. Lond. A*, **215**, 253–274.

204. Rossmann,M.G. (1990) The molecular replacement method. *Acta Crystallogr A Found Crystallogr*, **46**, 73–82.

205. Rossmann,M.G. (2001) Molecular replacement – historical background. *Acta Crystallogr D Biol Crystallogr*, **57**, 1360–1366.

206. Evans,P. and McCoy,A. (2008) An introduction to molecular replacement. *Acta Crystallogr D Biol Crystallogr*, **64**, 1–10.

207. Rupp,B. (2009) Biomolecular Crystallography 0 edn Garland Science.

208. Taylor,G. (2003) The phase problem. *Acta Crystallogr D Biol Crystallogr*, **59**, 1881–1890.

209. Cowtan,K. (2003) Phase Problem in X -ray Crystallography, and Its Solution. In *Encyclopedia of Life Sciences*. Wiley.

210. Rajashankar,K. and Dauter,Z. (2014) Data Collection for Crystallographic Structure Determination. In Anderson,W.F. (ed), *Structural Genomics and Drug Discovery*,

Methods in Molecular Biology. Springer New York, New York, NY, Vol. 1140, pp. 211–237.

211. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) *Phaser* crystallographic software. *J Appl Crystallogr*, **40**, 658–674.

212. Brünger,A.T. (1996) Recent Developments for Crystallographic Refinement of Macromolecules. In *Crystallographic Methods and Protocols*. Humana Press, New Jersey, Vol. 56, pp. 245–266.

213. Tronrud,Dale.E. (2004) Introduction to macromolecular refinement. *Acta Crystallogr D Biol Crystallogr*, **60**, 2156–2168.

214. Steiner,R.A. and Rupp,B. (2012) Model building, refinement and validation. *Acta Crystallogr D Biol Crystallogr*, **68**, 325–327.

215. Emsley,P. and Cowtan,K. (2004) *Coot* : model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, **60**, 2126–2132.

216. Murshudov,G.N., Skubák,P., Lebedev,A.A., Pannu,N.S., Steiner,R.A., Nicholls,R.A., Winn,M.D., Long,F. and Vagin,A.A. (2011) *REFMAC* 5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr*, **67**, 355–367.

217. Beilsten-Edmands,J., Winter,G., Gildea,R., Parkhurst,J., Waterman,D. and Evans,G. (2020) Scaling diffraction data in the *DIALS* software package: algorithms and new approaches for multi-crystal scaling. *Acta Crystallogr D Struct Biol*, **76**, 385–399.

218. Williamson,L.J., Galchenkova,M., Best,H.L., Bean,R.J., Munke,A., Awel,S., Pena,G., Knoska,J., Schubert,R., Dörner,K., *et al.* (2023) Structure of the *Lysinibacillus sphaericus* Tpp49Aa1 pesticidal protein elucidated from natural crystals using MHz-SFX. *Proc. Natl. Acad. Sci. U.S.A.*, **120**, e2203241120.

219. Andaleeb,H., Ullah,N., Falke,S., Perbandt,M., Brognaro,H. and Betzel,C. (2020) High-resolution crystal structure and biochemical characterization of a GH11 endoxylanase from Nectria haematococca. *Sci Rep*, **10**, 15658.

220. Parkinson,G., Vojtechovsky,J., Clowney,L., Brünger,A.T. and Berman,H.M. (1996) New parameters for the refinement of nucleic acid-containing structures. *Acta Crystallogr D Biol Crystallogr*, **52**, 57–64.

221. Lu,X.-J. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Research*, **31**, 5108–5121.

222. Langenegger,S.M., Malinovskii,V.L., Wenger,D., Werder,S. and Häner,R. (2007) DNA Containing Non-Nucleosidic Phenanthrene Building Blocks with Asymmetrical Linkers. *Nucleosides, Nucleotides & Nucleic Acids*, **26**, 901–903.

223. Langenegger,S.M. and Häner,R. (2002) The Effect of a Non-nucleosidic Phenanthrene Building Block on DNA Duplex Stability. *HCA*, **85**, 3414–3421.

224. Beaucage,S.L. and Caruthers,M.H. (1981) Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Letters*, **22**, 1859–1862.

225. Matthews,B.W. (1968) Solvent content of protein crystals. *Journal of Molecular Biology*, **33**, 491–497.

226. Kiliszek,A., Kierzek,R., Krzyzosiak,W.J. and Rypniewski,W. (2009) Structural insights into CUG repeats containing the 'stretched U–U wobble': implications for myotonic dystrophy. *Nucleic Acids Research*, **37**, 4149–4156.

227. Shi,X., Khade,P.K., Sanbonmatsu,K.Y. and Joseph,S. (2012) Functional Role of the Sarcin–Ricin Loop of the 23S rRNA in the Elongation Cycle of Protein Synthesis. *Journal of Molecular Biology*, **419**, 125–138.

228. Schulze-Gahmen,U. and Hurley,J.H. (2018) Structural mechanism for HIV-1 TAR loop recognition by Tat and the super elongation complex. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, 12973–12978.

229. Lebars,I., Legrand,P., Aimé,A., Pinaud,N., Fribourg,S. and Di Primo,C. (2008) Exploring TAR–RNA aptamer loop–loop interaction by X-ray crystallography, UV spectroscopy and surface plasmon resonance. *Nucleic Acids Research*, **36**, 7146–7156.

230. Lebars,I., Richard,T., Di Primo,C. and Toulmé,J.-J. (2007) NMR structure of a kissing complex formed between the TAR RNA element of HIV-1 and a LNA-modified aptamer. *Nucleic Acids Research*, **35**, 6103–6114.

231. Kolb,G., Reigadas,S., Castanotto,D., Faure,A., Ventura,M., Rossi,J.J. and Toulme,J.J. (2006) Endogenous Expression of an Anti-TAR Aptamer Reduces HIV-1 Replication. *RNA Biology*, **3**, 150–156.

232. Beaurain,F. (2003) Molecular dynamics reveals the stabilizing role of loop closing residues in kissing interactions: comparison between TAR-TAR* and TAR-aptamer. *Nucleic Acids Research*, **31**, 4275–4284.

233. Kurzynska-Kokorniak,A., Pokornowska,M., Koralewska,N., Hoffmann,W., Bienkowska-Szewczyk,K. and Figlerowicz,M. (2016) Revealing a new activity of the human Dicer DUF283 domain in vitro. *Sci Rep*, **6**, 23989.

234. Correll,C.C. (2003) The common and the distinctive features of the bulged-G motif based on a 1.04 A resolution RNA structure. *Nucleic Acids Research*, **31**, 6806–6818.

235. Antao,V.P., Lai,S.Y. and Tinoco,I. (1991) A thermodynamic study of unusually stable RNA and DNA hairpins. *Nucl Acids Res*, **19**, 5901–5905.

236. Jucker,F.M., Heus,H.A., Yip,P.F., Moors,E.H.M. and Pardi,A. (1996) A Network of Heterogeneous Hydrogen Bonds in GNRA Tetraloops. *Journal of Molecular Biology*, **264**, 968–980.

237. Nichols,P.J., Krall,J.B., Henen,M.A., Vögeli,B. and Vicens,Q. (2023) Z-RNA biology: a central role in the innate immune response? *RNA*, **29**, 273–281.

238. Gessner,R.V., Quigley,G.J., Wang,A.H.J., Van Der Marel,G.A., Van Boom,J.H. and Rich,A. (1985) Structural basis for stabilization of Z-DNA by cobalt hexaammine and magnesium cations. *Biochemistry*, **24**, 237–240.

239. Ennifar,E. (2003) A crystallographic study of the binding of 13 metal ions to two related RNA duplexes. *Nucleic Acids Research*, **31**, 2671–2682.

240. Cheatham,T.E. and Kollman,P.A. (1997) Insight into the stabilization of A-DNA by specific ion association: spontaneous B-DNA to A-DNA transitions observed in molecular dynamics simulations of d[ACCCGCGGGT]2 in the presence of hexaamminecobalt(III). *Structure*, **5**, 1297–1311.

241. Crooke,S.T. (1998) An Overview of Progress in Antisense Therapeutics. *Antisense and Nucleic Acid Drug Development*, **8**, 115–122.

242. De Mesmaeker,A., Haener,R., Martin,P. and Moser,H.E. (1995) Antisense Oligonucleotides. *Acc. Chem. Res.*, **28**, 366–374.

243. Zhang,Z., Chen,S., Li,Z. and Yu,H. (2023) Functional nucleic acids with synthetic sugar or nucleobase moieties. *Advanced Agrochem*, **2**, 236–245.

244. D'Esposito,R.J., Myers,C.A., Chen,A.A. and Vangaveti,S. (2022) Challenges with Simulating Modified RNA: Insights into Role and Reciprocity of Experimental and Computational Approaches. *Genes*, **13**, 540.

245. Chen,T., Hongdilokkul,N., Liu,Z., Thirunavukarasu,D. and Romesberg,F.E. (2016) The expanding world of DNA and RNA. *Current Opinion in Chemical Biology*, **34**, 80–87.

246. Kowalska,J., Osowniak,A., Zuberek,J. and Jemielity,J. (2012) Synthesis of nucleoside phosphosulfates. *Bioorganic & Medicinal Chemistry Letters*, **22**, 3661–3664.

247. Dabrowski-Tumanski,P., Kowalska,J. and Jemielity,J. (2013) Efficient and Rapid Synthesis of Nucleoside Diphosphate Sugars from Nucleoside Phosphorimidazolides. *Eur J Org Chem*, **2013**, 2147–2154.

248. Strenkowska,M., Wanat,P., Ziemniak,M., Jemielity,J. and Kowalska,J. (2012) Preparation of Synthetically Challenging Nucleotides Using Cyanoethyl P-Imidazolides and Microwaves. *Org. Lett.*, **14**, 4782–4785.

11. Attachments

1. Attachment 1: Marcin Ryczek, Martyna Pluta, Leszek Błaszczyk*, Agnieszka Kiliszek*, *Overview of Methods for Large-Scale RNA Synthesis,* Applied Sciences **2022**, 12, 1543;

2. Attachment 2: Martyna Mateja-Pluta, Leszek Błaszczyk, Magdalena Bejger, Kazuhiko Nakatani, Agnieszka Kiliszek*, *Naphthyridine carbamate dimer ligand induces formation of Z-RNA-like fold of disease-related RNA and exhibits a molecular glue characteristics in crystal lattice formation,* Nucleic Acids Research, **2025**, 53, gkaf924

3. Attachment 3: Declarations determining the candidate's contribution to the creation of the publications constituted in the doctoral dissertation

applied sciences

MDPI

Review

Overview of Methods for Large-Scale RNA Synthesis

Marcin Ryczek, Martyna Pluta, Leszek Błaszczyk * and Agnieszka Kiliszek *

Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznań, Poland; mryczek@ibch.poznan.pl (M.R.); mpluta@ibch.poznan.pl (M.P.)
* Correspondence: blaszcz@ibch.poznan.pl (L.B.); kiliszek@ibch.poznan.pl (A.K.)

Abstract: In recent years, it has become clear that RNA molecules are involved in almost all vital cellular processes and pathogenesis of human disorders. The functional diversity of RNA comes from its structural richness. Although composed of only four nucleotides, RNA molecules present a plethora of secondary and tertiary structures critical for intra and intermolecular contacts with other RNAs and ligands (proteins, small metabolites, etc.). In order to fully understand RNA function it is necessary to define its spatial structure. Crystallography, nuclear magnetic resonance and cryogenic electron microscopy have demonstrated considerable success in determining the structures of biologically important RNA molecules. However, these powerful methods require large amounts of sample. Despite their limitations, chemical synthesis and in vitro transcription are usually employed to obtain milligram quantities of RNA for structural studies, delivering simple and effective methods for large-scale production of homogenous samples. The aim of this paper is to provide an overview of methods for large-scale RNA synthesis with emphasis on chemical synthesis and in vitro transcription. We also present our own results of testing the efficiency of these approaches in order to adapt the material acquisition strategy depending on the desired RNA construct.

Keywords: RNA; RNA crystallography; large-scale RNA synthesis; ribozymes

Citation: Ryczek, M.; Pluta, M.; Błaszczyk, L.; Kiliszek, A. Overview of Methods for Large-Scale RNA Synthesis. *Appl. Sci.* **2022**, *12*, 1543. https://doi.org/10.3390/app12031543

Academic Editor: Chih-Ching Huang

Received: 28 October 2021
Accepted: 28 January 2022
Published: 31 January 2022

1. Introduction

Ribonucleic acid (RNA) is one of the major players involved in cell homeostasis [1]. Specific chemical features make RNA an important biological molecule, an interesting research object, and a tool adapted for human use [2,3]. In early nucleic acid research, there was no clear distinction between RNA and deoxyribonucleic acid (DNA); they were seen as "nuclein", a phosphor-rich substance different from proteins and present in large quantities in every living organism [4]. The pioneering work of many researchers, including Phoebus Levene, Albrech Kossel, and their co-workers, allowed to distinguish DNA and RNA molecules and to identify their basic components, such as ribose rings, nitrogen-containing bases, and phosphate groups [5–7]. Since then, RNA has been revealed as a multifaceted molecule that can form complex three-dimensional structures, presents catalytic properties, and serves as a regulatory element for governing gene expression [8–16]. On the other hand, RNA has also been identified as a toxic factor involved in the pathogenesis of numerous human disorders [17–19]. Finally, RNA became a tool applied in biotechnology and medicine [20–25]. The best-known examples are the mRNA-based vaccines used to fight the COVID-19 pandemic [26].

Knowledge of the three-dimensional structure of RNA is crucial for explaining its biological functions and conformational dynamics and for providing model platforms for drug development [15,27,28]. The pioneering studies on nucleic acid structure originated from research by Watson, Crick, and Franklin, who proposed the double-DNA-helix model [29,30]. The first atomic transfer RNA (tRNA) structure was determined by Robert Holley's group using single-crystal x-ray crystallography, opening up the field of structural studies of RNA molecules [31–33]. Subsequent discoveries of RNA properties (e.g., catalytic activity) and the development of methods for chemical and enzymatic synthesis increased

interest in RNA and RNA–protein research, resulting in the three-dimensional structures of ribozymes, riboswitches, ribosomes, small nucleolar RNAs (snoRNAs), and other types of RNA being discovered [34–56]. These fundamental findings helped to explain the principles of RNA functions, folding pathways, the roles of water and ion molecules, and to reveal ligand-binding sites [57].

Structural studies of RNA molecules start with the design, synthesis, and purification of the macromolecule. In this step, the purity and homogeneity of the sample are critical for successful crystallization and diffraction experiments [58,59]. Another obstacle is the amount of material required for crystallization; milligrams of homogeneous RNA are usually needed [60–63]. Although a number of well-established protocols exist, obtaining a sufficient amount of pure RNA is challenging [60–63]. While chemical synthesis allows the production of large quantities of RNA, it is suitable only for relatively short RNA oligomers and requires a laboratory equipped with specific infrastructure (synthesizer, fume hood, vacuum evaporator, etc.). Moreover, modified phosphoramidites can be expensive and often non-commercially available. Longer RNAs can be effectively synthesized using in vitro transcription but this method suffers from considerable 3′-end heterogeneity and requires specific sequence requirements at the 5′-end. This paper provides an overview of large-scale RNA-synthesis methods used in structural studies, supplemented by our own results with the practical implementation of available protocols. In the first part, we describe methods for RNA synthesis and discuss their advantages and limitations. Next, we present the implementation of these methods and guidelines for adapting RNA-synthesis strategies.

2. Methods of RNA Synthesis

RNA can be obtained in three ways: purification from biological sources, chemical synthesis using a solid-phase method, or enzymatic synthesis by in vitro transcription. Isolation from cells is suitable for complex or large and highly abundant RNAs (ribosomes or tRNAs), resulting in native samples having all the important posttranscriptional modifications. Shorter or less-abundant RNAs can be synthesized chemically or enzymatically. These methods are more universal and preferred for structural studies of RNA molecules.

2.1. Chemical Synthesis

Solid-phase chemical synthesis using an automated synthesizer was established by Bruce Merrifield and was further applied for the synthesis of deoxyribonucleic and ribonucleic acids [64,65]. It is based on the cyclic elongation of the DNA/RNA chain on a solid support, e.g., controlled-pore glass or highly crosslinked polystyrene. It uses phosphoramidites decorated with different protection groups in order to block their reactivity during chemical synthesis [66]. There are several types of commercially available phosphoramidites. Most of them possess mild base-labile protection of exocyclic amino group of nucleobase and phosphate moiety blocked by 2-cyanoethyl *N*,*N*,*N*′,*N*′-tetra-isopropyl [66]. In the case of 2′-hydroxyl, the most widely employed class of protection groups are TOM (triisopropylsilyl-oxy-methyl) and TBDMS (tert-butyldimethylsilyl). They are fluoride-labeled groups which are removed after synthesis using TBAF (tetra-n-butylammonium fluoride) or TEA·3HF (triethylamine trihydrofluoride). The 5′ OH group is usually blocked by DMT (dimethoxytrityl) which is compatible with 2′OH silyl blockage [67]. The selection of the 2′-protection group is the most critical step affecting the yield and time of RNA synthesis. Recently, more potent activators such as ETT (5-ethylthio-1H-tetrazole) and BTT (5-benzylthio-1H-tetrazole) helped to increase the rate of the coupling step [68]. Other strategies of phosphoramidites protection have also been developed, resulting in shorter and higher efficiency of RNA synthesis. However, they are expensive, non-commercially available or require different protocols of solid-phase synthesis (for review see [66,69]). The standard solid-phase synthesis occurs from the 3′ to 5′-end, and one elongation cycle results in a chain extension by one nucleotide. Each cycle includes four steps: detritylation, coupling, capping, and oxidation (Figure 1) [66]. It starts with removing the dimethoxytrityl (DMT) protection group from the 5′ OH group of the nucleotide attached to the solid

support. This is followed by a coupling reaction, resulting in the chain elongation of oligomers with a 5′ DMT group originating from added phosphoramidite. Since the yield of coupling is less than 100%, not all RNA chains are elongated. Thus, in the next step, the non-extended oligomers are 5′ capped, preventing their further extension in the next cycle. Finally, the phosphate group is oxidized and converted into a stable moiety [62,66,70].

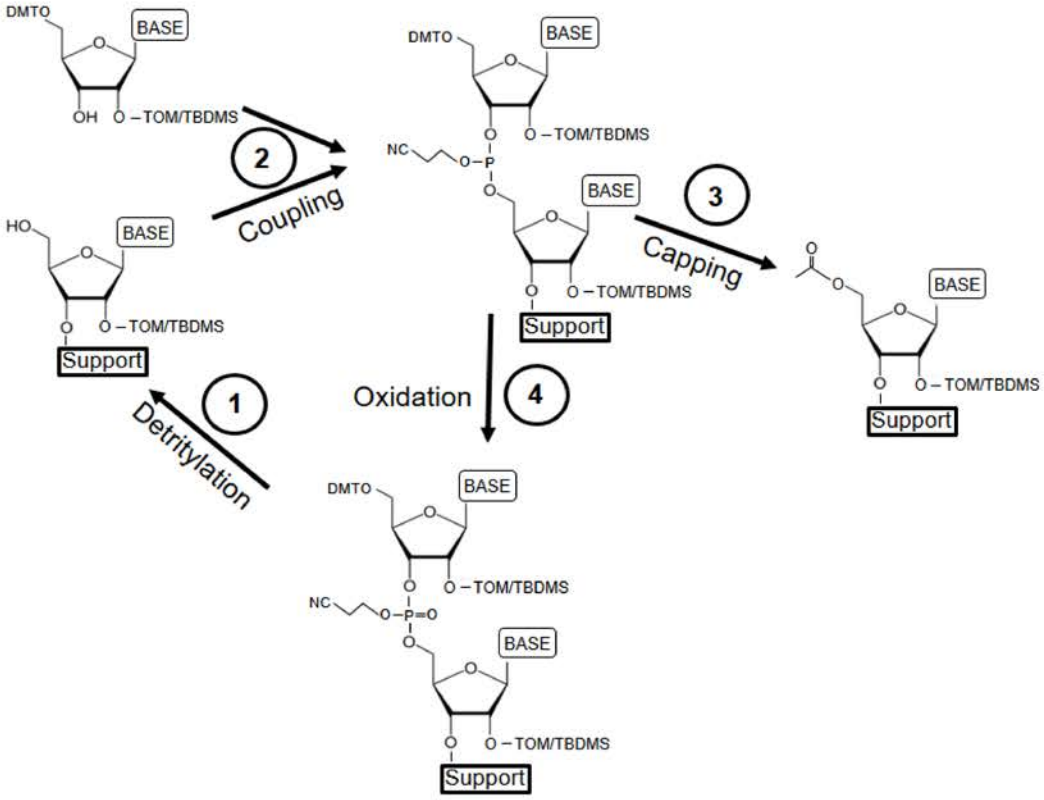


Figure 1. Overview of the solid-phase chemical RNA synthesis cycle.

Chemical synthesis is suitable for relatively short oligomers (up to 40 nt-long). Although the yield of a single elongation cycle can reach 99%, the overall efficiency drops systematically with each cycle. For 25 nt-long oligomers, the overall yield of synthesis is about 79%, assuming an average single coupling efficiency of 99%. When the coupling efficiency drops to 97%, the overall yield is only approximately 48%. For longer oligomers (50 nt), the overall yield is only 37%, even if the average performance of a single coupling is 98% [71]. Therefore, chemical synthesis for longer RNAs is impractical, unless other reasons are considered, such as the introduction of modified nucleotides. In contrast to in vitro transcription, chemical modifications can be easily incorporated into RNA chains during automated synthesis. If the oligomer is too long to be efficiently synthesized, two shorter modified RNA chains can be obtained separately and ligated using split–join ligation or T4 RNA ligase [72–74]. The other advantage of chemical synthesis is fewer constraints on the sequence of RNA and lower heterogeneity in comparison to in vitro transcription.

In our laboratory, chemical synthesis is routinely used for the preparation of RNA oligomers. As an example, we have obtained 11 nt-long RNA oligomers using two approaches: DMT-OFF, in which the 5′ DMT group is removed from the last synthesized nucleotide, and DMT-ON, in which the 5′ DMT group is preserved and used for an additional purification step (Supplementary Materials). Both oligomers were cleaved from a solid support, and protective groups were removed according to standard procedures [66,67]. The DMT-ON oligomer was further purified using cartridges with an affinity for dimethoxy groups, while the DMT-OFF oligomer was only desalted. A comparison of the high-performance liquid chromatography (HPLC) chromatograms of samples showed high purity and homogeneity for DMT-ON RNA (compare Figure 2A,B). The DMT-OFF

oligomer contained many short products, indicating the need for an additional purification step using thin-layer chromatography, polyacrylamide gel electrophoresis (PAGE), or HPLC. We also synthesized a 45 nt-long RNA oligomer (Figure 2C). Even though it was purified using the DMT-ON method, the amount of abortive products generated during the synthesis of longer RNA was very large, resulting in lower efficiency for the purification method.

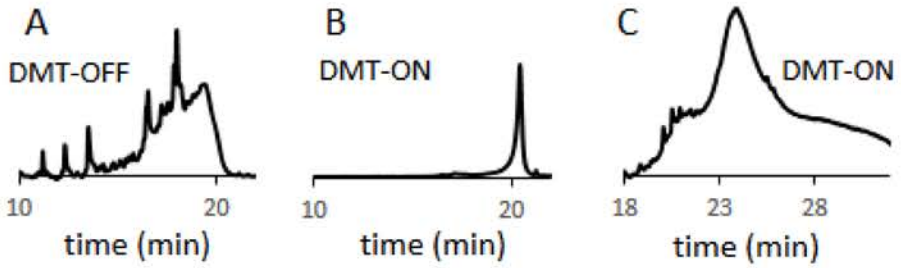


Figure 2. HPLC chromatograms of chemically synthesized RNAs: (**A**,**B**) 11 nt-long RNA oligomer synthesized in DMT-OFF and DMT-ON mode, respectively; (**C**) 45 nt-long RNA synthesized in DMT-ON mode.

In our hands, the average yield from 1 µmol-scale RNA synthesis resulted in at least 200 nmol of pure oligomer of 8 to 22 nt. Short oligomers purified by the DMT-ON method were crystallized, and many of them resulted in crystals exhibiting a high diffraction potential (Figure 3 and data not shown). Longer RNAs were subjected to an additional purification step before crystallization.

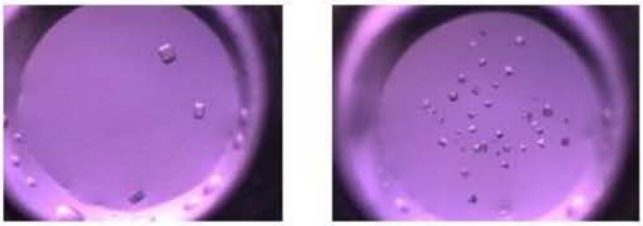


Figure 3. Examples of crystals that were grown from chemically synthesized RNAs.

2.2. In Vitro RNA Synthesis Using T7 RNA Polymerase

In vitro transcription is a common method used for synthesizing RNA molecules. It utilizes a DNA template with a promoter sequence for T7 RNA polymerase followed by a sequence encoding the target molecule [75]. The enzyme binds the template at the promoter region and starts the RNA synthesis. The elongation of the RNA chain terminates when the enzyme drops off the 3′-end of the DNA template [75]. In vitro synthesis can be efficient, resulting in milligrams of RNA sample, but a high yield depends on several factors. The most important is the type of nucleotides located immediately downstream of the T7 promoter. The transcript should start from at least one guanosine residue, but having two or three consecutive Gs is better [75–77]. Usually, the RNA sequence can be adjusted and mutations can be easily introduced into the DNA template. If the RNA sequence cannot be changed, catalytic RNAs can be employed (see below).

In structural studies, a serious drawback of in vitro RNA synthesis is the presence of short, abortive transcripts due to polymerase slippage and transcripts with additional nucleotides added to the 3′-end that are not encoded by the DNA template [78]. Separating short products from full-length oligomers is usually easy, while it is more difficult and time-consuming with transcripts extended by an additional one or two nucleotides [58]. The amounts of N+1 and N+2 products range from 10 to 50% of the whole synthesis, reducing the homogeneity, yield, and crystallization potential of the RNA sample [59].

In vitro transcription is conducted with a mixture of nucleoside triphosphates (NTPs), resulting in RNA having tri-phosphate at the 5′-end. In some instances, it can negatively

influence the crystallization potential of the RNA by preventing the packing of RNA molecules into a crystal lattice or introducing disorder [79]. Thus, the conversion of the 5′ triphosphate into a 5′ hydroxyl group may be required to improve the crystallization and quality of the crystals. This can be easily performed by in vitro transcription in the presence of guanosine residues or with the use of alkaline phosphatase [59]. If a 5′ monophosphate group is required for further RNA modification, the transcription reaction can be supplemented with guanosine monophosphate (GMP).

In vitro transcription can be employed for the production of capped RNAs. Cap is a modified nucleotide (7-methylguanosine, m^7G) linked by the 5′,5′-triphosphate bond to the first transcribed nucleotide [80,81]. Cap structure increases the stability of cellular and exogenous mRNAs (such as vaccine mRNA) [82,83]. It is also required for mRNA splicing, export and initiation of translation [82,84,85]. Therefore, capped RNAs are widely used in studies concerning vital cellular processes [86–88]. The increasing demand on simple and efficient methods for synthesis of capped mRNAs resulted in the development of several protocols [86,89–91]. For long RNAs, cap structure is commonly introduced during in vitro transcription in the presence of chemically synthesized cap analogs such as m^7G, anti-reverse cap analog (ARCA) and its derivatives [86,90–94]. However, this method suffers from low efficiency which rarely exceeds 80% under optimized conditions. Moreover, it is limited to Cap 0 and the first transcribed nucleotide must be purine [95]. Recently, other capping reagents have been developed, extending the availability of different cap structures and increasing the yield of cap incorporation [95,96]. Short capped RNAs are synthesized using solid-phase synthesis but it requires non-standard phosphoramidites or modified solid supports [97–99].

2.3. Ribozymes for RNA Production with Homogeneous Ends

Ribozymes are self-cleaving RNAs found mainly in small-RNA pathogens of plants [100–102]. They are involved in the replication process, with a rolling-circle mechanism. As a result, long, linear concatamers of replicated genome of the pathogen are produced, which are cleaved by the catalytic domain of the ribozyme into unit-length progeny [15,101,103]. There are four types of small ribozyme motifs: hammerhead, hairpin, Varkud satellite (VS), and hepatitis delta virus (HDV) [104]. The glucosamine-6-phosphate riboswitch (GlmS) ribozyme is also classified as a small ribozyme, but in contrast to other catalytically active RNAs, it is activated by the binding of the ligand molecule [105]. All ribozymes catalyze the RNA-cleavage reaction, producing a 5′ hydroxyl group and 2′,3′-cyclic phosphate (Figure 4) [102]. For optimal activity, they require divalent metal ions such as Mg^{2+}. Ribozymes usually act in *cis* mode when a substrate and ribozyme are present in single RNA chains, and cleavage separates RNA into two species (the catalytic domain of the ribozyme and target RNA). Ribozymes can also act in trans, recognizing a substrate by hybridizing to the complementary sequence and cleaving the RNA. Although ribozymes are commonly applied in different research areas, each type has its own specific requirements, mostly concerning the sequence of the RNA substrate (see below).

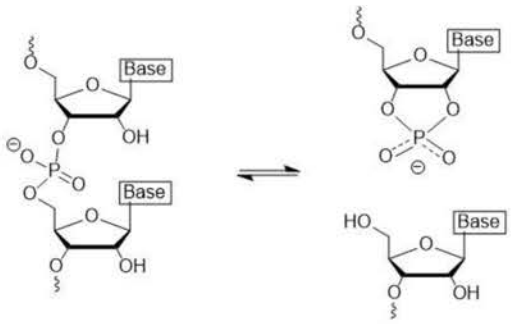


Figure 4. Ribozyme-mediated cleavage of RNA.

Their relatively small size makes ribozymes attractive tools for use in structural studies for the production of RNA molecules with homogeneous ends or for enhancing the efficiency of in vitro transcription [61,63]. In the first case, they are incorporated into a DNA template upstream, downstream, or at both positions of the target sequence. During in vitro transcription, the RNA chain is cleaved, resulting in a mixture of target RNA and the catalytic domain of the ribozyme [61]. They need to be separated and purified using PAGE, gel filtration chromatography, or affinity tags [106,107].

When the ribozyme is located downstream of the target sequence, the 3′-end of the transcript RNA is homogeneous, but it possesses a 2′,3′-cyclic phosphate group (Figure 5A). The phosphate group can be removed with T4 polynucleotide kinase if the 3′ OH group is required for downstream experiments [108]. The ribozyme is placed at the 5′-end of the target RNA to increase the yield of in vitro transcription or to produce RNA with homogeneous 5′-ends (Figure 5B). Higher efficiency in RNA synthesis is achieved by introducing the GGGAGA sequence at the beginning of the catalytic domain of the ribozyme, which ensures high processivity for the RNA polymerase [109]. This approach is also used when the sequence of the target RNA is critical and cannot be adjusted according to the T7 RNA polymerase's requirements. Since homogeneity and high yield in RNA synthesis are key factors for successful crystallization, one should consider RNA constructs with ribozymes at both ends (Figure 5C) [61]. Ribozymes can also be used in trans. This requires the co-transcription of the catalytic domain of the ribozyme and target RNA [110,111].

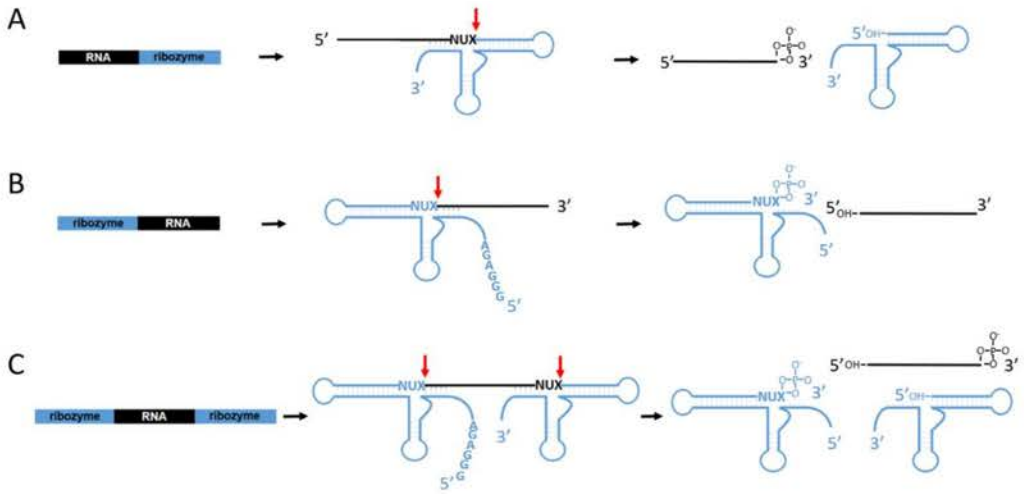


Figure 5. Production of homogeneous RNAs using ribozymes located at (**A**) 5′−end, (**B**) 3′−end, and (**C**) both ends of the target RNA. Cleavage sites are indicated with red arrows.

2.3.1. Hammerhead Ribozyme

The hammerhead ribozyme is one of the most studied catalytic motifs found in viroids and satellite RNAs [112]. It is relatively small (40–50 nucleotides), with a secondary structure consisting of three helices branching from the conserved junction region (Figure 6A). The length of the helical regions can be significantly altered without influencing ribozyme activity (one helix can be reduced to only two base pairs). The overall fold of the ribozyme resembles a Y-shape, and it is stabilized by loop–loop interactions [113,114]. A cleavage site is located at the 3′-end of the target RNA (next to helix I). It occurs after a NUX triplet (N stands for any nucleotide, and X is any nucleotide except guanosine), and the optimal cleavage sequence is AUC, GUC, UUC [110,115]. Another important factor influencing the cleavage efficiency is the presence of divalent cations such as Mg^{2+} or Mn^{2+} at millimolar concentrations. They are involved in the proper folding of the ribozyme and serve as co-factors during the cleavage reaction [116].

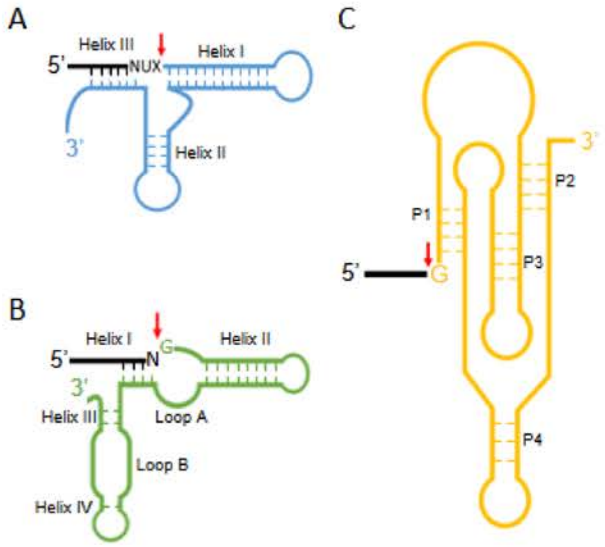


Figure 6. Secondary structure models of (**A**) hammerhead ribozyme, (**B**) hairpin ribozyme, and (**C**) HDV ribozyme. Target RNA is marked with black line. Cleavage sites are indicated with red arrows.

The hammerhead ribozyme has been widely used to generate RNA molecules with homogeneous ends. The efficient cleavage of RNAs ranging from 36 to 140 nt has been observed for ribozymes located at either the 5′ or 3′-end, with optimum activity at a 24 mM concentration of Mg^{2+} ions [61]. The hammerhead ribozyme has been successfully applied in crystallographic studies to synthesize part of the human 7SL RNA (Protein Data Bank—PDB: 1L9A), 4.5S rRNA domain IV (PDB: 1DUH), and tRNAIle (PDB: 6UFH) [117–119].

2.3.2. Hairpin Ribozyme

The hairpin ribozyme was identified in the minus strand of the tobacco ringspot virus satellite RNA (TRSV). It possesses four helical regions and two internal bulges (Figure 6B) [120]. The secondary structure of the hairpin ribozyme resembles the shape of a paperclip, but in three-dimensional space, the helices are coaxially arranged, representing two hairpin structures linked by a four-way junction [121]. Similar to the hammerhead ribozyme, helical regions of hairpin ribozyme can be engineered with different lengths and sequences without influencing the cleavage efficiency, as long as the integrity of the helices is maintained. The cleavage site is located in one of the conserved bulge regions. The optimal cleavage sequence is RYN↓GUC (R is a purine, Y is a pyrimidine, and N is any nucleotide) and occurs between N and G residues [100].

Due to different sequence constraints, the hairpin ribozyme has been proposed as an alternative to the hammerhead ribozyme. Although the cleavage efficiency has been observed to be high, the presence of the hairpin ribozyme leads to the generation of nonspecific products. This activity was shown to be reduced by a higher concentration (30 mM) of Mg^{2+} ions [61].

2.3.3. HDV-like Ribozymes

An HDV ribozyme (approximately 85 nt) is found in the genomic and antigenomic strands of the hepatitis delta virus [122,123]. Its secondary structure consists of four helical regions forming two coaxial stacks, which are linked by two single-stranded junctions (Figure 6C). The tertiary structure of the HDV ribozyme folds into a nested double pseudoknot, constraining the overall shape and forming the active site of the ribozyme [123,124]. The cleavage site is located in helix P1 and typically occurs at a guanosine residue. The HDV ribozyme is active in the presence of divalent ions such as Mg^{2+}, Mn^{2+}, or Ca^{2+} [125]. The sequence upstream of the cleavage site is unrestricted. Therefore, the HDV ribozyme can be applied to generate 3′ homogeneous ends for any RNA. On the other hand, highly structured

target RNAs can negatively influence the proper folding of the ribozyme domain, resulting in lower cleavage efficiency. In this case, incubating it at 65–70 °C, adding urea, or performing cycles of denaturation–renaturation can enhance the cleavage efficiency [126–128].

The lack of a sequence requirement has resulted in the HDV ribozyme being widely used by structural researchers in the preparation of RNAs for crystallization. Examples include domain IV of the ribosomal 4.5S RNA from E. coli (PDB: 1DUH), the nadA riboswitch (PDB: 7D82), and the constitutive decay element (CDE1) from human 3′ untranslated region (PDB: 6XWW) [118,129,130].

2.3.4. VS RNA Ribozyme

The Varkud satellite RNA ribozyme was found in mitochondrial plasmids of certain strains of Neuropsora [104,131]. It is the longest known small ribozyme (150 nucleotides), but its sequence can be reduced to approximately 120 nt without influencing its cleavage activity. The tertiary fold of the VS ribozyme is highly structured. It is composed of six double-stranded regions forming multihelix stacks, connected by three-way junctions and involving "kissing-loop" interactions [132]. Cleavage sites occur between A and G/A/U residues [133]. The VS ribozyme can act in cis and in trans, but only the trans system is suitable for the large-scale preparation of RNAs with homogeneous 3′-ends [110]. One of the examples is the preparation of the isotope-labelled RNA for NMR measurements [134].

2.3.5. GlmS Ribozyme

The GlmS ribozyme is a widespread, 145 nt-long, conserved element in Gram-positive bacteria located upstream of the glmS gene (encoding glutamine/fructose-6-phosphate aminotransferase) [105,135]. It was discovered as a riboswitch, but further characterization revealed that, instead of RNA protection, it induces RNA decay in the presence of metabolites. Thus, GlmS has been classified as a ribozyme activated by the binding of the ligand (glucosamine-6-phosphate, GlcN6P). The secondary structure of the GlmS ribozyme is composed of four double-stranded regions connected by four junctions, while its tertiary fold forms a compact structure with three near-parallel stacks comprising a total of three pseudoknots [105,136]. The cleavage site is located between conserved guanosine and adenosine residues, which are also involved in co-factor binding and the positioning of a scissile phosphate group [105]. GlmS has been successfully used for crystallographic studies of Xrn1-resistant RNAs (xrRNAs) from Flaviviridae (PDB: 4PQV) [137].

In order to test the efficiency of in vitro transcription using ribozymes, we designed and synthesized several constructs differing by location, type of ribozyme, and length and sequence of the target RNA (Table 1 and Supplementary Materials).

Table 1. RNA constructs used in this study.

Construct	Ribozyme	Length of Target RNA
RNA-HH 1	3′ hammerhead	47 nt
RNA-HH 2	3′ hammerhead	30 nt
RNA-HH 3	3′ hammerhead	83 nt
RNA-HH 4	3′ hammerhead	86 nt
RNA-HH 5	3′ hammerhead	113 nt
RNA-HH 6	3′ hammerhead	43 nt
HH-RNA-HH	5′ hammerhead/3′ hammerhead	43 nt
HH-RNA-HP	5′ hammerhead/3′ hairpin	43 nt
RNA-HDV 1	3′ HDV	47 nt
RNA-HDV 2	3′ HDV	51 nt

The RNA-HH 1 to 6 constructs possessed the hammerhead (HH) ribozyme at the 3′-end and HH-RNA-HH hammerhead ribozyme at both ends. The HH-RNA-HP construct had the hammerhead ribozyme at the 5′ and hairpin (HP) ribozyme at the 3′-end, and RNA-HDV was equipped with the HDV-like ribozyme at the 3′-end. The constructs were prepared as chemically synthesized DNA or obtained by polymerase chain reaction (PCR).

They were cloned into plasmid DNA, which was linearized prior to in vitro transcription. All the RNA-HH constructs showed a high cleavage efficiency independent of the sequence of the target RNA (Figure 7A–C). We estimated that more than 95% of the transcripts were cleaved. We added a 5′-end hammerhead ribozyme to the RNA-HH6 (a new construct, named HH-RNA-HH) to increase the yield from in vitro transcription. Surprisingly, the efficiency of the cleavage was decreased to about 50% (Figure 7C, line 2). We observed a band corresponding to an uncleaved intermediate RNA product attached to the 3′ hammerhead ribozyme. This suggests that the presence of the 5′ ribozyme affects the proper folding of 3′ hammerhead, reducing its cleavage activity.

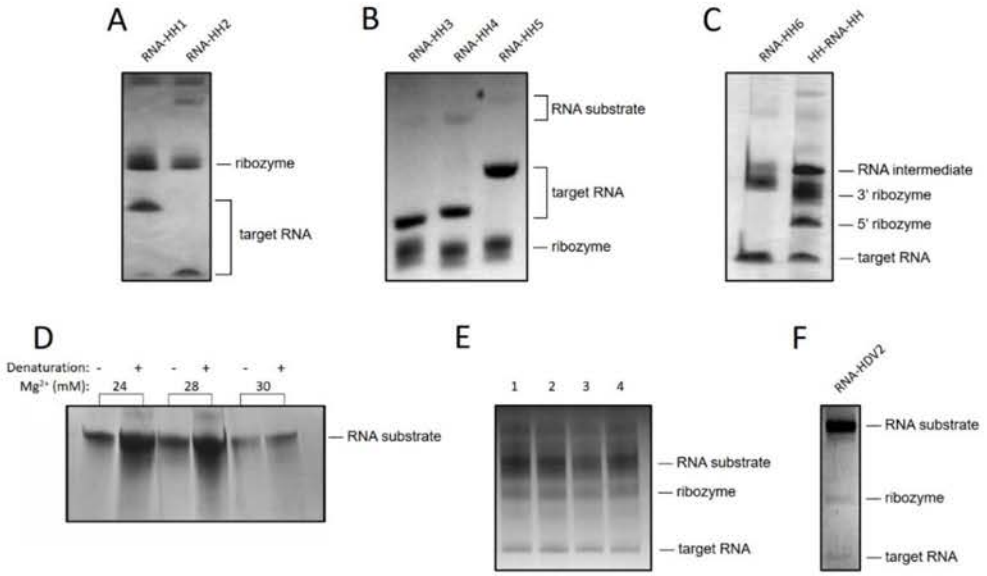


Figure 7. Assessment of ribozyme cleavage efficiency monitored by PAGE. (**A**–**C**) line 1, RNA-HH 1 to 6 constructs; (**C**) lane 2, construct HH-RNA-HH; (**D**) construct HH-RNA-HP (see text for details); (**E**) RNA-HDV 1 (see text for details); (**F**) RNA-HDV 2.

We replaced the 3′ hammerhead with the hairpin ribozyme in order to test whether other types of ribozymes at the 3′-end could increase the yield of the RNA product. However, this modification resulted in no cleavage of the HH-RNA-HP construct by either 5′- or 3′-end-located ribozymes (Figure 7D, lane 1). Optimizing the reaction conditions by increasing the concentration of Mg^{2+} ions or performing cycles of denaturation–renaturation did not improve the cleavage efficiency (Figure 7D, lanes 2–6).

Finally, we tested the RNA-HDV 1 and 2 constructs with the HDV-like ribozyme at the 3′-end. Although we detected RNA products, the cleavage efficiency was very low (Figure 7E, lane 1, and Figure 7F). Incubating the sample at 65 °C for 1 h after in vitro transcription (Figure 7E, lane 2), increasing the Mg^{2+} concentration (Figure 7E, lane 3), or incubating it in 4 M urea (Figure 7E, lane 4) did not improve the cleavage efficiency, suggesting potential misfolding of the catalytic domain of the HDV ribozyme.

2.4. Preparation of RNA with Homogeneous 3′-End Using Modified DNA Template

The alternative method used for reducing RNA 3′-end heterogeneity is the modification of the DNA template for the in vitro transcription reaction. It requires incorporating one or two C2′-methoxy groups into the 5′-end of the antisense strand of the DNA template (Figure 8A) [78]. It was shown that C2′-methoxy modified DNA significantly reduced the N+1 activity of RNA polymerase and increased the amount of proper RNA product compared to transcription carried out with an unmodified template [78,138]. Since the modifications are located on the sugar moiety, there is a low risk of incorporating nucleotides other than those encoded in the sequence. Such 2′-methoxy-modified phosphoramidites are commercially available. The DNA oligomer can be synthesized using an in-house

synthesizer or purchased from external companies. For longer DNA templates, 2′-methoxy modifications can be incorporated during the PCR by using a 2′-methoxy-modified reverse primer [78,138]. Recent reports suggest that the reduction of the side products can be further improved by the addition of DMSO (dimethyl sulphoxide) [139,140]. This modification has been used for synthesis of 5′ leader of the HIV-1 genome for NMR studies [141].

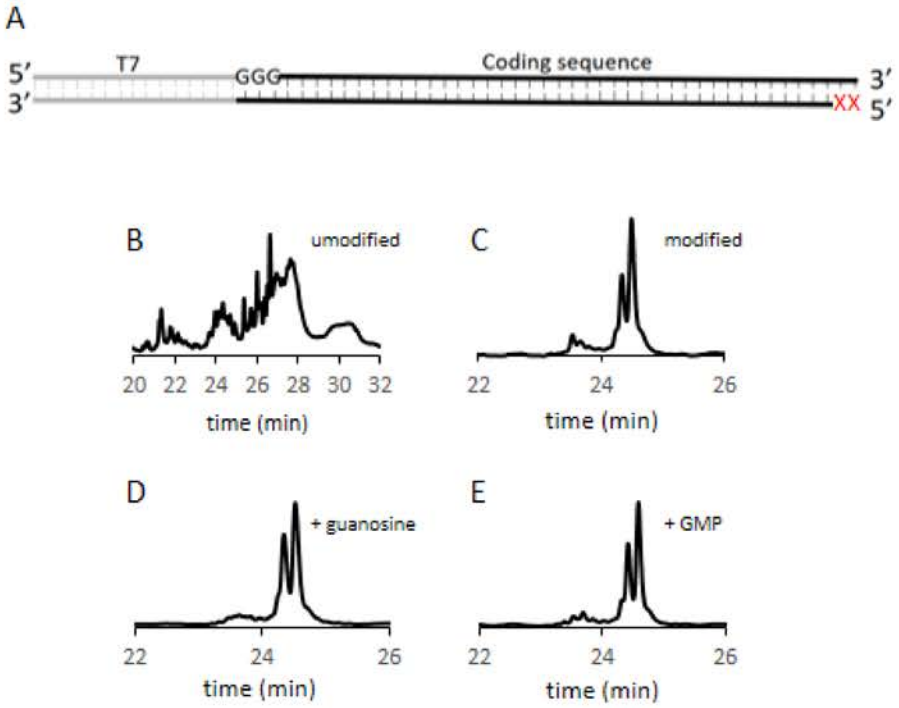


Figure 8. In vitro transcription of RNA using C2′-methoxy-modified DNA. (**A**) Schematic representation of DNA template. Modified nucleotides are marked with red Xs. Chromatograms showing RNA synthesis with (**B**) unmodified template, (**C**) C2′-methoxy-modified template, (**D**) C2′-methoxy-modified template in the presence of 10 mM guanosine, and (**E**) C2′-methoxy-modified template in the presence of 15 mM GTP.

We designed and synthesized a DNA template with two 2′-methoxy-modified nucleotides at the 5′-end of the antisense strand. Next, we performed in vitro transcription of modified and unmodified DNA templates in order to compare the homogeneity of the RNA products. HPLC analysis showed a significant reduction in RNA heterogeneity. With the modified DNA template, only two peaks were observed, and with the unmodified template, RNA products of different lengths were detected (Figure 8B,C).

Using the same method, we obtained RNAs with either a 5′ OH or monophosphate group at the 5′-end by conducting in vitro transcription in the presence of guanosine or GMP, respectively. The HPLC analysis showed no difference in chromatograms compared to the control transcription reactions (Figure 8D,E). The presence of a monophosphate at the 5′-end was further confirmed by the ligation reaction (data not shown).

3. Discussion

This paper provides an overview of methods for large-scale RNA synthesis. We also describe our experience during the preparation of homogeneous samples for crystallization experiments (summarized in Table 2). Many factors should be taken into account when selecting methods for the production of milligram quantities of RNA suitable for structural studies. They include the length of the oligomer, the presence of specific chemical modifications, the oligomer manufacturing time, cost, and equipment availability.

Table 2. Factors that should be considered for the efficient chemical and enzymatic synthesis of RNA.

Method	Important Factors
Chemical synthesis	• Higly efficient for short oligomers (up to 20 nt) • High-grade phosphoramidites (preferably TOM protection) • Acetonitryl with lowest possible content of water (up to 20 ppm) • Preferable synthesis in DMT-ON mode
In vitro transcription	• Concentration of DNA template (50–100 µg/mL) and reaction time (12–16 h) • Hammerhead ribozyme showed highest cleavage efficiency • DNA template with C2′-methoxy modifications reduces 3′-end heterogeneity

From our experience, chemical synthesis allows the preparation of sufficient quantities of RNA up to 20 nucleotides. Longer (>40 nt) oligomers can be synthesized but with much lower efficiency. Chemically synthesized RNA can also be purchased from external companies. However, since crystallization experiments often require testing large numbers of different constructs, using in-house synthesizers significantly reduces the cost of RNA production. We observed that in order to obtain a high yield of chemically synthesized RNA, it is crucial to dissolve reagents in acetonitrile with the lowest possible water content (up to 20 ppm). We recommend TOM rather than TBDMS protected phosphoramidites. They are characterized by higher coupling efficiency due to lower steric hindrance and prevented relocation from 2′ to 3′ moiety of ribose ring. The 2′ to 3′ migration can take place when using the TBDMS group, leading to the formation of non-biologically active 2′-5′ linkages. The costs of TOM and TBDMS phosphoramidites are comparable. When modified phosphoramidites are produced in-house they should have the highest possible purity, because some H-phosphonate or other impurities can still be present, reducing the coupling efficiency. RNA oligomers should be synthesized with the DMT-ON method, which significantly increases the homogeneity of the sample. DMT-ON purification on dedicated columns can remove abortive products, and the resulting RNA is suitable for crystallization or nuclear magnetic resonance measurements (Figure 2). With longer oligomers, additional purification steps are usually required to obtain homogeneous RNA.

In vitro transcription is suitable for longer RNA, but its main drawback is the non-specific activity of the RNA polymerase generating 3′-end heterogeneity. Homogeneous samples can be achieved by using ribozymes cleaving at the 3′ site of the RNA construct. In our hands, the hammerhead ribozyme was characterized by the highest (~90%) cleavage efficiency, while the HDV-like ribozyme cleaved RNA less efficiently (Figure 7). Despite the careful design of RNA constructs to fulfill all the requirements for the proper folding of the catalytic domains of ribozymes, the cleavage efficiency was low when ribozymes were present at both ends of the RNA. An extreme example is the HP-RNA-HH construct, where the replacement of the 5′ hammerhead with the hairpin ribozyme led to the complete inhibition of cleavage (Figure 7). The presence of the 5′ hairpin ribozyme could interfere with the folding of the 3′ hammerhead ribozyme. As a result, both ribozymes became inactive. The ribozyme approach requires the purification of the RNA product from uncleaved RNA and the catalytic domain of the ribozyme. Therefore, the lengths of the ribozyme and the RNA of interest should differ in such a way that they can be separated with gel electrophoresis or HPLC. An attractive alternative to the ribozyme approach is in vitro transcription of the C2′-methoxy-modified DNA. This does not have specific requirements for the proper folding of catalytic RNAs, and a DNA template can be obtained from external companies or synthesized in-house. Using this method, we observed significant improvement in terms of the RNA's 3′-end homogeneity (Figure 8). Although N+1 activity of RNA polymerase was still present, RNAs were easily separated using HPLC or gel electrophoresis (data not shown).

Concerning factors influencing the yield of RNA transcription, we noted that the most important were DNA template concentration (50–100 µg/mL) and reaction incubation time (12–16 h). Although the adjusting concentration of magnesium ions and ribonucleotides often improves the yield of transcription, we did not observe significant differences in the amount of synthesized RNA (data not shown) [142].

4. Conclusions

RNA is involved in the regulation of almost all aspects of cell physiology. Structural richness of RNA molecules has been explored, not only by nature but also adapted by humans in the construction of tools for molecular biology, biomedicine and nanotechnology. Therefore, investigation of RNA three-dimensional structures is fundamentally important and is a key to understanding the RNA-function relationship.

Here, we reviewed well-established methods for large-scale production of homogenous RNA with emphasis on chemical synthesis and in vitro transcription. We also present the results of the incorporation of these methods to obtain RNA suitable for crystallographic experiments. Although preparation of such RNA is still a serious bottleneck, chemical synthesis and in vitro transcription provide simple and effective ways of sample production.

In recent years, a number of new methods for large-scale RNA synthesis have emerged, such as circular RNAs, and position-selective labeling of RNA and polymerase chain transcription (for review see [62,143]). Although still under development, they show new research avenues with great potential to open up the field of structural studies of RNA.

We anticipate that this paper will be useful for many researchers. It will help in choosing a method for large-scale RNA production depending on the chemical nature of the target molecule and avoiding difficulties during sample preparation.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/app12031543/s1, Materials and Methods (DNA /RNA synthesis and purification, in vitro transcription and HPLC analysis).

Author Contributions: Conceptualization, M.R., A.K. and L.B.; methodology, M.R., M.P., L.B. and A.K.; investigation, M.R, M.P. and L.B.; writing—original draft preparation, M.R, L.B and A.K.; writing—review and editing, A.K and L.B.; visualization, M.R.; supervision, A.K.; project administration, A.K.; funding acquisition, A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science Centre, Poland, grant number UMO-2017/26/E/NZ1/00950.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Sharp, P.A. The Centrality of RNA. *Cell* **2009**, *136*, 577–580. [CrossRef] [PubMed]
2. Michelini, F.; Jalihal, A.P.; Francia, S.; Meers, C.; Neeb, Z.T.; Rossiello, F.; Gioia, U.; Aguado, J.; Jones-Weinert, C.; Luke, B.; et al. From "Cellular" RNA to "Smart" RNA: Multiple Roles of RNA in Genome Stability and Beyond. *Chem. Rev.* **2018**, *118*, 4365–4403. [CrossRef] [PubMed]
3. Strobel, E.J.; Watters, K.E.; Loughrey, D.; Lucks, J.B. RNA systems biology: Uniting functional discoveries and structural tools to understand global roles of RNAs. *Curr. Opin. Biotechnol.* **2016**, *39*, 182–191. [CrossRef]
4. Frixione, E.; Ruiz-Zamarripa, L. The "scientific catastrophe" in nucleic acids research that boosted molecular biology. *J. Biol. Chem.* **2019**, *294*, 2249–2255. [CrossRef] [PubMed]
5. Levene, P.A.; Tipson, R.S. The ring structure of thymidine. *Science* **1935**, *81*, 98. [CrossRef]
6. Levene, P.A. On the biochemistry of nucleic acids. *J. Am. Chem. Soc.* **1910**, *32*, 231–240. [CrossRef]
7. Kossel, A.; Neumann, A. Ueber das Thymin, ein Spaltungsproduct der Nucleïnsäure. *Berichte Dtsch. Chem. Gesellschaft* **1893**, *26*, 2753–2756. [CrossRef]

8. Jones, C.P.; Ferré-D'Amaré, A.R. Long-Range Interactions in Riboswitch Control of Gene Expression. *Annu. Rev. Biophys.* **2017**, *46*, 455–481. [CrossRef]
9. Winkler, W.; Nahvi, A.; Breaker, R.R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **2002**, *419*, 952–956. [CrossRef]
10. Kruger, K.; Grabowski, P.J.; Zaug, A.J.; Sands, J.; Gottschling, D.E.; Cech, T.R. Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell* **1982**, *31*, 147–157. [CrossRef]
11. Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **1983**, *35*, 849–857. [CrossRef]
12. Statello, L.; Guo, C.-J.; Chen, L.-L.; Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 96–118. [CrossRef] [PubMed]
13. Hendrix, D.K.; Brenner, S.E.; Holbrook, S.R. RNA structural motifs: Building blocks of a modular biomolecule. *Q. Rev. Biophys.* **2005**, *38*, 221–243. [CrossRef] [PubMed]
14. Leontis, N.B.; Lescoute, A.; Westhof, E. The building blocks and motifs of RNA architecture. *Curr. Opin. Struct. Biol.* **2006**, *16*, 279–287. [CrossRef]
15. Serganov, A.; Patel, D.J. Ribozymes, riboswitches and beyond: Regulation of gene expression without proteins. *Nat. Rev. Genet.* **2007**, *8*, 776–790. [CrossRef] [PubMed]
16. Stavast, C.; Erkeland, S. The Non-Canonical Aspects of MicroRNAs: Many Roads to Gene Regulation. *Cells* **2019**, *8*, 1465. [CrossRef]
17. Malik, I.; Kelley, C.P.; Wang, E.T.; Todd, P.K. Molecular mechanisms underlying nucleotide repeat expansion disorders. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 589–607. [CrossRef] [PubMed]
18. Esteller, M. Non-coding RNAs in human disease. *Nat. Rev. Genet.* **2011**, *12*, 861–874. [CrossRef] [PubMed]
19. Zhang, Y.-Z.; Wu, W.-C.; Shi, M.; Holmes, E.C. The diversity, evolution and origins of vertebrate RNA viruses. *Curr. Opin. Virol.* **2018**, *31*, 9–16. [CrossRef]
20. Lundstrom, K. RNA Viruses as Tools in Gene Therapy and Vaccine Development. *Genes* **2019**, *10*, 189. [CrossRef]
21. Schmidt, C.M.; Smolke, C.D. RNA Switches for Synthetic Biology. *Cold Spring Harb. Perspect. Biol.* **2019**, *11*, a032532. [CrossRef] [PubMed]
22. Ouellet, J. RNA Fluorescence with Light-Up Aptamers. *Front. Chem.* **2016**, *4*, 1–12. [CrossRef] [PubMed]
23. Pfeiffer, F.; Mayer, G. Selection and Biosensor Application of Aptamers for Small Molecules. *Front. Chem.* **2016**, *4*, 25. [CrossRef]
24. Guo, P. The emerging field of RNA nanotechnology. *Nat. Nanotechnol.* **2010**, *5*, 833–842. [CrossRef]
25. Bennett, C.F.; Krainer, A.R.; Cleveland, D.W. Antisense Oligonucleotide Therapies for Neurodegenerative Diseases. *Annu. Rev. Neurosci.* **2019**, *42*, 385–406. [CrossRef]
26. Meo, S.A.; Bukhari, I.A.; Akram, J.; Meo, A.S.; Klonoff, D.C. COVID-19 vaccines: Comparison of biological, pharmacological characteristics and adverse effects of pfizer/BioNTech and moderna vaccines. *Eur. Rev. Med. Pharmacol. Sci.* **2021**, *25*, 1663–1679. [CrossRef] [PubMed]
27. Warner, K.D.; Hajdin, C.E.; Weeks, K.M. Principles for targeting RNA with drug-like small molecules. *Nat. Rev. Drug Discov.* **2018**, *17*, 547–558. [CrossRef] [PubMed]
28. Hermann, T. Small molecules targeting viral RNA. *Wiley Interdiscip. Rev. RNA* **2016**, *7*, 726–743. [CrossRef] [PubMed]
29. Watson, J.D.; Crick, F.H.C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171*, 737–738. [CrossRef] [PubMed]
30. Franklin, R.E.; Gosling, R.G. Molecular Configuration in Sodium Thymonucleate. *Nature* **1953**, *171*, 740–741. [CrossRef] [PubMed]
31. Clark, B.F.C. The crystal structure of tRNA. *J. Biosci.* **2006**, *31*, 453–457. [CrossRef] [PubMed]
32. Holley, R.W.; Apgar, J.; Everett, G.A.; Madison, J.T.; Marquisee, M.; Merrill, S.H.; Penswick, J.R.; Zamir, A. Structure of a ribonucleic acid. *Science* **1965**, *147*, 1462–1465. [CrossRef] [PubMed]
33. Rich, A. The Era of RNA Awakening: Structural biology of RNA in the early years. *Q. Rev. Biophys.* **2009**, *42*, 117–137. [CrossRef] [PubMed]
34. Pley, H.W.; Flaherty, K.M.; McKay, D.B. Three-dimensional structure of a hammerhead ribozyme. *Nature* **1994**, *372*, 68–74. [CrossRef] [PubMed]
35. Ren, A.; Micura, R.; Patel, D.J. Structure-based mechanistic insights into catalysis by small self-cleaving ribozymes. *Curr. Opin. Chem. Biol.* **2017**, *41*, 71–83. [CrossRef]
36. Klein, D.J.; Edwards, T.E.; Ferré-D'Amaré, A.R. Cocrystal structure of a class I preQ1 riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase. *Nat. Struct. Mol. Biol.* **2009**, *16*, 343–344. [CrossRef]
37. Cate, J.H.; Yusupov, M.M.; Yusupova, G.Z.; Earnest, T.N.; Noller, H.F. X-ray crystal structures of 70S ribosome functional complexes. *Science* **1999**, *285*, 2095–2104. [CrossRef]
38. Wimberly, B.T.; Brodersen, D.E.; Clemons, W.M.; Morgan-Warren, R.J.; Carter, A.P.; Vonrhein, C.; Hartsch, T.; Ramakrishnan, V. Structure of the 30S ribosomal subunit. *Nature* **2000**, *407*, 327–339. [CrossRef]
39. Ban, N.; Nissen, P.; Hansen, J.; Moore, P.B.; Steitz, T.A. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **2000**, *289*, 905–920. [CrossRef]
40. Schluenzen, F.; Tocilj, A.; Zarivach, R.; Harms, J.; Gluehmann, M.; Janell, D.; Bashan, A.; Bartels, H.; Agmon, I.; Franceschi, F.; et al. Structure of Functionally Activated Small Ribosomal Subunit at 3.3 Å Resolution. *Cell* **2000**, *102*, 615–623. [CrossRef]

41. Yusupov, M.M.; Yusupova, G.Z.; Baucom, A.; Lieberman, K.; Earnest, T.N.; Cate, J.H.D.; Noller, H.F. Crystal structure of the ribosome at 5.5 Å resolution. *Science* **2001**, *292*, 883–896. [CrossRef] [PubMed]
42. Oubridge, C.; Ito, N.; Evans, P.R.; Teo, C.H.; Nagai, K. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **1994**, *372*, 432–438. [CrossRef] [PubMed]
43. Allain, F.H.T.; Howe, P.W.A.; Neuhaus, D.; Varani, G. Structural basis of the RNA-binding specificity of human U1A protein. *EMBO J.* **1997**, *16*, 5764–5774. [CrossRef] [PubMed]
44. Bonilla, S.L.; Sherlock, M.E.; MacFadden, A.; Kieft, J.S. A viral RNA hijacks host machinery using dynamic conformational changes of a tRNA-like structure. *Science* **2021**, *374*, 955–960. [CrossRef] [PubMed]
45. Akiyama, B.M.; Laurence, H.M.; Massey, A.R.; Costantino, D.A.; Xie, X.; Yang, Y.; Shi, P.Y.; Nix, J.C.; Beckham, J.D.; Kieft, J.S. Zika virus produces noncoding RNAs using a multi-pseudoknot structure that confounds a cellular exonuclease. *Science* **2016**, *354*, 1148–1152. [CrossRef] [PubMed]
46. Cate, J.H.; Gooding, A.R.; Podell, E.; Zhou, K.; Golden, B.L.; Kundrot, C.E.; Cech, T.R.; Doudna, J.A. Crystal structure of a group I ribozyme domain: Principles of RNA packing. *Science* **1996**, *273*, 1678–1685. [CrossRef]
47. Imai, S.; Kumar, P.; Hellen, C.U.T.; D'Souza, V.M.; Wagner, G. An accurately preorganized IRES RNA structure enables eIF4G capture for initiation of viral translation. *Nat. Struct. Mol. Biol.* **2016**, *23*, 859–864. [CrossRef]
48. Brown, J.A.; Bulkley, D.; Wang, J.; Valenstein, M.L.; Yario, T.A.; Steitz, T.A.; Steitz, J.A. Structural insights into the stabilization of MALAT1 noncoding RNA by a bipartite triple helix. *Nat. Struct. Mol. Biol.* **2014**, *21*, 633–640. [CrossRef] [PubMed]
49. Błaszczyk, L.; Rypniewski, W.; Kiliszek, A. Structures of RNA repeats associated with neurological diseases. *Wiley Interdiscip. Rev. RNA* **2017**, *8*, e1412. [CrossRef]
50. Ferré-D'Amaré, A.R.; Zhou, K.; Doudna, J.A. Crystal structure of a hepatitis delta virus ribozyme. *Nature* **1998**, *395*, 567–574. [CrossRef]
51. Golden, B.L.; Gooding, A.R.; Podell, E.R.; Cech, T.R. A preorganized active site in the crystal structure of the Tetrahymena ribozyme. *Science* **1998**, *282*, 259–264. [CrossRef] [PubMed]
52. Serganov, A.; Yuan, Y.R.; Pikovskaya, O.; Polonskaia, A.; Malinina, L.; Phan, A.T.; Hobartner, C.; Micura, R.; Breaker, R.R.; Patel, D.J. Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chem. Biol.* **2004**, *11*, 1729–1741. [CrossRef] [PubMed]
53. Thore, S.; Leibundgut, M.; Ban, N. Structure of the Eukaryotic Thiamine Pyrophosphate Riboswitch with Its Regulatory Ligand. *Science* **2006**, *312*, 1208–1211. [CrossRef] [PubMed]
54. Serganov, A.; Polonskaia, A.; Phan, A.T.; Breaker, R.R.; Patel, D.J. Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature* **2006**, *441*, 1167–1171. [CrossRef] [PubMed]
55. Roth, A.; Breaker, R.R. The Structural and Functional Diversity of Metabolite-Binding Riboswitches. *Annu. Rev. Biochem.* **2009**, *78*, 305–334. [CrossRef]
56. Duchardt-Ferner, E.; Gottstein-Schmidtke, S.R.; Weigand, J.E.; Ohlenschläger, O.; Wurm, J.P.; Hammann, C.; Suess, B.; Wöhnert, J. What a Difference an OH Makes: Conformational Dynamics as the Basis for the Ligand Specificity of the Neomycin-Sensing Riboswitch. *Angew. Chem. Int. Ed.* **2016**, *55*, 1527–1530. [CrossRef] [PubMed]
57. Westhof, E. Twenty years of RNA crystallography. *RNA* **2015**, *21*, 486–487. [CrossRef] [PubMed]
58. Holbrook*, S.R.; Holbrook, E.L.; Walukiewicz, H.E. Crystallization of RNA. *Cell. Mol. Life Sci.* **2001**, *58*, 234–243. [CrossRef] [PubMed]
59. Golden, B.L.; Kundrot, C.E. RNA crystallization. *J. Struct. Biol.* **2003**, *142*, 98–107. [CrossRef]
60. Ferré-D'Amaré, A.R.; Doudna, J.A. Methods to Crystallize RNA. *Curr. Protoc. Nucleic Acid Chem.* **2000**, *1*, 1–13. [CrossRef]
61. Price, S.R.; Ito, N.; Oubridge, C.; Avis, J.M.; Nagai, K. Crystallization of RNA-protein complexes I. Methods for the large-scale preparation of RNA suitable for crystallographic studies. *J. Mol. Biol.* **1995**, *249*, 398–408. [CrossRef] [PubMed]
62. Baronti, L.; Karlsson, H.; Marušič, M.; Petzold, K. A guide to large-scale RNA sample preparation. *Anal. Bioanal. Chem.* **2018**, *410*, 3239–3252. [CrossRef] [PubMed]
63. Ahmed, Y.L.; Ficner, R. RNA synthesis and purification for structural studies. *RNA Biol.* **2014**, *11*, 427–432. [CrossRef] [PubMed]
64. Merrifield, R.B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154. [CrossRef]
65. Josephson, S.; Lagerholm, E.; Palm, G.; Mörch, L.; Norin, T. Automatic Synthesis of Oligodeoxynucleotides and Mixed Oligodeoxynucleotides using the Phosphoamidite Method. *Acta Chem. Scand.* **1984**, *38*, 539–545. [CrossRef] [PubMed]
66. Flamme, M.; McKenzie, L.K.; Sarac, I.; Hollenstein, M. Chemical methods for the modification of RNA. *Methods* **2019**, *161*, 64–82. [CrossRef]
67. Bartosik, K.; Debiec, K.; Czarnecka, A.; Sochacka, E.; Leszczynska, G. Synthesis of Nucleobase-Modified RNA Oligonucleotides by Post-Synthetic Approach. *Molecules* **2020**, *25*, 3344. [CrossRef]
68. Somoza, Á. Protecting groups for RNA synthesis: An increasing need for selective preparative methods. *Chem. Soc. Rev.* **2008**, *37*, 2668. [CrossRef]
69. Beaucage, S.L.; Reese, C.B. Recent Advances in the Chemical Synthesis of RNA. *Curr. Protoc. Nucleic Acid Chem.* **2009**, *38*, 1–31. [CrossRef]
70. Sekine, M. Recent development of chemical synthesis of RNA. In *Synthesis of Therapeutic Oligonucleotides*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 41–65, ISBN 9789811319129.
71. Reese, C.B. Oligo- and poly-nucleotides: 50 years of chemical synthesis. *Org. Biomol. Chem.* **2005**, *3*, 3851–3868. [CrossRef]

72. Moore, M.; Charles, Q. Use of site-specifically modified RNAs constructed by RNA ligation. In *RNA: Protein Interactions. A Practical Approach*; Smith, C., Ed.; Oxford University Press: Oxford, UK, 1998; pp. 75–108, ISBN 0-19-963651-6.
73. Höbartner, C.; Micura, R. Chemical Synthesis of Selenium-Modified Oligoribonucleotides and Their Enzymatic Ligation Leading to an U6 SnRNA Stem-Loop Segment. *J. Am. Chem. Soc.* **2004**, *126*, 1141–1149. [CrossRef] [PubMed]
74. Kiliszek, A.; Błaszczyk, L.; Kierzek, R.; Rypniewski, W. Stabilization of RNA hairpins using non-nucleotide linkers and circularization. *Nucleic Acids Res.* **2017**, *45*, e92. [CrossRef]
75. Beckert, B.; Masquida, B. Synthesis of RNA by In Vitro Transcription. In *RNA. Methods in Molecular Biology (Methods and Protocols)*; Nielsen, H., Ed.; Humana Press: Totowa, NJ, USA, 2011; Volume 703, ISBN 978-1-58829-913-0.
76. Coleman, T.M.; Wang, G.; Huang, F. Superior 5′ homogeneity of RNA from ATP-initiated transcription under the T7 phi 2.5 promoter. *Nucleic Acids Res.* **2004**, *32*, 2–5. [CrossRef] [PubMed]
77. Dunn, J.J.; Studier, F.W. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.* **1983**, *166*, 477–535. [CrossRef]
78. Kao, C.; Zheng, M.; Rüdisser, S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3′ terminus of RNAs transcribed by T7 RNA polymerase. *RNA* **1999**, *5*, 1268–1272. [CrossRef] [PubMed]
79. Pujari, N.; Saundh, S.L.; Acquah, F.A.; Mooers, B.H.M.; Ferré-D'Amaré, A.R.; Leung, A.K.-W. Engineering Crystal Packing in RNA Structures I: Past and Future Strategies for Engineering RNA Packing in Crystals. *Crystals* **2021**, *11*, 952. [CrossRef] [PubMed]
80. Shatkin, A. Capping of eucaryotic mRNAs. *Cell* **1976**, *9*, 645–653. [CrossRef]
81. Ramanathan, A.; Robb, G.B.; Chan, S.H. mRNA capping: Biological functions and applications. *Nucleic Acids Res.* **2016**, *44*, 7511–7526. [CrossRef]
82. Galloway, A.; Cowling, V.H. mRNA cap regulation in mammalian cell function and fate. *Biochim. Biophys. Acta Gene Regul. Mech.* **2019**, *1862*, 270–279. [CrossRef]
83. Pardi, N.; Hogan, M.J.; Porter, F.W.; Weissman, D. mRNA vaccines—A new era in vaccinology. *Nat. Rev. Drug Discov.* **2018**, *17*, 261–279. [CrossRef]
84. Hodel, A.E.; Gershon, P.D.; Quiocho, F.A. Structural Basis for Sequence-Nonspecific Recognition of 5′-Capped mRNA by a Cap-Modifying Enzyme. *Mol. Cell* **1998**, *1*, 443–447. [CrossRef]
85. Shirokikh, N.E.; Preiss, T. Translation initiation by cap-dependent ribosome recruitment: Recent insights and open questions. *WIREs RNA* **2018**, *9*, 1–45. [CrossRef] [PubMed]
86. Grudzien-Nogalska, E.; Stepinski, J.; Jemielity, J.; Zuberek, J.; Stolarski, R.; Rhoads, R.E.; Darzynkiewicz, E. Synthesis of Anti-Reverse Cap Analogs (ARCAs) and their Applications in mRNA Translation and Stability. *Methods Enzymol.* **2007**, *431*, 203–227. [CrossRef] [PubMed]
87. Kim, J.; Eygeris, Y.; Gupta, M.; Sahay, G. Self-assembled mRNA vaccines. *Adv. Drug Deliv. Rev.* **2021**, *170*, 83–112. [CrossRef] [PubMed]
88. Pelletier, J.; Schmeing, T.M.; Sonenberg, N. The multifaceted eukaryotic cap structure. *Wiley Interdiscip. Rev. RNA* **2021**, *12*, e1636. [CrossRef] [PubMed]
89. Ensinger, M.J.; Martin, S.A.; Paoletti, E.; Moss, B. Modification of the 5′-terminus of mRNA by soluble guanylyl and methyl transferases from vaccinia virus. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 2525–2529. [CrossRef]
90. Yisraeli, J.K.; Melton, D.A. Synthesis of long, capped transcripts in Vitro by SP6 and T7 RNA polymerases. *Methods Enzymol.* **1989**, *180*, 42–50.
91. Nielsen, D.A.; Shapiro, D.J. Preparation of capped RNA transcripts using T7 RNA polymerase. *Nucleic Acids Res.* **1986**, *14*, 5936. [CrossRef]
92. Konarska, M.M.; Padgett, R.A.; Sharp, P.A. Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell* **1984**, *38*, 731–736. [CrossRef]
93. Jemielity, J.; Fowler, T.; Zuberek, J.; Stepinski, J.; Lewdorowicz, M.; Niedzwiecka, A.; Stolarski, R.; Darzynkiewicz, E.; Rhoads, R.E. Novel "anti-reverse" cap analogs with superior translational properties. *RNA* **2003**, *9*, 1108–1122. [CrossRef]
94. Stepinski, J.; Waddell, C.; Stolarski, R.; Darzynkiewicz, E.; Rhoads, R.E. Synthesis and properties of mRNAs containing the novel " anti-reverse " cap analogs. *RNA* **2001**, *7*, 1486–1495. [CrossRef] [PubMed]
95. Sikorski, P.J.; Warminski, M.; Kubacka, D.; Ratajczak, T.; Nowis, D.; Kowalska, J.; Jemielity, J. The identity and methylation status of the first transcribed nucleotide in eukaryotic mRNA 5′ cap modulates protein expression in living cells. *Nucleic Acids Res.* **2020**, *48*, 1607–1626. [CrossRef] [PubMed]
96. Henderson, J.M.; Ujita, A.; Hill, E.; Yousif-Rosales, S.; Smith, C.; Ko, N.; McReynolds, T.; Cabral, C.R.; Escamilla-Powers, J.R.; Houston, M.E. Cap 1 Messenger RNA Synthesis with Co-transcriptional CleanCap® Analog by In Vitro Transcription. *Curr. Protoc.* **2021**, *1*, 1–17. [CrossRef] [PubMed]
97. Thillier, Y.; Decroly, E.; Morvan, F.; Canard, B.; Vasseur, J.-J.; Debart, F. Synthesis of 5′ cap-0 and cap-1 RNAs using solid-phase chemistry coupled with enzymatic methylation by human (guanine- N 7)-methyl transferase. *RNA* **2012**, *18*, 856–868. [CrossRef] [PubMed]
98. Jemielity, J.; Heinonen, P.; Lönnberg, H.; Darzynkiewicz, E. A novel approach to solid phase chemical synthesis of Oligonucleotides mRNA CAP analogs. *Nucleosides Nucleotides Nucleic Acids* **2005**, *24*, 601–605. [CrossRef]
99. Kadokura, M.; Wada, T.; Seio, K.; Moriguchi, T.; Huber, J.; Lührmann, R.; Sekine, M. Solid-phase synthesis of a 5′-terminal TMG-capped trinucleotide block of U1 snRNA. *Tetrahedron Lett.* **2001**, *42*, 8853–8856. [CrossRef]

100. Tanner, N.K. Ribozymes: The characteristics and properties of catalytic RNAs. *FEMS Microbiol. Rev.* **1999**, *23*, 257–275. [CrossRef]
101. Weinberg, C.E.; Weinberg, Z.; Hammann, C. Novel ribozymes: Discovery, catalytic mechanisms, and the quest to understand biological function. *Nucleic Acids Res.* **2019**, *47*, 9480–9494. [CrossRef]
102. Doherty, E.A.; Doudna, J.A. Ribozyme Structures and Mechanisms. *Annu. Rev. Biochem.* **2000**, *69*, 597–615. [CrossRef]
103. Müller, S.; Appel, B.; Balke, D.; Hieronymus, R.; Nübel, C. Thirty-five years of research into ribozymes and nucleic acid catalysis: Where do we stand today? *F1000Research* **2016**, *5*, 1511. [CrossRef]
104. Maurel, M.-C.; Leclerc, F.; Hervé, G. Ribozyme Chemistry: To Be or Not To Be under High Pressure. *Chem. Rev. Am. Chem. Soc.* **2020**, *120*, 4898–4918. [CrossRef] [PubMed]
105. Ferré-D'Amaré, A.R. The glmSribozyme: Use of a small molecule coenzyme by a gene-regulatory RNA. *Q. Rev. Biophys.* **2010**, *43*, 423–447. [CrossRef] [PubMed]
106. Kieft, J.S.; Batey, R.T. A general method for rapid and nondenaturing purification of RNAs. *RNA* **2004**, *10*, 988–995. [CrossRef] [PubMed]
107. Salvail-Lacoste, A.; Di Tomasso, G.; Piette, B.L.; Legault, P. Affinity purification of T7 RNA transcripts with homogeneous ends using ARiBo and CRISPR tags. *RNA* **2013**, *19*, 1003–1014. [CrossRef] [PubMed]
108. Schürer, H.; Lang, K.; Schuster, J.; Mörl, M. A universal method to produce in vitro transcripts with homogeneous 3′ ends. *Nucleic Acids Res.* **2002**, *30*, e56. [CrossRef] [PubMed]
109. Morl, M.; Hartmann, R.K. Production of RNAs with Homogeneous 5′ and 3′ Ends. In *Handbook of RNA Biochemistry*; Wiley-VCH: Weinheim, Germany, 2008; pp. 22–35, ISBN 3527308261.
110. Ferré-D'Amaré, A.R.; Doudna, J.A. Use of cis- and trans-ribozymes to remove 5′ and 3′ heterogeneities from milligrams of in vitro transcribed RNA. *Nucleic Acids Res.* **1996**, *24*, 977–978. [CrossRef]
111. Wichlacz, A.; Łęgiewicz, M.; Ciesiołka, J. Generating in vitro transcripts with homogenous 3′ ends using trans-acting antigenomic delta ribozyme. *Nucleic Acids Res.* **2004**, *32*, e39. [CrossRef] [PubMed]
112. De La Peña, M.; García-Robles, I.; Cervera, A. The hammerhead Ribozyme: A long history for a short RNA. *Molecules* **2017**, *22*, 78. [CrossRef]
113. De La Peña, M.; Dufour, D.; Gallego, J. Three-way RNA junctions with remote tertiary contacts: A recurrent and highly versatile fold. *RNA* **2009**, *15*, 1949–1964. [CrossRef]
114. Martick, M.; Scott, W.G. Tertiary Contacts Distant from the Active Site Prime a Ribozyme for Catalysis. *Cell* **2006**, *126*, 309–320. [CrossRef]
115. Amarzguioui, M.; Prydz, H. Hammerhead ribozyme design and application. *Cell. Mol. Life Sci.* **1998**, *54*, 1175–1202. [CrossRef] [PubMed]
116. Scott, W.G.; Horan, L.H.; Martick, M. *The Hammerhead Ribozyme: Structure, Catalysis, and Gene Regulation*, 1st ed.; Elsevier: Amsterdam, The Netherlands, 2013; Volume 120, ISBN 9780123812865.
117. Oubridge, C.; Kuglstatter, A.; Jovine, L.; Nagai, K. Crystal structure of SRP19 in complex with the S domain of SRP RNA and its implication for the assembly of the signal recognition particle. *Mol. Cell* **2002**, *9*, 1251–1261. [CrossRef]
118. Jovine, L.; Hainzl, T.; Oubridge, C.; Scott, W.G.; Li, J.; Sixma, T.K.; Wonacott, A.; Skarzynski, T.; Nagai, K. Crystal structure of the Ffh and EF-G binding sites in the conserved domain IV of Escherichia coli 4.5S RNA. *Structure* **2000**, *8*, 527–540. [CrossRef]
119. Battaglia, R.A.; Grigg, J.C.; Ke, A. Structural basis for tRNA decoding and aminoacylation sensing by T-box riboregulators. *Nat. Struct. Mol. Biol.* **2019**, *26*, 1106–1113. [CrossRef] [PubMed]
120. Walter, N.G.; Burke, J.M. The hairpin ribozyme: Structure, assembly and catalysis. *Curr. Opin. Chem. Biol.* **1998**, *2*, 24–30. [CrossRef]
121. Rupert, P.B.; Ferré-D'Amaré, A.R. Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature* **2001**, *410*, 780–786. [CrossRef] [PubMed]
122. Webb, C.-H.T.; Lupták, A. HDV-like self-cleaving ribozymes. *RNA Biol.* **2011**, *8*, 719–727. [CrossRef]
123. Tanner, N.K.; Schaff, S.; Thill, G.; Petit-Koskas, E.; Crain-Denoyelle, A.M.; Westhof, E. A three-dimensional model of hepatitis delta virus ribozyme based on biochemical and mutational analyses. *Curr. Biol.* **1994**, *4*, 488–498. [CrossRef]
124. Chen, J.-H.; Yajima, R.; Chadalavada, D.M.; Chase, E.; Bevilacqua, P.C.; Golden, B.L. A 1.9 Å Crystal Structure of the HDV Ribozyme Precleavage Suggests both Lewis Acid and General Acid Mechanisms Contribute to Phosphodiester Cleavage. *Biochemistry* **2010**, *49*, 6508–6518. [CrossRef]
125. Suh, Y.A.; Kumar, P.K.R.; Taira, K.; Nishikawa, S. Self-cleavage activity of the genomic HDV ribozyme in the presence of various divalent metal ions. *Nucleic Acids Res.* **1993**, *21*, 3277–3280. [CrossRef]
126. Doudna, J.A. Preparation of homogeneous ribozyme RNA for crystallization. *Methods Mol. Biol.* **1997**, *74*, 365–370. [CrossRef] [PubMed]
127. Peselis, A.; Gao, A.; Serganov, A. Preparation and crystallization of riboswitches. *Methods Mol. Biol.* **2016**, *1320*, 21–36. [CrossRef] [PubMed]
128. Steckelberg, A.L.; Akiyama, B.M.; Costantino, D.A.; Sit, T.L.; Nix, J.C.; Kieft, J.S. A folded viral noncoding RNA blocks host cell exoribonucleases through a conformationally dynamic RNA structure. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 6404–6409. [CrossRef] [PubMed]

129. Chen, H.; Egger, M.; Xu, X.; Flemmich, L.; Krasheninina, O.; Sun, A.; Micura, R.; Ren, A. Structural distinctions between NAD+riboswitch domains 1 and 2 determine differential folding and ligand binding. *Nucleic Acids Res.* **2020**, *48*, 12394–12406. [CrossRef]
130. Binas, O.; Tants, J.N.; Peter, S.A.; Janowski, R.; Davydova, E.; Braun, J.; Niessing, D.; Schwalbe, H.; Weigand, J.E.; Schlundt, A. Structural basis for the recognition of transiently structured AU-rich elements by Roquin. *Nucleic Acids Res.* **2020**, *48*, 7385–7403. [CrossRef]
131. Saville, B.J.; Collins, R.A. A site-specific self-cleavage reaction performed by a novel RNA in neurospora mitochondria. *Cell* **1990**, *61*, 685–696. [CrossRef]
132. Suslov, N.B.; Dasgupta, S.; Huang, H.; Fuller, J.R.; Lilley, D.M.J.; Rice, P.A.; Piccirilli, J.A. Crystal Structure of the vs. ribozyme. *Nat. Chem. Biol.* **2015**, *11*, 840–846. [CrossRef]
133. Guo, H.C.T.; Collins, R.A. Efficient trans-cleavage of a stem-loop RNA substrate by a ribozyme derived from Neurospora vs. RNA. *EMBO J.* **1995**, *14*, 368–376. [CrossRef]
134. Duss, O.; Maris, C.; von Schroetter, C.; Allain, F.H.T. A fast, efficient and sequence-independent method for flexible multiple segmental isotope labeling of RNA using ribozyme and RNase H cleavage. *Nucleic Acids Res.* **2010**, *38*, e188. [CrossRef]
135. Barrick, J.E.; Corbino, K.A.; Winkler, W.C.; Nahvi, A.; Mandal, M.; Collins, J.; Lee, M.; Roth, A.; Sudarsan, N.; Jona, I.; et al. New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 6421–6426. [CrossRef]
136. Klein, D.J.; Ferré-D'Amaré, A.R. Structural basis of glmS ribozyme activation by glucosamine-6-phosphate. *Science* **2006**, *313*, 1752–1756. [CrossRef]
137. Chapman, E.G.; Costantino, D.A.; Rabe, J.L.; Moon, S.L.; Wilusz, J.; Nix, J.C.; Kieft, J.S. The structural basis of pathogenic subgenomic flavivirus RNA (sfRNA) production. *Science* **2014**, *344*, 307–310. [CrossRef] [PubMed]
138. Kao, C.; Rüdisser, S.; Zheng, M. A simple and efficient method to transcribe RNAs with reduced 3′ heterogeneity. *Methods* **2001**, *23*, 201–205. [CrossRef] [PubMed]
139. Helmling, C.; Keyhani, S.; Sochor, F.; Fürtig, B.; Hengesbach, M.; Schwalbe, H. Rapid NMR screening of RNA secondary structure and binding. *J. Biomol. NMR* **2015**, *63*, 67–76. [CrossRef]
140. Karlsson, H.; Baronti, L.; Petzold, K. A robust and versatile method for production and purification of large-scale RNA samples for structural biology. *RNA* **2020**, *26*, 1023–1037. [CrossRef]
141. Keane, S.C.; Van, V.; Frank, H.M.; Sciandra, C.A.; McCowin, S.; Santos, J.; Heng, X.; Summers, M.F. NMR detection of intermolecular interaction sites in the dimeric 5′-leader of the HIV-1 genome. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 13033–13038. [CrossRef]
142. Kanwal, F.; Chen, T.; Zhang, Y.; Simair, A.; Rujie, C.; Sadaf Zaidi, N.U.S.; Guo, X.; Wei, X.; Siegel, G.; Lu, C. Large-Scale in Vitro Transcription, RNA Purification and Chemical Probing Analysis. *Cell. Physiol. Biochem.* **2018**, *48*, 1915–1927. [CrossRef]
143. Zhang, H.; Keane, S.C. Advances that facilitate the study of large RNA structure and dynamics by nuclear magnetic resonance spectroscopy. *Wiley Interdiscip. Rev. RNA* **2019**, *10*, 1–15. [CrossRef] [PubMed]

Nucleic Acids Research, 2025, **53**, gkaf924
https://doi.org/10.1093/nar/gkaf924
Molecular and Structural Biology

Naphthyridine carbamate dimer ligand induces formation of Z-RNA-like fold of disease-related RNA and exhibits a molecular glue characteristics in crystal lattice formation

Martyna Mateja-Pluta[1], Leszek Błaszczyk[1], Magdalena Bejger[1], Kazuhiko Nakatani[2], Agnieszka Kiliszek[1,*]

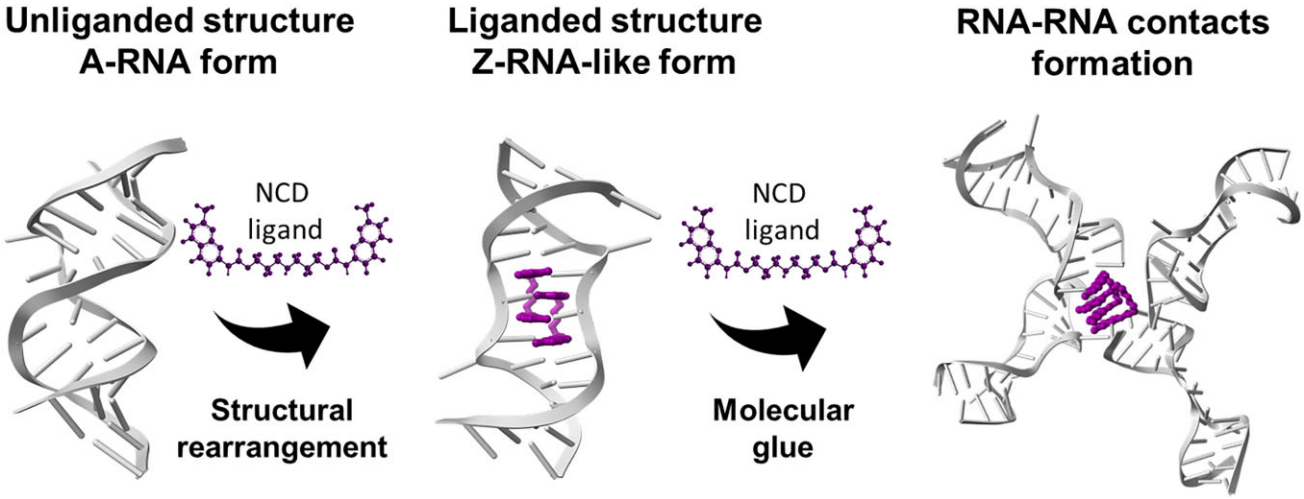
[1]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Z. Noskowskiego 12/14, 61-704, Poland
[2]Department of Regulatory Bioorganic Chemistry, SANKEN (The Institute of Scientific and Industrial Research), Osaka University, 8-1 Mihogaoka, Ibaraki 567-0047, Japan

*To whom correspondence should be addressed. Email: kiliszek@ibch.poznan.pl

Abstract

The naphthyridine carbamate dimer (NCD) is a small molecule that recognizes disease-related RNA containing UGGAA repeats associated with spinocerebellar ataxia type 31 (SCA 31) and alleviates the disease phenotype *in vitro* and *in vivo*. In this study, we use X-ray crystallography to elucidate the mode of NCD binding in detail. We determine the crystal structures of the RNA–NCD complex and a structure of unliganded RNA. The NCD interacts differently than in previously reported nuclear magnetic resonance structure, forming pseudo-canonical base pairs with guanosine residues located on the same RNA strand. Furthermore, in one of the complexes, the ligand is located between symmetry-related RNA molecules, exhibiting a molecular glue characteristics in crystal lattice formation. The comparison of RNA–NCD and ligand-free models allows the identification of structural changes in RNA upon ligand binding from A-form to Z-RNA-like form. These observations extend our understanding of the interactions between RNA and small compounds and can be useful as a reference model in the development of bioinformatics tools for RNA–ligand structure predictions.

Graphical abstract



Introduction

Biologically active small synthetic molecules are organic compounds with molecular sizes typically <1000 Da [1]. Since the synthesis of urea in 1828, the development of synthetic molecules has evolved into a sophisticated research field supported by structural methods, such as X-ray crystallography, nuclear magnetic resonance (NMR), and bioinformatics. Currently, it is experiencing a breakthrough, mainly due to innovations in new moieties, implementation of high-throughput techniques, and advances in bioinformatics [2–6].

New modalities have provided an opportunity to focus on the underexplored field of RNA-targeting. RNA is a vital molecule that is abundant in cells and is involved in almost all steps of gene expression. Alterations in the RNA structure and function can lead to the development of serious disorders. In this regard, RNA is an attractive target for drug development

Received: April 17, 2025. **Revised:** August 18, 2025. **Accepted:** August 27, 2025

and an exciting area of many ongoing studies. However, there are significant challenges and risks, including target selection, hit validation, optimization of small molecules, and translation into the cellular environment. One of the difficulties arises from the specific structural characteristics of the RNA, which is a negatively charged molecule that prefers to bind ligands with positively charged functional groups or metal ions, or water molecules from the environment. Another aspect concerns the architecture of the binding pockets in RNA, which are different from that observed in proteins. Typically, they are shallow, polar and solvated [7]. Moreover, RNA exhibits conformational dynamics that can lead to the coexistence of different structures of the same molecule, or structural rearrangements upon ligand binding [8]. Finally, our understanding of the structural nature of RNA–ligand interactions is very limited. There are >37 000 biomolecule structures in complex with ligands deposited in the Protein Data Bank, but only 565 represent RNA–ligand complexes. Although experimentally determined structures provide invaluable knowledge that supports the selection of compounds from libraries for high-throughput screening, optimization of small molecules and the improvement of bioinformatics structural predictions could also accelerate this process.

In recent years, we developed a class of small molecules called mismatch binding ligands (MBLs) that recognize non-canonical base pairs in RNA or mismatches in DNA [9–13]. They were designed to contain aromatic moieties complementary to the hydrogen-bonding pattern of the nucleobases [11,14]. Structural analysis using NMR and X-ray crystallography confirmed that MBLs interact directly with nucleotides forming Watson–Crick like pairs [8, 11, 12, 15, 16]. In some cases, ligand binding was associated with conformational changes in the nucleic acid structure [11, 12]. Recently, one of the first-generation molecules, naphthyridine carbamate dimer (NCD), was shown to recognize the pathogenic RNA associated with spinocerebellar ataxia type 31 (SCA 31) [17]. SCA 31 is an autosomal spinocerebellar neurodegenerative disorder. It is caused by abnormally repeated 5′-UGGAA-3′/5′-UGGAA-3′ RNA pentads, which sequester many cellular proteins by forming toxic nuclear foci [18]. *In vitro and in vivo* studies have shown that NCD ligand possesses pharmaceutical properties. Feeding NCD to larvae of the Drosophila SCA31 model alleviated the disease phenotype induced by toxic UGGAA repeats. The previous studies have demonstrated that NCD acts as a molecular glue facilitating the assembly and stabilization of higher-order RNA and DNA architectures. NCD and its multimeric derivatives were used to induce folding of DNA into tetrahedron structure, for modulation of RNA aptamer activity and to enhance the catalytic properties of the hammerhead ribozyme [19–22]. An NMR study showed that two NCD molecules bind to 5′-UGGAA-3′/5′-UGGAA-3′ RNA. Each NCD ligand interacted with two guanosine residues located on opposite RNA strands and formed noncanonical guanosine–naphthyridine pairs [17]. To further elucidate the mode of NCD binding, we performed an X-ray crystallographic analysis of RNA oligomers containing the 5′-UGGAA-3′/5′-UGGAA-3′ motif.

Here, we report two crystal structures of the RNA–NCD complex and one, so far unreported, structure of unliganded RNA. In the models of complexes, we observed two NCD molecules bound to the internal loop of the UGGAA motif. The mode of binding is different from that of previously reported NMR structure. Furthermore, in one of the complexes, an additional binding site for NCD was found. The ligand was located between the symmetry-related RNA molecules and helped forming additional interactions in the crystal lattice. This shows that NCD exhibits a molecular glue characteristics and facilitate the formation of crystal contacts. In the unliganded model, the 5′-UGGAA-3′/5′-UGGAA-3′ motif formed an internal loop that maintained the helical character of RNA. Three out of the four guanosine residues of the motif stacked one above another, attracting the cobaltamine $[Co(NH_3)_6]^{3+}$ cation. The comparison between bound and unbound models revealed that the binding of the NCD molecules caused major structural changes in the RNA which were confirmed by biophysical methods. These observations extend our understanding of the interactions between RNA and small compounds in the context of potential therapeutics against RNA-driven disorders, and can be useful as a reference model in the development of bioinformatics tools for RNA–ligand structure predictions.

Materials and methods

Synthesis, purification, and crystallization of the oligomers

All oligomers were synthesized by the solid-phase method in an Applied Biosystem DNA/RNA synthesizer using TOM-protected phosphoramidites [23]. The synthesis was conducted using the DMT-ON procedure, leading to oligomers with a trityl group attached at the 5′end. The RNA oligonucleotides were purified using Glen-Pak cartridges according to the appropriate protocol, desalted, and lyophilized under vacuum conditions using Speed-Vac [24–26]. Prior to crystallization, RNA was dissolved in 10 mM sodium cacodylate (pH 7.0) to a final concentration of 1 mM. Oligonucleotides were denatured at 95°C for 5 min and then gradually cooled to room temperature. Crystallization of RNA complexes with NCDs was conducted using the sitting drop vapor-diffusion technique at 19°C. RNA and NCD ligands were mixed in a 1:2 molar ratio. The optimal crystallization conditions for ROG2 oligomer were 0.012 M sodium chloride, 0.08 M potassium chloride, 0.04 M sodium cacodylate trihydrate (pH 5.5), 45% v/v (±)-2-methyl-2,4-pentanediol, 0.002 M hexammine cobalt(III) chloride, for ROG2–NCD 0.08 M potassium chloride, 0.02 M magnesium chloride hexahydrate, 0.04 M sodium cacodylate trihydrate (pH 6.0), 45% v/v (±)-2-methyl-2,4-pentanediol, 0.012 M spermine tetrahydrochloride and for ROG1–NCD 0.01 M magnesium acetate tetrahydrate, 0.05 M sodium cacodylate trihydrate (pH 6.5), 1.3 M lithium sulfate monohydrate. Crystals of ROG1–NCD appeared within 3 months, ROG2–NCD within 1 month while unliganded ROG2 within 2 weeks. The crystals of ROG1–NCD were cryoprotected with 20% glycerol and flash frozen, whereas ROG2–NCD and ROG2 were directly transferred from the crystallization drop to liquid nitrogen. The crystals were transferred to a synchrotron core facility by using a dry shipper.

X-ray data collection, structure solution, and refinement

The X-ray diffraction data were obtained using the beam lines at BESSY II (Berlin) [27–29], DESY (Hamburg) [30, 31], and ESRF (Grenoble) [32, 33]. The data were integrated and scaled using the XDS program suite, facilitated by the xdsgui inter-

face [34]. The complex structure of ROG1–NCD complex was solved by the molecular replacement (MR) method using the PHASER program [35, 36]. The search model was the crystal structure of the CUG repeat duplex (PDB code: 3GLP) [37]. The structures of ROG2–NCD complex and ROG2 oligomer were solved by MR with PHASER, using ROG1–NCD (PDB code: 9I9W) as a search model. The refinement were carried out using Refmac5 from the CCP4 program suite [38–42]. Restraints for the NCD ligand were generated using the Grade Web server, a tool developed by Global Phasing (https://grade.globalphasing.org). Manual model building and electron density map inspection were conducted using Coot [43, 44]. The helical parameters were calculated using 3DNA [45, 46]. The X-ray data and refinement statistics are summarized in Supplementary Table S1.

Circular dichroism measurement

Circular dichroism (CD) experiments were carried out on a J-725 CD spectropolarimeter (JASCO) using a 10 mm path length cell at room temperature. Prior to measurement, RNA (10 μM) was refolded in sodium cacodylate buffer (40 mM, pH 7.0) containing 100 mM NaCl for 5 min at 95°C and snap-cooled on ice for 10 min, followed by incubation for 10 min at room temperature. The CD titration spectra of ROG2 with variable concentrations of NCD (10, 20, 30, 40, and 50 μM) were measured at ambient temperature. For each sample, five spectral scans were accumulated in the 215–370 nm range.

Differential scanning calorimetry measurements

The RNA oligomers were dissolved in 100 mM sodium chloride and 40 mM sodium cacodylate buffer adjusted to pH 7.0. The final concentration of RNA was 100 μM with presence and absence of NCD ligand (200 μM) and cobaltamine chloride (200 μM) To equilibrate the RNA with the reference solutions, the samples were dialyzed overnight at 4°C. Differential scanning calorimetry (DSC) experiments were conducted using a MicroCal PEAQ-DSC calorimeter (Malvern Instruments, Ltd.). Each measurement was carried out in five cycles of heating and cooling in the range of 2–110°C at a scan rate of 1°C/min. Initially, reference scans of the buffer were performed to establish the instrument's thermal history and achieve near-perfect baseline repeatability. The results were analyzed using dedicated software implemented by Malvern Instruments. The melting temperature (T_m) was calculated by applying two-state model fitting.

Ultraviolet melting temperature measurements

Ultraviolet (UV) thermal melting studies were performed on a Jasco V-700 spectrometer with a thermoprogrammer. The RNA oligomers were dissolved in a buffer containing 100 mM sodium chloride and 40 mM sodium cacodylate (pH 7.0). Unliganded duplexes of ROG1 and ROG2 oligomer as well as RNA–Co(NH_3) and RNA–NCD complexes were prepared in three different concentrations (10, 30, and 100 μM). The ratio between RNA and ligand molecules (NCD and cobaltammine) was 1:2. The UV absorption versus temperature was measured at 260 nm at a heating rate of 0.5°C/min in the range 4–90°C [47]. The melting curves were analyzed using Microsoft Excel software to calculate the first derivative of the UV-melting curves and melting temperature.

Native and denaturing gel electrophoresis

For native gel electrophoresis 10 pmol of RNA in 10 mM sodium cacodylate (pH 7.0) and 100 mM sodium chloride was renatured for 2 min at 95°C, snap cooled in 4°C for 5 min and transferred into room temperature. Next, RNA was mixed with increasing concentrations of NCD ligand (0.62–80 μM) and incubated for 12 h at room temperature. The final RNA concentration in the reaction was 0.3 μM. Samples were combined with ficoll (2%) and loaded on 20% polyacrylamide gel with 0.5× TB buffer (Tris-Borate). After electrophoresis gel was stained with SYBR Gold and visualized using Amersham Typhoon Biomolecular Imager (Cytiva). For evaluation of RNA homogeneity samples were renatured in the same manner except that 10, 30, 100 μM RNA concentration was used. RNA was visualized using toluidine blue staining.

In case of denaturing gel electrophoresis 1 nmol of RNA was combined with equal volume of 8 M urea and loaded on 15% polyacrylamide gel in 1× TBE buffer (Tris-Borate-EDTA). After electrophoresis RNA was visualized using UV light.

Results

In this study, we used a NCD ligand containing two naphthyridine units (NU) connected by a relatively long aliphatic linker (3-(carbamoyloxy)propylaminopropyl carbamate) consisting of eleven atoms (Fig. 1A). The distribution of nitrogen atoms in NU was designed to form base pairs with the guanosine residues [15, 48, 49]. As a result the NCD was found to selectively bind two guanosines located in single-stranded or bulged regions of nucleic acids [50, 51]. Based on a previous study, we designed and synthesized two RNA oligomers (ROG1 and ROG2). Each of the oligomers formed self-complementary duplex: ROG1 $(GGCACUGGAAGUGC)_2$ and ROG2 $(GGCACUGGAAGUGCC)_2$; containing the 5′-UGGAA-3′/5′-UGGAA-3′ pentad motif selectively recognized by NCD (Fig. 1B) [17, 52]. ROG1 was shorter than ROG2 by one cytosine residue at the 3′ end. We crystallized two RNA–ligand complexes: ROG1–NCD (1.50 Å resolution) and ROG2–NCD (2.55 Å resolution) and one unliganded duplex of ROG2 oligomer (2.79 Å). All X-ray data and refinement statistics are presented in Supplementary Table S1. The electron density maps of all the crystal models were unambiguous resulting in well-defined positions and conformations of the RNA and NCD ligands (Supplementary Fig. S1).

General overview of RNA–ligand structures

In both liganded structures, the asymmetric unit contained one RNA duplex consisting of chains A and B. The superposition of the ROG1–NCD and ROG2–NCD models showed similarity, with an r.m.s.d (root mean square deviation) value of 0.8 Å. The highest conformational similarity was observed in the central part of the structures embedding the 5′-UGGAA-3′/5′-UGGAA-3′ motif (r.m.s.d was 0.67 Å of main chain). The pentad motif bound two NCD molecules that interacted with all four guanosine residues (Fig. 1C). As a result, six base pairs were formed: two 6U–10A pairs and four pseudo-canonical naphthyridine-guanine pairs. The 9A residues were flipped out and engaged in crystal lattice contacts with symmetry-related RNA molecules (Fig. 1C and Supplementary Fig. S2). The pen-

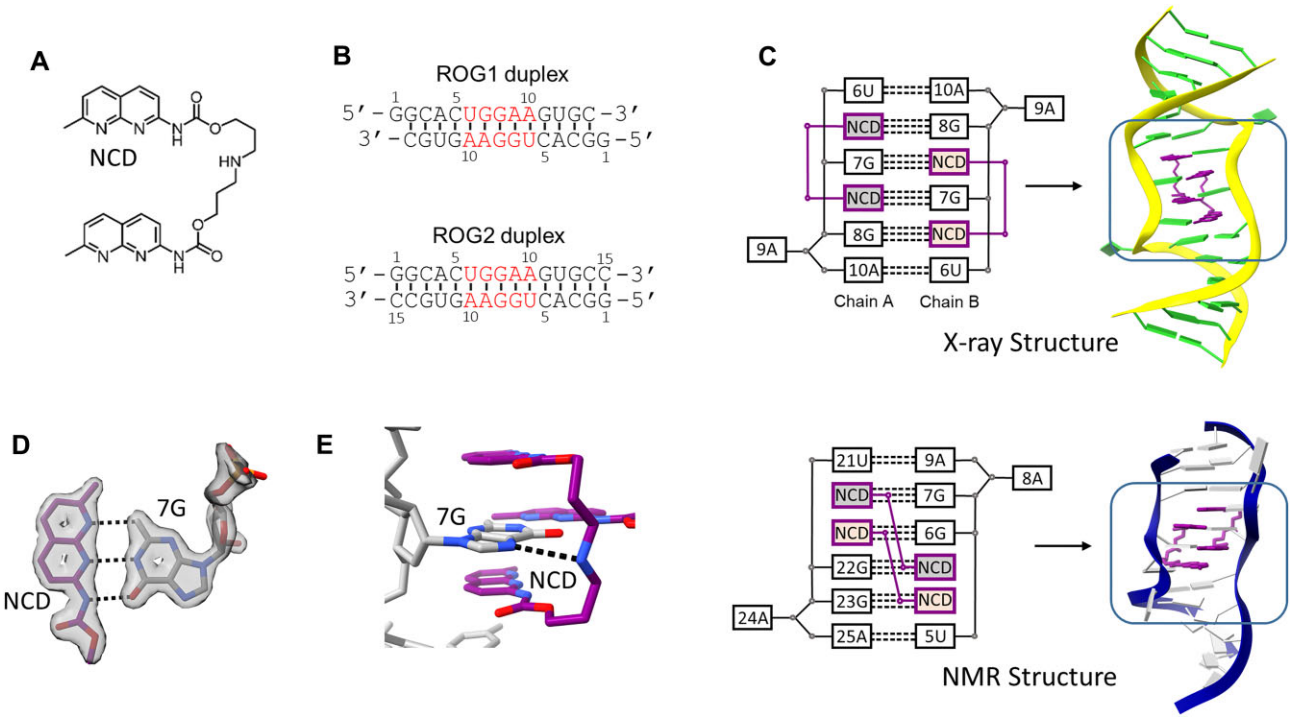


Figure 1. Interactions between NCD ligand and RNA. (**A**) Chemical structure of NCD molecule. (**B**) Duplexes of ROG1 and ROG2 oligomers (the UGGAA motif is presented in red). (**C**) Secondary structure diagrams and tertiary structures of interactions between the UGGAA motif and NCD molecules (purple) in X-ray crystallographic (upper part) and NMR (lower part) models. (**D**) Pseudo-canonical base pair formed between the NU of NCD and the guanosine residue. The $2F_o–F_c$ electron density map (gray) is contoured at the 1 σ level. H-bonds are represented by black dashed lines. (**E**) H-bond between the amino group of the NCD linker and the imino group of the neighboring G residue.

tad motif showed structural symmetry, as reflected by identical interactions between the ligand and guanosine residues.

In the ROG1–NCD complex, a third ligand molecule was observed between the 5′ ends of the symmetry-related helices, participating in crystal lattice formation (*vide section: Role of NCD ligand in crystal lattice formation*).

Interactions between NCD and UGGAA motif

The NCD molecules were located in the minor groove of the 5′-UGGAA-3′/5′-UGGAA-3′ motif. Each ligand bound two guanosine residues located on the same RNA strand, forming identical pseudo-canonical NCD-G pairs (Fig. 1C). Each NU unit interacted with the guanine base *via* three hydrogen bonds (Fig. 1D). The first hydrogen bond was formed between the N1 imino group of naphthyridine and the N2 amine group of guanine. A second bond was formed between the N8 imino group of NU and the secondary nitrogen atom N1 of the guanine amino group. A third interaction occurred between the amide group of the NU moiety and O6 carbonyl of the G base. Additional RNA–ligand interaction was observed between the amine group of the NCD linker and the N7 imine group of the G7 residue (Fig. 1E). All four pseudo-canonical base pairs formed a stack consisting of alternating pairs: NCD-G8, G7-NCD, NCD-G7, G8-NCD. This arrangement resulted in the location of G7 residues from chains A and B between two NU units from the same ligand molecule (Fig. 1C). Likewise, one of the NU units was located between the U6 and G7 residues, and the second NU unit was located between the G7 and G8 residues. In other words, the U6, G7, and G8 residues from each RNA strand were separated by wedged NU units of NCD ligands, increasing the distance between neighboring residues from the typical value of approximately 3.2 Å to approximately 6.5 Å. Despite the formation of pseudo-canonical pairs, only limited stacking interactions between the naphthyridines and nitrogen bases were observed.

Role of NCD ligand in crystal lattice formation

In the ROG1–NCD model, the third ligand molecule connected the symmetry-related RNA helices. The contacts in the crystal lattice were formed by two NCD molecules (one from the asymmetric unit and second, NCD', from the symmetry-related complex) and four RNA duplexes (one from the asymmetric unit and three from symmetry-related RNAs) (Fig. 2). The NCD was bound to the 1G residue of chain B and 1G of chain A'. Similarly, the NCD' linked the 1G residue of chain B" and 1G of chain A"'. In total, four pseudo-canonical G-NCD pairs were observed, and similar to the 5′-UGGAA-3′/5′-UGGAA-3′ motif, they were arranged in an alternating manner (the NCD-G pairs were separated by symmetry-related NCD'-G pairs) (Fig. 2). The observed lattice interactions resulted in the alignment of two helices, one above the other, while the other two helices were perpendicular to the RNA helix from the asymmetric unit. In ROG2 and ROG2–NCD models, lattice contacts were formed by stacking interactions between symmetry-related duplexes, resulting in the formation of pseudo-continuous helices.

NCD was shown earlier to act as a molecular glue for DNA and RNA molecules [19–22]. To further investigate the properties of NCD in facilitating the interaction between ROG1 duplexes in solution, we performed native gel electrophoresis of ROG1 RNA in the presence of increasing concentrations of NCD ligand (Supplementary Fig. S3A). A slower migrating

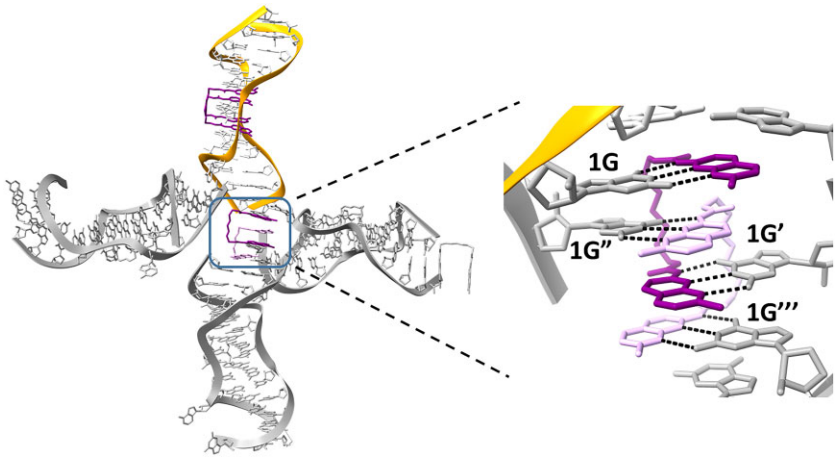


Figure 2. Crystal lattice formation by NCD ligand. Two ligand molecules, NCD (dark purple) and NCD′ (light purple), bind four guanosine residues of four symmetry-related RNA molecules.

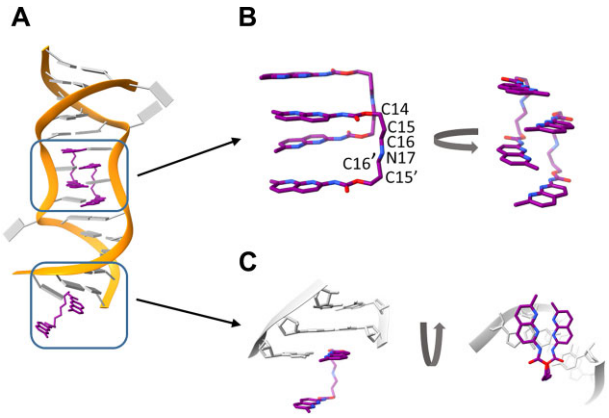


Figure 3. Conformation of ligand molecules in the ROG1–NCD structure. (**A**) Binding sites of the NCD in the ROG1 helix. (**B**) Conformation of the NCD molecules that bind to the UGGAA motif. (**C**) Structural arrangement of the NCD participating in the crystal lattice contacts formation.

bands were observed within a range of NCD:ROG1 ratio of 1.8:1–30:1, indicating the formation of higher-order NCD–ROG1 complexes. At the two highest NCD concentrations, the intensity of all bands was greatly reduced suggesting the NCD-dependent RNA aggregation.

Ligand conformation

All ligand molecules exhibited a square bracket-like conformation. The aliphatic part of the linker was elongated, whereas the NUs with carbamate (R-NH-CO-O-R') groups were perpendicular to the linker (Fig. 3A). In the case of NCD ligands bound to the pentad motif, the aliphatic part of the spacer consisted of six atoms (C14, C15, C16, N17, C16', and C15'). They were arranged in a zigzag line with a length of 6.3 Å. The NU units of each NCD ligand were parallel and stacked one above another (Fig. 3B). Moreover, the nitrogen atoms of NU were oriented in the same direction (*cis* geometry).

In the ROG1–NCD model, the aliphatic linker of the ligand connecting symmetry-related molecules showed a similar zigzag conformation but consisted of seven atoms (C14, C15, C16, N17, C16', C15', and C14'). This resulted in a longer distance (7.4 Å) between the two NU planes. Additionally, the NUs were twisted relative to each other by 17° and their nitrogen edges showed a trans geometry (Fig. 3C).

Unliganded RNA duplex structure

The unliganded ROG2 oligomer crystallized as a self-complementary duplex. Similar to the ROG2–NCD structure, the 5′-UGGAA-3′/5′-UGGAA-3′ motif was located in the central part of the helix and flanked by five Watson–Crick base pairs.

The distal parts of the helix exhibited an A'-RNA form, with nucleotides in the C3'-endo or C2'-*exo* conformation (Fig. 4A). The values of helical parameters were within A'-RNA range: Zp > 1.5°, twist = 31.8°, and rise = 3.2Å [53–55]. The GGA residues of the pentad motif formed an internal 3 × 3 loop (GGA/AGG) showing double-stranded character in terms of strand separation, but deviated from the A-RNA helix, which was reflected in most of the helical parameters (e.g. stretch, twist, or rise) (Supplementary Table S2). Moreover, the presence of an internal loop was associated with a bending of the helix by 100° (Fig. 4A). In contrast to the liganded structures, the arrangement of the pentad motif residues in the unliganded model was not symmetric. It consisted of two flanking 6U–10A Watson–Crick base pairs and two cross-strand stacks of nucleotides. The first stack was located in the major groove of the helix and consisted of three guanosine residues: 7G and 8G from chain B, and 7G from chain A (Fig. 4B). The second stack, located in the minor groove, consisted of 9A from chain B, 9A' from symmetry-related chain A', and 8G from chain A (Fig. 4B and Supplementary Fig. S4). Both stacks could be extended by neighboring U-A pairs. The 6U from chain A overlapped with the first stack, whereas 10A from chain B overlapped with the second stack. The conformation of the nucleotides in the first stack was typical for the A-RNA form. In the second stack, the 9A exhibited a *syn* conformation. The 9A from chain A' had C2'-endo sugar puckering and an *anti* conformation of the nucleobase. The remaining 8G was shifted toward the minor groove and showed a high *anti-conformation*, reflected in the χ torsion angle having −104.2° value instead of approximately 180° [55].

The cross-strand stacks of nucleotides were almost parallel to each other and formed only one H-bond between nucleobases (*exo*-amino group of 7G in chain B and the N7

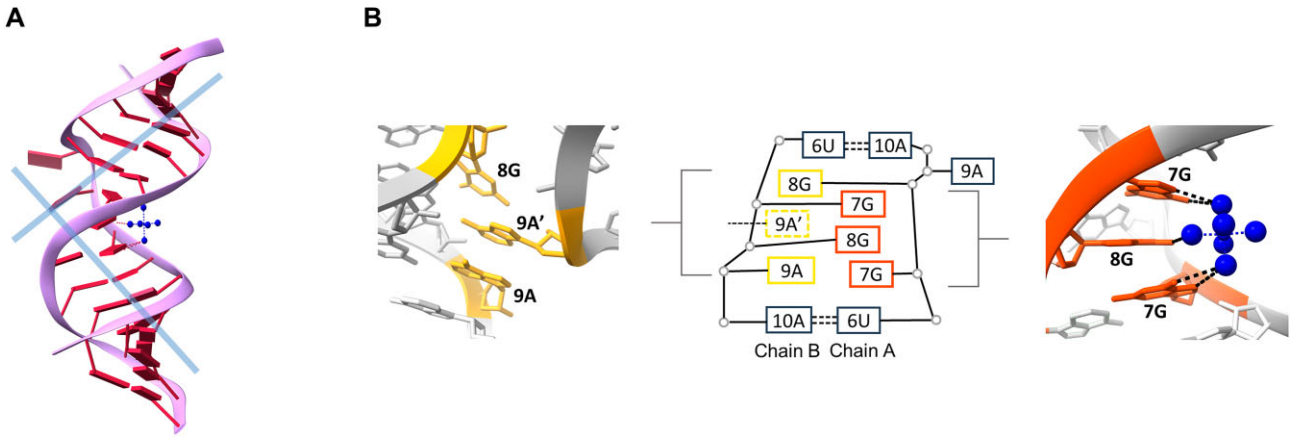


Figure 4. Unliganded structure of the ROG2 duplex. (**A**) The bending of the helix is marked by blue lines representing helical axes. The hexammine cobalt cation (blue) is located in the major groove. (**B**) 3 × 3 internal loop formed by the UGGAA motif. One of the nucleotides stack (yellow) is located in the minor groove while the second (orange) in the major groove. Hexammine cobalt cation (blue) interacts with guanosine residues of 3 × 3 internal loop.

imino group of 8G in chain A). Other interactions were observed between the guanine nucleobases and phosphate backbone (Supplementary Fig. S5). Additionally, the guanosine residues of the first stack interacted with the cobalt hexammine $[Co(NH_3)_6]^{3+}$ cation located in the major groove of the helix (Fig. 4).

Structural changes of RNA associated with ligand binding

The comparison between the unliganded ROG2 and the ROG1/ROG2–NCD models showed that the flanking base pairs of the pentad motif were similar in terms of helical parameters (Supplementary Table S3). The superposition of the flanking base pairs of ROG2 and ROG2–NCD showed r.m.s.d. of 0.3 Å for residues 11–15 of chain A and residues 1–5 of chain B and r.m.s.d. of 0.9 Å for residues 1–5 of chain A and 11–15 of chain B. Major differences between the unliganded and liganded models were observed for the 5′-UGGAA-3′/5′-UGGAA-3′ motif. Only the flanking U-A pairs remained the same, whereas other residues underwent substantial rearrangement. In the liganded models, the pentad loop was unwound, resulting in loss of helicity (Fig. 1C). Both 9A residues were flipped out, while guanosines interacting with the NCD ligands formed a double-stranded stem resembling the Z-RNA like structure [56] (Supplementary Fig. S6). The sugar-phosphate backbone followed a zigzag course, which was reflected in α (O3′—P—O5′—C5′) and ζ (C3′—O3′—P—O5′) torsion angles of 7G and 8G residues of both RNA strands (Supplementary Table S4). This unusual conformation resulted in a large distance between the 6U, 7G, and 8G residues (6.4–7.0 Å). The resulting gaps were filled with the naphthyridine rings of the NCD ligands. As in Z-RNA form, guanosines of the pentad motif had *syn* conformation of the N-glycosidic bond and alternating sugar puckering from C3'-endo for the 7G residue to C2'-endo for the 8G residue. However, although the pentad motif resembled many structural features of the Z-RNA form, the structure remained right-handed, as in A-RNA (Supplementary Fig. S6).

Physicochemical data

DSC was used to determine the thermal properties of the RNA samples. Since the ROG2 oligomer contained one binding site for NCD molecules, it was used for DSC evaluation. Three samples were measured: unliganded ROG2, the ROG2–NCD complex, and ROG2 in the presence of hexamminecobalt [ROG2-$Co(NH_3)_6$], which was present in the crystal structure. In general, the DSC spectra of all the samples in all cycles showed two peaks (Tm1 and Tm2) (Fig. 5A). The Tm1 peak was observed at 43°C for unliganded ROG2, 43°C for ROG2–NCD and 46°C for ROG2-$Co(NH_3)_6$ complex after the third round of denaturation–renaturation cycle. The Tm2 peak was observed at nearly the same temperature (79.0°C), regardless the absence or presence of the ligand molecule (Supplementary Table S5A). In the case of the ROG2–NCD complex, the first and second round of denaturation–renaturation cycle resulted in one major peak which was higher and slightly asymmetric than those of the other samples. Almost no decay of signal was observed for unliganded ROG2 across all five rounds of denaturation–renaturation cycle, whereas in the presence of hexamminecobalt and NCD gradual diminishing of the peak was observed (Fig. 5A). To determine whether the decay of the signal was associated with RNA degradation, we performed denaturing gel electrophoresis of the samples after DSC measurements. The DSC ROG2–NCD band was diffused in comparison to the control line, whereas for DSC ROG2-$Co(NH_3)_6$, it was well-defined (Supplementary Fig. S3B). The diffused band observed for the ROG2–NCD sample suggests temperature-dependent NCD decomposition leading to disruption of RNA–ligand complex. This is reflected in the DSC spectra because after the third cycle, the DSC profile resembles that of the unliganded ROG2 duplex.

To compare the thermal stability of the ROG1 and ROG2 duplexes, we used the UV-melting method (Fig. 5B and C and Supplementary Table S5B). The measurements were performed for ROG1, ROG2, and their complexes with $[Co(NH_3)]^{3+}$ and NCD ligands at three different RNA concentration (10, 30, and 100 μM) to assess its influence on RNA melting behavior. At 100 μM RNA, the UV melting curves of unliganded ROG1 and ROG2 exhibited two distinct peaks. The first peak occurred at approximately 33°C for ROG1 and 48°C for ROG2. At lower RNA concentrations (10 and 30 μM), the first transition was difficult to resolve due to its low energetic contribution, resulting in a broad and shallow peak. In contrast, the second peak was

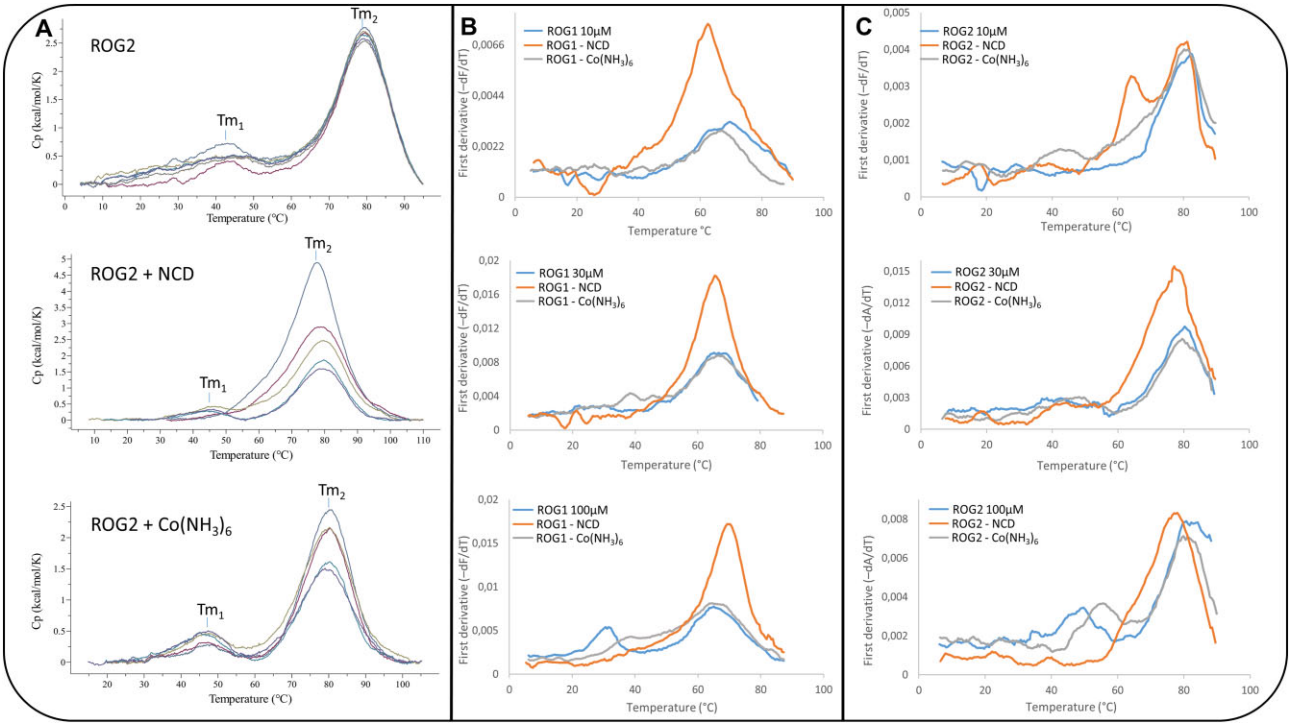


Figure 5. Physicochemical evaluation of ROG1 and ROG2 duplexes and their interactions with NCD and $Co(NH_3)_6$ ligands using (**A**) DSC and (**B**, **C**) thermal denaturation monitored by UV spectroscopy. The first derivative of absorption (−dA/dT) was plotted against temperature.

well defined at all RNA concentrations, occurring at approximately 67°C for ROG1 and 80°C for ROG2. In the presence of $[Co(NH_3)]^{3+}$ (at 100 μM RNA), a shift in the first peak was observed (Δtm = 6°C for ROG1 and Δtm = 7°C for ROG2). In general, UV melting of the RNA–NCD complexes revealed a single, asymmetric peak (at 65°C for ROG1–NCD and 77°C for ROG2–NCD). The only exception was observed for 10 μM ROG2–NCD, where two transitions appeared: the first at ∼62°C and the second at ∼80°C (Fig. 5C and Supplementary Table S5B).

To exclude the influence of the RNA concentration on the RNA heterogeneity we performed native gel electrophoresis. The results clearly shows high structural homogeneity of both duplexes regardless the RNA concentration (Supplementary Fig. S3C).

Structural rearrangements in ROG2 upon ligand titration were investigated using CD (Supplementary Fig. S7). The most significant changes in the CD spectra were observed in the 245–225 nm range. The binding of NCD induced an inversion of the peak at 233 nm from a negative to positive value. Other spectral changes were observed at 263 nm for the decreasing peak and in the 355–325 nm range.

Discussion

NCD binding induces Z-RNA-like conformation

The crystal model of the unliganded ROG2 oligomer allowed the tracking of the structural rearrangements of RNA upon NCD binding, which was not possible in an earlier study [17]. The conformational changes were significant and mostly involved the 5′-UGGAA-3′/5′-UGGAA-3′ motif. In the native structure, it formed an internal loop consisting of two flanking U-A pairs and two stacks of nucleotide residues. One consisting of three guanosines and the second consisting of two adenosines and one guanosine (Fig. 4B). Although the stacked nucleotides were not engaged in base pairing, they were well defined and ordered. The helical character of the duplex was maintained; however, the major groove was narrowed and the helix was bent (Fig. 4A). Binding of NCD molecules resulted in unwinding of the pentad motif and stretching of the helix, similar to the Z-RNA form (Supplementary Fig. S6). All guanosines of the pentad motif switched from *anti* to *syn* conformation and formed H-bonds with the NUs of the ligand, whereas 9A residues from both strands were flipped out. Consequently, the pentad motif gained internal symmetry, which was not observed in the unliganded structure.

Structural changes of RNA were also observed in the CD spectra. The inversion of the peak at 235 nm from negative to positive upon ligand saturation could indicate unfolding of the pentad motif, leading to the loss of helicity (Supplementary Fig. S7). The slight changes in the shape of the spectra in the 325–355 nm range were most likely related to the chirality of the NCD molecule. Upon binding to RNA the achiral NCD became the chiral component of the complex. A similar effect was observed in our previous study [8].

The biological impact of NCD binding to RNA containing $r(UGGAA)_n$ was observed in HeLa cells and a Drosophila model of SCA31 [17, 49]. HeLa cells expressing toxic $r(UGGAA)_{76}$ RNA and exposed to the NCD ligand exhibited a significant reduction in RNA foci formation while in the Drosophila, NCD showed suppression of disease phenotype. NCD's binding to mutated transcripts can either block RNA-binding protein (RBP) sites or alter the RNA structure so it is no longer recognized by RBPs. It is conceivable that $r(UGGAA)_{76}$ forms a hairpin structure with a long stem composed of U-A and A-U pairs interrupted by GGA internal loops. Based on our results NCD binding stabilizes stem struc-

ture by forming a paired region with the guanosine residues of the GGA motif, inducing a conformational change in the RNA from an A to a Z-like form. Consequently, the internal loop becomes unavailable for interactions, and the global RNA structure differs from its unbound state. This can explain why in the presence of NCD ligand r(UGGAA)$_n$ repeats are no longer a toxic factor in SCA31 pathogenesis in model organism [17]. The family of mismatch binding molecules (MBL) compounds, including NCD, can induce structural rearrangement of RNA [9–11, 17, 50]. The crystallographic and NMR structures of complexes of nucleic acids with other ligands such as naphthyridine-azaquinolone molecules and cyclic mismatch-binding ligands showed that ligands formed pseudo-canonical base pairs with target nucleotides [11, 14, 16, 57]. As a consequence, the neighboring residues were flipped out in order to fit into the structural changes of the DNA/RNA helix imposed by ligand binding. We observed that NCD binding caused an unexpected loss of helicity, which would be difficult to predict using *in silico* approaches. Hence, the presented structures can serve as excellent reference models for the evaluation and improvement of existing and development of future approaches for computational prediction of RNA–ligand interactions.

Comparison of crystal structures with the NMR model

In a previous study, the NMR structure of the RNA–NCD complex has been described [17]. Similar to the crystallographic models, two NCD molecules were bound to guanosine residues of the 5′-UGGAA-3′/5′-UGGAA-3′ motif. Both adenosine residues were flipped out and two flanking A-U pairs were formed (Fig. 1C). However, the arrangement of the NCD molecules in the NMR model was different from that in the X-ray structures. Each ligand molecule was base-paired with guanosine residues located on opposite RNA strands. As a result, the G residues involved in ligand binding were not separated by the NUs. Consequently, two naphthyridines from each ligand were adjacent and located between the uridine and two consecutive guanosines (Fig. 1C). The conformations of the NCD linkers also differed, representing a disordered C-shape rather than an ordered zigzag-like conformation and the amino group of the linker did not form H-bond with N7 imino group of guanosine residue as in X-ray model (Fig. 1E).

It is difficult to explain the distinct folding of the RNA–NCD complex in the NMR and crystallographic structures. Crystallographic models were determined based on electron density maps, providing unambiguous information on the atomic positions of the NCD and RNA molecules and eliminating other binding possibilities. The NMR model with neighboring G6–G7 and G22–G23 guanosine residues was based on the observation of weak NOE signals of exchange of the imino protons of guanosine residues, suggesting their close proximity (Fig. 1C) [17]. Using NMR restraints, only the arrangement of neighboring guanosine residues resulted in reasonable structures. The existence of two modes of NCD binding could arise from the kinetic or thermodynamic effects present in the experimental conditions. In solution, a starting point for both crystallographic and NMR experiments, the internal loop of the UGGAA motif probably shows structural dynamics. Thus, the initiation of NCD binding can influence the subsequent events during the full complex formation. In one possible scenario, the interactions between the ligand and RNA could start with the formation of H-bonds between one NU and one of the guanosine residues, most likely 8G from chain B located in the minor groove (Fig. 4B). The formation of interactions between the second NU and second G residue on the same or opposite strand would lead to the formation of two different conformations of the RNA–ligand complexes. When second NU binds the G residue on the opposite strand, an NMR-like structure will be observed. When the second guanosine is located on the same strand, the complex will form crystallographic-like structure. The formation of an NMR-like structure may be a kinetically driven process, whereas the crystallographic structure is thermodynamically dependent. Over time, the kinetic product of the RNA–ligand complex can be converted to a thermodynamically stable structure due to partial dissociation of ligand and formation of crystallization-like interactions. The RNA–ligand complexes were prepared under similar conditions for NMR and crystallography analysis [20 mM sodium phosphate (pH 6.9) and 100 mM NaCl for NMR and 10 mM sodium cacodylate (pH 7.0) and 100 mM NaCl for crystallography]. However, for crystallization, the RNA–ligand complexes were further mixed with a buffer containing precipitants such as 45% MPD and 1.3 M lithium sulphate. This may change the balance between the kinetic and thermodynamic products compared to the NMR experiment. Alternatively, both types of complexes are formed in solution but usually only one of the conformational state forms crystals while in NMR spectra overlapping signal of two conformers could be difficult to distinguish [58].

The architecture of the internal loop

The unliganded structure of ROG2 contained a purine (GGA/AGG) 3 × 3 internal loop, in which nucleotides were unpaired and formed two cross-strand stacks of residues (Fig. 4B). The search using RNA FRABASE revealed that the GGA/AGG loop was not observed in RNA structures deposited in the PDB database [59]. Instead we found three other types of 3 × 3 purine-rich loops: GAA/AAG, GAA/AGG, and GAG/AGG. Unlike in the loop of ROG2 RNA, their residues were paired and formed noncanonical G-A, A-A, and G-G pairs. However, using the RNA 3D motif atlas, we identified a 3 × 3 unpaired loop with structural characteristics similar to those of the loop in ROG2 RNA. The loop was located in archaebacterial rRNA from the large ribosomal subunit (LSU) [60, 61]. Three adenosine and cytosine residues formed one cross-strand stack, whereas guanosine and cytosine formed a second stack (Supplementary Fig. S8). In general, the unpaired 3 × 3 purine loop can affect the structure and function of RNA molecules in many ways by: (i) maintaining the double stranded character of RNA in terms of stacking interactions, (ii) inducing of bending of RNA helix, (iii) exposing nucleobases to the solvent moiety, (iv) involving in tertiary and quaternary interactions, or (v) serving as a ligand-binding site. In the rRNA of LSU, the two stacked adenosine residues were engaged in interactions with the neighboring RNA helix, whereas in ROG2, the residues of the 3 × 3 loop were either involved in the crystal lattice contacts and/or bound ligand molecules (NCD or $[Co(NH_3)_6]^{3+}$).

Local stabilization of the RNA structure by NCD and cobaltamine molecules

The binding of cobaltammine and NCD molecules has influence on the thermal stability of the internal loop of RNA. The UV-melting and DSC spectra of unliganded ROG1 and

ROG2 showed two transition points (Fig. 5). The first peak corresponds to the melting of the internal UGGAA loop, while the second peak represents the denaturation of the double-stranded region of the duplex. The shift of the first peak observed in the presence of $[Co(NH_3)]^{3+}$ supports its role in local stabilization of the internal UGGAA loop. This interpretation is consistent with crystallographic data of the unliganded ROG2 oligomer, which shows direct interactions between $[Co(NH_3)]^{3+}$ and the guanosine residues of the UGGAA motif. In the presence of the NCD ligand, the first peak was no longer distinct while the second peak became asymmetric in the UV and DSC profiles. These alterations in the melting behavior imply that the first transition was shifted and merged with the second, indicating that NCD locally stabilizes the internal loop. In case of UV profile at the 10 μM ROG2–NCD concentration, the peaks were not merged. This can be explained by the fact that the melting point of RNA duplexes is concentration-dependent leading to the migration of the peaks [62].

NCD exhibits a molecular glue characteristics

Our previous studies showed that NCD and NCD derivatives serve as molecular glue for the assembly of DNA molecules into large tetrahedron structures and modulation of the function of structured RNAs (ribozymes, riboswitches, and pseudoknots) [19–22]. In the present study, we showed that this property of NCD can be applied for crystal lattice formation. In the ROG2–NCD model, two NCD ligands bridged four neighboring RNA molecules into contact, leading to the formation of a crystal lattice in all directions. NCD is a promising molecular tool because it exhibits several important modalities. First, it possesses two selective and specific RNA binding sites for unpaired guanosine residues, forming strong and direct interactions with the target RNA. Second, the NCD linker is not only a passive element connecting NUs but also an important contributor to the selectivity and specificity of the ligand. It prefers the elongated zigzag conformation and has the ability to form hydrogen bonds with RNA having an amino group (Fig. 1E). The zigzag conformation results in a linker length of ∼7 Å, which is exactly double the value of the rise parameter corresponding to the distance between the neighboring bases in the RNA helix. Therefore, the gap between the two NCD-G pairs could be filled by the NCD-G pair of the second ligand. This observation is in agreement with recent results aimed at optimizing the linker length [63]. The longer linker resulted in weaker NCD binding, while the reduction of one atom (from 11 to 10 atoms) caused stronger binding than the parental NCD molecule. In the structure of ROG1–NCD, the zigzag line of the linker contains six or seven atoms, suggesting that the removal of one atom could limit the conformational freedom of the linker and increase the affinity of the ligand. Third, NCD prefers a tandem mode of binding, which means that exclusively two NCD molecules interact with RNA molecules. The resulting four NCD-G pairs alternate, and this arrangement is stabilized by interactions of the amino groups of the NCD linkers with the N8 imino group of guanosines. Fourth, the protonation state of NCD remains stable under neutral, slightly acidic, and slightly basic pH conditions, indicating that the ligand maintains its binding properties in a physiological environment. Due the presence of the carbamate group (R-NH-CO-O-R'), under strongly acidic or basic conditions, the molecule undergoes decomposition.

The crystallization of RNA molecules is challenging. Successful crystallization depends on the proper design of the RNA sequence which will induce contacts between the RNA molecules and crystal lattice formation. One of the methods to enhance RNA crystallization was to use 5′ or 3′ overhanging nucleotides that base-pair with the other end of the neighboring RNA molecule forming pseudo-infinite helices [64]. The use of NCD as a "molecular glue" for crystal lattice formation may be more advantageous because it connects four RNA molecules instead of two, as in the case of overhanging nucleotides. This resulted in formation of crystal lattice in all directions and improved the data resolution from 2.55 Å (ROG2–NCD) to 1.5 Å for ROG1–NCD. We believe that the exploitation of the NCD in RNA crystallography could contribute to determining the structures of many functional RNAs, increasing our knowledge of RNA structure and function relationship.

Acknowledgements

We are very grateful to Professor Gota Kawai of Chiba Institute of Technology for his time and the productive discussions, which were particularly beneficial in understanding the variations between NMR and crystallographic data.

Author contributions: M.M.-P.: Investigation, Formal analysis, Methodology, Visualization, Writing—original draft. L.B.: Investigation, Visualization, Writing—review & editing. M.B.: Investigation, Formal analysis. K.N.: Resources, Writing—review & editing. A.K.: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Supervision, Validation, Writing—review & editing

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

National Science Centre [UMO-2022/45/B/NZ7/03543, UMO-2017/26/E/NZ1/00950 to A.K.]. For Open Access publishing, the authors have applied a CC-BY public copyright license to any Author Accepted Manuscript (AAM) version arising from this submission. The synchrotron data were collected at P13 beamline operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany) [proposal number MX-989]. We would like to thank Selina Storm for the assistance in using the beamline. We acknowledge the European Synchrotron Radiation Facility (ESRF) for provision of synchrotron radiation facilities under proposal number mx-2432 (https://doi.org/10.15151/ESRF-ES-1028835001) and we would like to thank Igor Melnikov for assistance and support in using beamline ID30A-3. The access to ESRF was financed by the Polish Ministry of Science and Higher Education – decision no. 2021/WK/11. We acknowledge the MCB Structural Biology Core Facility (supported by the TEAM TECH CORE FACILITY/2017-4/6 grant from the Foundation for Polish Science) for valuable support. Funding to pay the Open Access publication

charges for this article was provided by National Science Centre [UMO-2022/45/B/NZ7/03543].

Data availability

Atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data Bank under accession numbers 9IF1, 9IF0 and 9I9W. Diffraction images have been deposited in Macromolecular Xtallography Raw Data Repository (MX-RDR, mxrdr.icm.edu.pl): doi:10.60884/QV3IGO for ROG2 RNA, doi:10.60884/NO2MXW for ROG2–NCD complex and doi:10.60884/SCBCTX for ROG1–NCD complex.

References

1. Carbone EJ, Jiang T, Nelson C *et al.* Small molecule delivery through nanofibrous scaffolds for musculoskeletal regenerative engineering. *Nanomed Nanotechnol Biol Med* 2014;**10**:1691–9. https://doi.org/10.1016/j.nano.2014.05.013
2. Chan KH, Wang Y, Zheng B *et al.* RNA-selective small-molecule ligands: recent advances in live-cell imaging and drug discovery. *Chem Med Chem* 2023;**18**:e202300271. https://doi.org/10.1002/cmdc.202300271
3. Pandey M, Fernandez M, Gentile F *et al.* The transformational role of GPU computing and deep learning in drug discovery. *Nat Mach Intell* 2022;**4**:211–21. https://doi.org/10.1038/s42256-022-00463-x
4. Sadybekov AV, Katritch V. Computational approaches streamlining drug discovery. *Nature* 2023;**616**:673–85. https://doi.org/10.1038/s41586-023-05905-z
5. Wu H, Bagal SK, Jin M. Directly targeting RNA with drug-like small molecules. In Chan HM, Arrowsmith C. (eds), *Epigenetic Drug Discovery*. Royal Society of Chemistry, Cambridge, United Kingdom, 2024, pp. 529–60. https://doi.org/10.1039/9781837674916
6. Morishita EC. Discovery of RNA-targeted small molecules through the merging of experimental and computational technologies. *Expert Opin Drug Discovery* 2023;**18**:207–26. https://doi.org/10.1080/17460441.2022.2134852
7. Tessaro F, Scapozza L. How 'protein-docking' Translates into the new emerging field of docking small molecules to nucleic acids? *Molecules* 2020;**25**:2749. https://doi.org/10.3390/molecules25122749
8. Błaszczyk L, Ryczek M, Das B *et al.* Antisense RNA C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms a triplex-like structure and binds small synthetic ligand. *Nucleic Acids Res* 2024;**52**:6707–17. https://doi.org/10.1093/nar/gkae376
9. Nakatani K. Recognition of mismatched base pairs in DNA. *Bull Chem Soc Jpn* 2009;**82**:1055–69. https://doi.org/10.1246/bcsj.82.1055
10. Nakatani K, Sando S, Kumasawa H *et al.* Recognition of guanine−guanine mismatches by the dimeric form of 2-amino-1, 8-naphthyridine. *J Am Chem Soc* 2001;**123**:12650–7. https://doi.org/10.1021/ja0109186
11. Mukherjee S, Błaszczyk L, Rypniewski W *et al.* Structural insights into synthetic ligands targeting A–A pairs in disease-related CAG RNA repeats. *Nucleic Acids Res* 2019;**47**:10906–13. https://doi.org/10.1093/nar/gkz832
12. Nakamori M, Panigrahi GB, Lanni S *et al.* A slipped-CAG DNA-binding small molecule induces trinucleotide-repeat contractions in vivo. *Nat Genet* 2020;**52**:146–59. https://doi.org/10.1038/s41588-019-0575-8
13. Mukherjee S, Dohno C, Nakatani K. Design and synthesis of cyclic mismatch-binding ligands (CMBLs) with variable linkers by ring-closing metathesis and their photophysical and DNA repeat binding properties. *Chem A Eur J* 2017;**23**:11385–96. https://doi.org/10.1002/chem.201702064
14. Nakatani K, Hagihara S, Goto Y *et al.* Small-molecule ligand induces nucleotide flipping in (CAG)n trinucleotide repeats. *Nat Chem Biol* 2005;**1**:39–43. https://doi.org/10.1038/nchembio708
15. Yamada T, Furuita K, Sakurabayashi S *et al.* NMR determination of the 2:1 binding complex of naphthyridine carbamate dimer (NCD) and CGG/CGG triad in double-stranded DNA. *Nucleic Acids Res* 2022;**50**:9621–31. https://doi.org/10.1093/nar/gkac740
16. Shibata T, Nakamachi A, Nakatani K. Enhancing binding affinity to CGG/CGG triad: optimizing naphthyridine carbamate dimer derivatives with varied linker lengths. *ChemMedChem* 2024;**19**:e202400351. https://doi.org/10.1002/cmdc.202400351
17. Shibata T, Nagano K, Ueyama M *et al.* Small molecule targeting r(UGGAA)n disrupts RNA foci and alleviates disease phenotype in Drosophila model. *Nat Commun* 2021;**12**:236. https://doi.org/10.1038/s41467-020-20487-4
18. Niimi Y, Takahashi M, Sugawara E *et al.* Abnormal RNA structures (structures (RNA foci) containing a penta-nucleotide repeat (foci) containing a penta-nucleotide repeat $(UGGAA)_n$ in the P urkinje cell nucleus is associated with spinocerebellar ataxia type 31 pathogenesis urkinje cell nucleus is associated with spinocerebellar ataxia type 31 pathogenesis. *Neuropathology* 2013;**33**:600–11. https://doi.org/10.1111/neup.12032
19. Hong C, Hagihara M, Nakatani K. Ligand-assisted complex formation of two DNA hairpin loops. *Angew Chem Int Ed* 2011;**50**:4390–3. https://doi.org/10.1002/anie.201100075
20. Hong C, Otabe T, Matsumoto S *et al.* Formation of a Ligand-assisted complex of two RNA hairpin loops. *Chem A Eur J* 2014;**20**:5282–7. https://doi.org/10.1002/chem.201304683
21. Dohno C, Nakatani K. Control of DNA hybridization by photoswitchable molecular glue. *Chem Soc Rev* 2011;**40**:5718. https://doi.org/10.1039/c1cs15062f
22. Dohno C, Atsumi H, Nakatani K. Ligand inducible assembly of a DNA tetrahedron. *Chem Commun* 2011;**47**:3499. https://doi.org/10.1039/c0cc05473a
23. Ryczek M, Pluta M, Błaszczyk L *et al.* Overview of methods for large-scale RNA synthesis. *Appl Sci* 2022;**12**:1543. https://doi.org/10.3390/app12031543
24. Kiliszek A, Pluta M, Bejger M *et al.* Structure and thermodynamics of a UGG motif interacting with Ba^{2+} and other metal ions: accommodating changes in the RNA structure and the presence of a G(*syn*)–G(*syn*) pair. *RNA* 2023;**29**:44–54. https://doi.org/10.1261/rna.079414.122
25. Rypniewski W, Banaszak K, Kuliński T *et al.* Watson–Crick-like pairs in CCUG repeats: evidence for tautomeric shifts or protonation. *RNA* 2016;**22**:22–31. https://doi.org/10.1261/rna.052399.115
26. Kiliszek A, Banaszak K, Dauter Z *et al.* The first crystal structures of RNA–PNA duplexes and a PNA-PNA duplex containing mismatches—Toward anti-sense therapy against TREDs. *Nucleic Acids Res* 2016;**44**:1937–43. https://doi.org/10.1093/nar/gkv1513
27. Gerlach M, Mueller U, Weiss MS. The MX beamlines BL14.1-3 at BESSY II. *JLSRF* 2016;**2**:A47. https://doi.org/10.17815/jlsrf-2-64
28. Mueller U, Förster R, Hellmig M *et al.* The macromolecular crystallography beamlines at BESSY II of the Helmholtz–Zentrum Berlin: current status and perspectives. *Eur Phys J Plus* 2015;**130**:141. https://doi.org/10.1140/epjp/i2015-15141-2
29. Mueller U, Darowski N, Fuchs MR *et al.* Facilities for macromolecular crystallography at the Helmholtz–Zentrum Berlin. *J Synchrotron Rad* 2012;**19**:442–9. https://doi.org/10.1107/S0909049512006395
30. Burkhardt A, Pakendorf T, Reime B *et al.* Status of the crystallography beamlines at PETRA III. *Eur Phys J Plus* 2016;**131**:56. https://doi.org/10.1140/epjp/i2016-16056-0
31. Cianci M, Bourenkov G, Pompidor G *et al.* P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable

beam focusing. *J Synchrotron Rad* 2017;**24**:323–32. https://doi.org/10.1107/S1600577516016465
32. Raimondi P. ESRF-EBS: the extremely brilliant source project. *Synchrotron Radiation News* 2016;**29**:8–15. https://doi.org/10.1080/08940886.2016.1244462
33. Pernot P, Theveneau P, Giraud T *et al.* New beamline dedicated to solution scattering from biological macromolecules at the ESRF. *J Phys Conf Ser* 2010;**247**:012009. https://doi.org/10.1088/1742-6596/247/1/012009
34. Kabsch W. XDS. *Acta Crystallogr D Biol Crystallogr* 2010;**66**:125–32. https://doi.org/10.1107/S0907444909047337
35. McCoy AJ, Grosse-Kunstleve RW, Adams PD *et al.* Phaser crystallographic software. *J Appl Crystallogr* 2007;**40**:658–74. https://doi.org/10.1107/S0021889807021206
36. Bunkóczi G, Echols N, McCoy AJ *et al.* Phaser.MRage: automated molecular replacement. *Acta Crystallogr D Biol Crystallogr* 2013;**69**:2276–86. https://doi.org/10.1107/S0907444913022750
37. Kiliszek A, Kierzek R, Krzyzosiak WJ *et al.* Structural insights into CUG repeats containing the 'stretched U–U wobble': implications for myotonic dystrophy. *Nucleic Acids Res* 2009;**37**:4149–56. https://doi.org/10.1093/nar/gkp350
38. Winn MD, Ballard CC, Cowtan KD *et al.* Overview of the *CCP* 4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 2011;**67**:235–42. https://doi.org/10.1107/S0907444910045749
39. Vagin AA, Steiner RA, Lebedev AA *et al.* *REFMAC* 5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr D Biol Crystallogr* 2004;**60**:2184–95. https://doi.org/10.1107/S0907444904023510
40. Murshudov GN, Skubák P, Lebedev AA *et al.* *REFMAC* 5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 2011;**67**:355–67. https://doi.org/10.1107/S0907444911001314
41. Kovalevskiy O, Nicholls RA, Long F *et al.* Overview of refinement procedures within *REFMAC* 5: utilizing data from different sources. *Acta Crystallogr D Struct Biol* 2018;**74**:215–27. https://doi.org/10.1107/S2059798318000979
42. Nicholls RA, Long F, Murshudov GN. Low-resolution refinement tools in *REFMAC* 5. *Acta Crystallogr D Biol Crystallogr* 2012;**68**:404–17. https://doi.org/10.1107/S090744491105606X
43. Emsley P, Lohkamp B, Scott WG *et al.* Features and development of *Coot*. *Acta Crystallogr D Biol Crystallogr* 2010;**66**:486–501. https://doi.org/10.1107/S0907444910007493
44. Emsley P, Cowtan K. *Coot* : model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 2004;**60**:2126–32. https://doi.org/10.1107/S0907444904019158
45. Zheng G, Lu X-J, Olson WK. Web 3DNA–a web server for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid structures. *Nucleic Acids Res* 2009;**37**:W240–W246. https://doi.org/10.1093/nar/gkp358
46. Lu X-J. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res* 2003;**31**:5108–21. https://doi.org/10.1093/nar/gkg680
47. Pasternak A, Wengel J. Thermodynamics of RNA duplexes modified with unlocked nucleic acid nucleotides. *Nucleic Acids Res* 2010;**38**:6697–706. https://doi.org/10.1093/nar/gkq561
48. Nakatani K. Possibilities and challenges of small molecule organic compounds for the treatment of repeat diseases. *Proc Jpn Acad Ser B* 2022;**98**:30–48. https://doi.org/10.2183/pjab.98.003
49. Shibata T, Nakatani K. A small molecule binding to TGGAA pentanucleotide repeats that cause spinocerebellar ataxia type 31. *Bioorg Med Chem Lett* 2023;**79**:129082. https://doi.org/10.1016/j.bmcl.2022.129082
50. Yamada T, Sakurabayashi S, Sugiura N *et al.* NMR analysis of^{15} N-labeled naphthyridine carbamate dimer (NCD) to contiguous CGG/CGG units in DNA. *Chem Commun* 2024;**60**:3645–8. https://doi.org/10.1039/D4CC00544A
51. Yamada T, Miki S, Ni L *et al.* CGG repeat DNA assisted dimerization of CGG/CGG binding molecule through intermolecular disulfide formation. *Chem Commun* 2018;**54**:13072–5. https://doi.org/10.1039/C8CC06757K
52. Ni L, Yamada T, Murata A *et al.* Mismatch binding ligand upregulated back-splicing reaction producing circular RNA in a cellular model. *Chem Commun* 2022;**58**:3629–32. https://doi.org/10.1039/D1CC06936E
53. Holbrook SR, Cheong C, Tinoco I *et al.* Crystal structure of an RNA double helix incorporating a track of non-Watson–Crick base pairs. *Nature* 1991;**353**:579–81. https://doi.org/10.1038/353579a0
54. Tanaka Y. A'-form RNA double helix in the single crystal structure of r(UGAGCUUCGGCUC). *Nucleic Acids Res* 1999;**27**:949–55. https://doi.org/10.1093/nar/27.4.949
55. Saenger W. *Principles of Nucleic Acid Structure*. Springer New York, New York, NY, 1984.
56. Nichols PJ, Krall JB, Henen MA *et al.* Z-RNA biology: a central role in the innate immune response? *RNA* 2023;**29**:273–81. https://doi.org/10.1261/rna.079429.122
57. Yamada T, Furuita K, Sakurabayashi S *et al.* NMR determination of the 2:1 binding complex of naphthyridine carbamate dimer (NCD) and CGG/CGG triad in double-stranded DNA. *Nucleic Acids Res* 2022;**50**:9621–31. https://doi.org/10.1093/nar/gkac740
58. Ennifar E, Dumas P. Polymorphism of bulged-out residues in HIV-1 RNA DIS kissing complex and structure comparison with solution studies. *J Mol Biol* 2006;**356**:771–82. https://doi.org/10.1016/j.jmb.2005.12.022
59. Popenda M, Szachniuk M, Blazewicz M *et al.* RNA FRABASE 2.0: an advanced web-accessible database with the capacity to search the three-dimensional fragments within RNA structures. *BMC Bioinformatics* 2010;**11**:231. https://doi.org/10.1186/1471-2105-11-231
60. Gabdulkhakov A, Nikonov S, Garber M. Revisiting the *Haloarculamarismortui* 50S ribosomal subunit model. *Acta Crystallogr D Biol Crystallogr* 2013;**69**:997–1004. https://doi.org/10.1107/S0907444913004745
61. Petrov AS, Bernier CR, Gulen B *et al.* Secondary structures of rRNAs from all three domains of life. *PLoS One* 2014;**9**:e88222. https://doi.org/10.1371/journal.pone.0088222
62. Xia T, Mathews DH, Turner DH. Thermodynamics of RNA secondary structure formation. In *Comprehensive Natural Products Chemistry*. Elsevier, London, United Kingdom, 1999, pp.21–47.
63. Shibata T, Nakamachi A, Nakatani K. Enhancing binding affinity to CGG/CGG triad: optimizing naphthyridine carbamate dimer derivatives with varied linker lengths. *ChemMedChem* 2024;**19**:e202400351. https://doi.org/10.1002/cmdc.202400351
64. Ferré-D'Amaré AR, Doudna JA. Methods to crystallize RNA. *Curr Protoc Nucleic Acid Chem* 2000;**1**:7.6.1. https://doi.org/10.1002/0471142700.nc0706s00

Received: April 17, 2025. **Revised:** August 18, 2025. **Accepted:** August 27, 2025

Naphthyridine carbamate dimer ligand induces formation of Z-RNA-like fold of disease-related RNA and exhibits a molecular glue characteristics in crystal lattice formation

Martyna Mateja-Pluta[1], Leszek Błaszczyk[1],Magdalena Bejger[1], Kazuhiko Nakatani[2],

Agnieszka Kiliszek[1]*

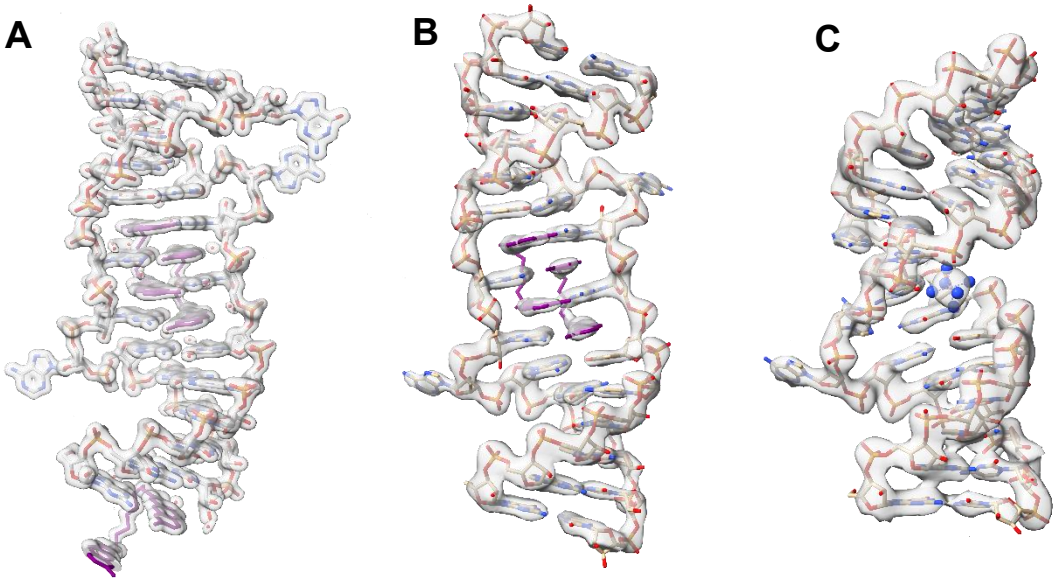
[1] Institute of Bioorganic Chemistry, Polish Academy of Sciences, Z. Noskowskiego 12/14, 61-704, Poland

[2] Department of Regulatory Bioorganic Chemistry, SANKEN (The Institute of Scientific and Industrial Research), Osaka University, 8-1 Mihogaoka, Ibaraki, 567-0047, Japan

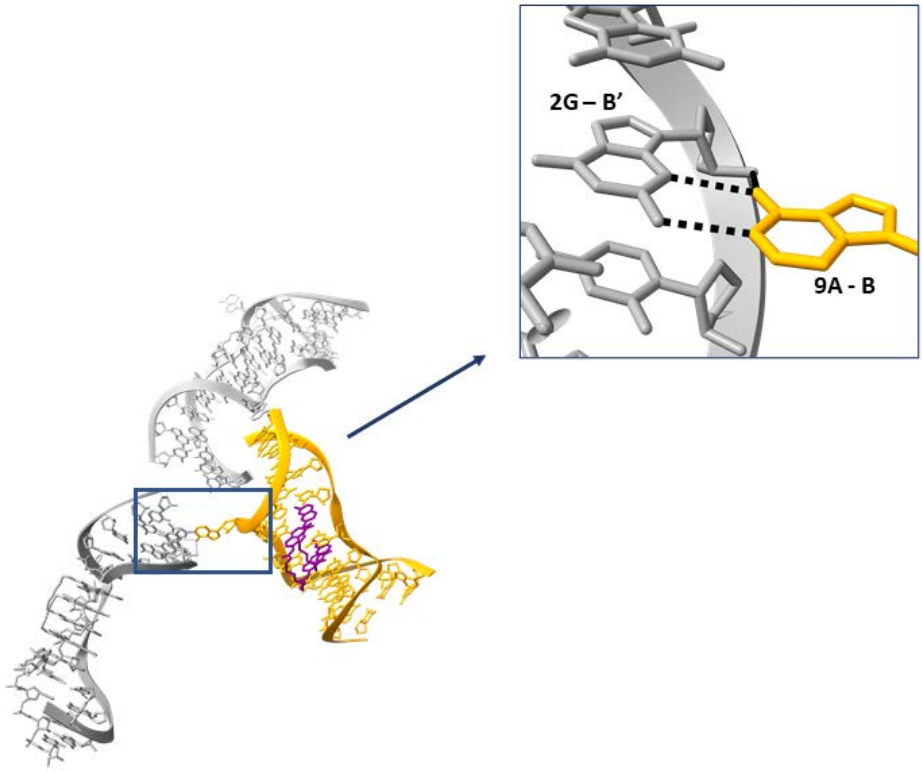
* To whom correspondence should be addressed. Tel: +48 61 852 8503 ext. 1278; Fax: +48 61 852 05 32; Email: kiliszek@ibch.poznan.pl

	ROG2	ROG2-NCD	ROG1-NCD
Data Collection			
X-ray source	Eiger1 X 4M Mx 2432	PILATUS 6M	Dectris PILATUS 6M-F,
Space group	$P4_12_12$	C121	$P4_12_12$
Cell parameters (Å)	a = 94.46, b = 25.31, c = 45.41	a = 94,46, b = 25.31, c = 45.41	a = 71.63, b = 71.63, c = 62.88
Resolution (Å)	42.52-2.79 (2.96-2.79)	46.11-2.55 (2.7-2.55)	47.3-1.5 (1.59-1.5)
R_{merge}	0.059 (0.364)	0.066 (1.159)	0.085 (1.792)
I/σ	16.72 (2.64)	16.17 (1.73)	24.64 (1.91)
$CC_{1/2}$	0.999 (0.982)	0.99 (0.844)	1.0 (0.89)
Completeness (%)	97.3 (82.3)	98.4 (98.9)	100 (100)
Redundancy	6.26 (5.84)	6.03 (6.08)	25.7 (24.9)
Number of unique reflections	2764	3544	26777
Refinement			
Software	Refmac 5.8.0258	Refmac 5.8.0258	Refmac 5.8.0258
Number of reflections: work/test	2485/277	3358/185	25879/899
Overall mean B value ($Å^2$)	28.256	65.69	26
R_{work}/R_{free}	0.2625/0.3085	0.201/0.2397	0.129/0.166
RNA atoms	656	725	1017
Water molecules	6	6	203
Ligand molecules	NCO	NCD	NCD
RMSD in bonds (Å)	0.009	0.008	0.012
RMSD in angles (°)	1.742	1.836	1.616
PDB code	9IF1	9IF0	9I9W
x-ray images	10.60884/QV3IGO	doi:10.60884/NO2MXW	doi:10.60884/SCBCTX

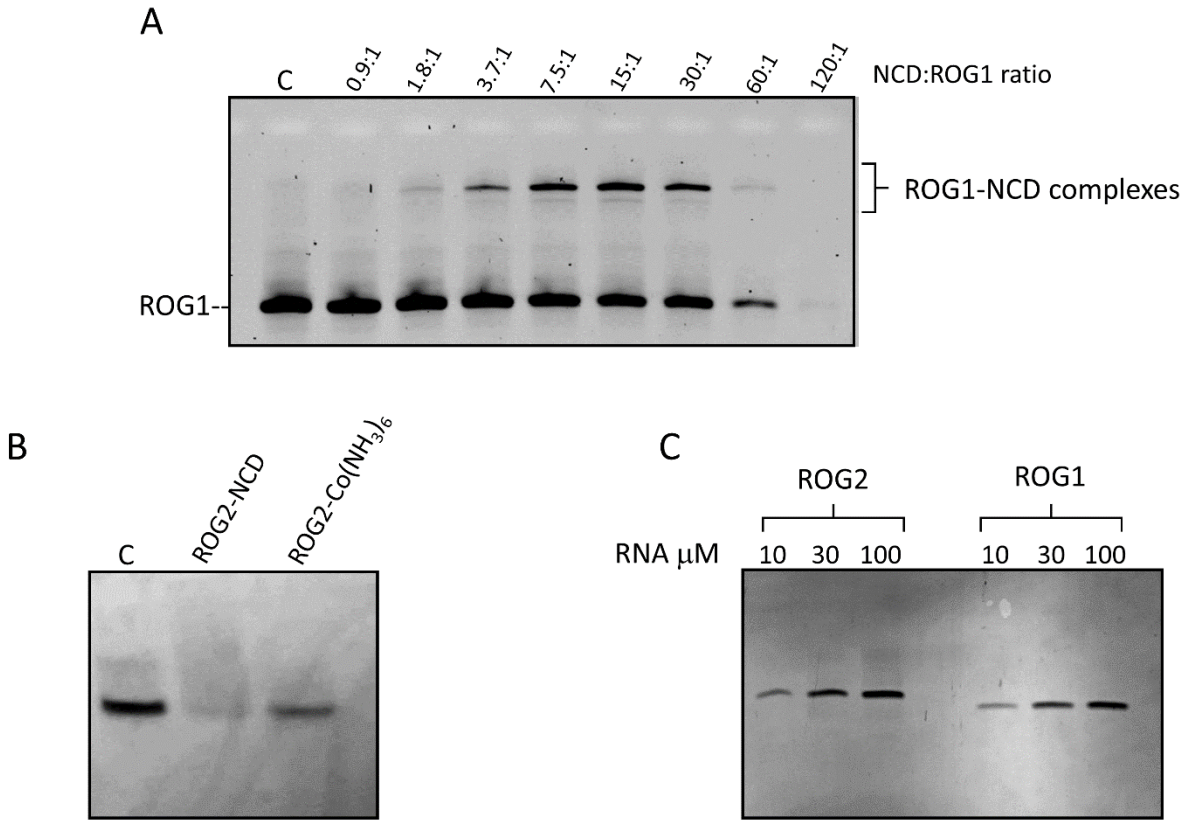
Supplementary Table S1: Summary of X-ray data and model refinement statistics.



Supplementary Figure S1: The $F_o - F_c$ omit map of (A) ROG1-NCD, (B) ROG2-NCD and (C) ROG2 contoured at the σ 1.0 level.



Supplementary Figure S2: The interactions between the flipped out 9A residue forming crystal lattice contacts in ROG1-NCD. The H-bonds are represented by black dashed lines.

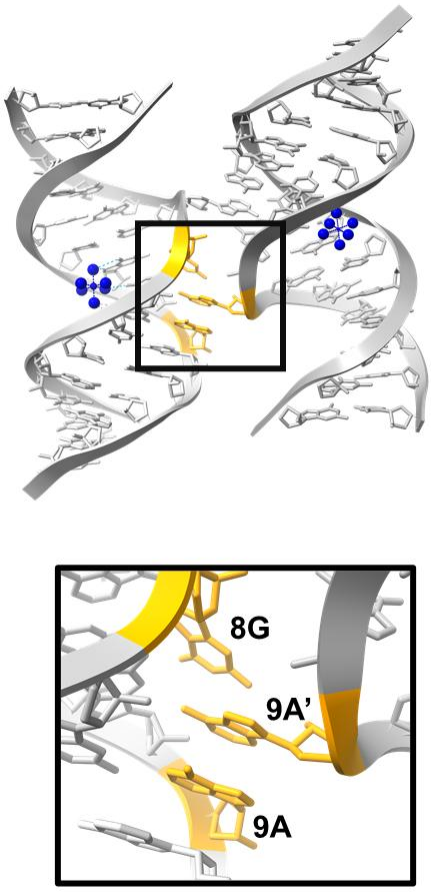


Supplementary Figure S3: Native and denaturing gel electrophoresis of RNA and RNA-ligand complexes. (A) Electrophoretic mobility shift assay of ROG1-NCD complexes. C – control reaction (unliganded ROG1 duplex). (B) Denaturing gel electrophoresis of ROG2-NCD and ROG2-$Co(NH_3)_6$ complexes after DSC measurements. C – control reaction (unliganded ROG2 duplex). (C) Native gel electrophoresis of ROG2 and ROG1 oligomers in three concentrations: 10, 30 and 100 μM.

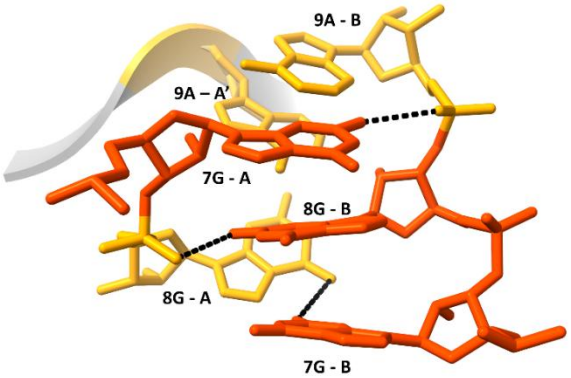
ROG2						
bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 G-C	-0,58	-0,18	0,18	-10,07	-15,38	3,22
2 G-C	0,02	-0,14	-0,05	-8,74	-14,88	-4,75
3 C-G	-0,03	-0,21	-0,09	4,8	-11,28	2,77
4 A-U	-0,06	-0,23	0,58	3,75	-10,31	-5,4
5 C-G	0,16	0	-0,1	2,95	-21,01	1,18
6 U-A	0,91	-0,01	0,58	-9,36	-7,9	2,86
* 7 G-G	6,75	3,89	0,16	37,3	-31,25	-59,45
* 8 A+G	4,33	4,39	0,37	-7,65	13,6	104,46
9 A-U	-0,34	-0,09	-0,29	-9,22	4,95	-6,6
10 G-C	-0,79	-0,02	0,01	-4,74	-9,46	1,45
11 U-A	0,16	-0,19	0,38	-9,08	-5,67	-2,78
12 G-C	0,35	-0,13	0,21	4,37	-8,8	-1,19
13 C-G	0,36	-0,3	0,45	0,87	-8,54	-0,81
14 C-G	0,8	-0,47	0,43	0,38	-5,27	-0,09
Ave.	0,04	-0,19	0,20	-1,55	-11,06	-0,64

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 GG/CC	-0,44	-1,71	3,29	-1,25	3,98	29,54
2 GC/GC	-0,21	-1,95	2,96	-3,28	-1,16	29,66
3 CA/UG	-0,49	-1,68	3,1	-7,28	6,21	33,58
4 AC/GU	0,28	-1,45	3,09	6,16	4,27	32,83
5 CU/AG	0,22	-1,42	3,55	-4,08	3,69	34,05
* 6 UG/GA	-2,75	-4,35	3,37	-14,72	18,48	29,64
* 7 GA/GG	0,35	0,31	6,9	-3,83	3,5	93,38
* 8 AA/UG	-1,06	-1,01	3,61	2,96	1,62	8,13
9 AG/CU	0,57	-2,02	3	-3,15	15,15	28,08
10 GU/AC	-0,6	-1,68	3,26	-1,89	6,38	32,24
11 UG/CA	0,19	-1,41	2,76	2,9	8,07	32,91
12 GC/GC	0,4	-1,65	3,19	-0,32	9,05	33,36
13 CC/GG	0,35	-1,64	3,18	1,35	6,99	31,36
Ave.	0,03	-1,66	3,14	-1,08	6,26	31,76

Supplementary Table S2: Helical parameters calculated using 3DNA, based on C1'- C1' vectors for unliganded structure ROG2.



Supplementary Figure S4: Crystal lattice contacts between two neighboring RNA molecules of unliganded structure ROG2. The flipped out 9A' residue stacks between 8G and 9A residues from UGGAA motif.



Supplementary Figure S5: Hydrogen bonds observed in cross-strand stacks in unliganded ROG2 structure

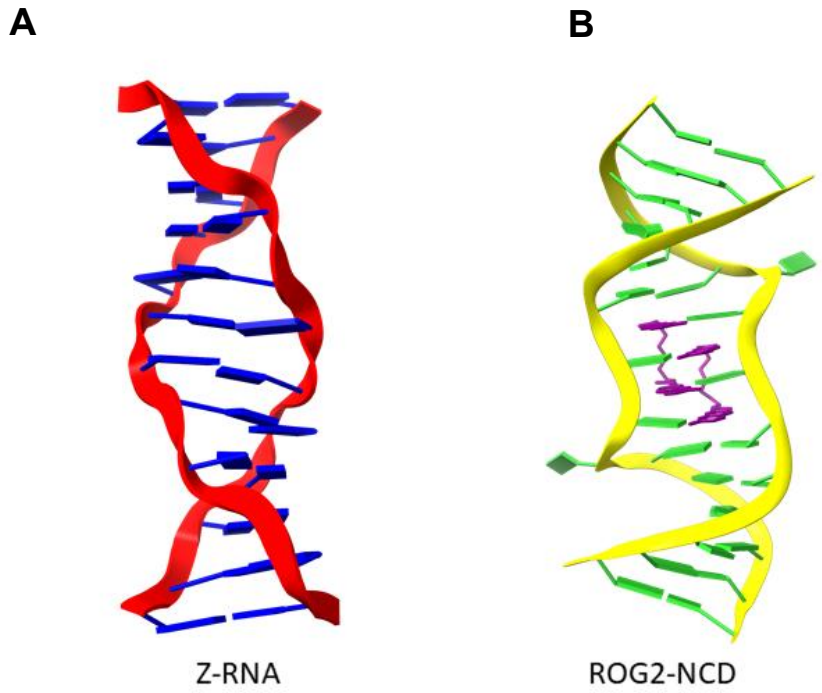
ROG 2-NCD						
bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 G-C	0,15	-0,13	0,46	-1,69	1,53	-2,01
2 G-C	-0,17	0,04	0,02	-1,21	-12,58	4,53
3 C-G	-0,13	-0,23	0,24	4,07	-13,51	-6,11
4 A-U	-0,35	-0,16	-0,56	-3,78	-18,63	-3,62
5 C-G	0,03	-0,04	0	-1,82	-25,78	3,12
6 U-A	0,17	0,28	0,15	-4,12	-16,42	6,19
7 A-U	0,12	-0,03	-0,04	8,5	-12,92	2,09
8 G-C	-0,41	0,03	-0,03	-3,3	-17,09	4,5
9 U-A	0,51	-0,22	0,64	-3,92	-18,25	6,09
10 G-C	0,18	0,01	0,1	-7,44	-12,65	2,34
11 C-G	0,42	-0,38	0,36	-4,26	-13,72	3,3
12 C-G	0,08	-0,02	0,16	0,89	-5,28	2,36
Ave.	0,03	-0,11	0,14	-2,25	-13,60	1,45

ROG1-NCD						
bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 G-C	-0,24	-0,11	0,02	-6,09	-8,78	-0,87
2 C-G	0,18	-0,16	0,08	1,58	-13,05	-1,4
3 A-U	0,04	-0,09	0,09	0,57	-12,28	1,74
4 C-G	0,26	-0,14	0,1	1,37	-18,58	1,2
5 U-A	-0,03	-0,08	0,17	-2,51	-4,25	-0,41
6 A-U	-0,31	0	0,08	0,46	-3,27	1,1
7 G-C	-0,28	-0,15	0	-0,4	-15,98	2,15
8 U-A	-0,01	-0,13	-0,01	-1,51	-13,63	5,07
9 G-C	-0,15	-0,14	-0,06	-5,64	-10,08	-0,27
10 C-G	0,05	-0,17	-0,11	1,62	-5,58	-0,66
Ave.	-0,02	-0,14	0,01	-1,06	-12,25	0,87

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 GG/CC	-0,3	-1,95	3,33	2,88	4,54	29,84
2 GC/GC	-0,3	-1,44	3,2	-4,45	1,13	33,14
3 CA/UG	-0,65	-1,46	3,35	3,19	13,22	29,63
4 AC/GU	0,07	-1,34	3,25	-3,46	5,18	31,96
5 CU/AG	0,61	-1,1	3,3	1,94	7,23	34,88
6 UA/UA	----	----	----	----	----	----
7 AG/CU	-0,51	-1,31	3,57	-1,77	10,33	35,31
8 GU/AC	0,22	-1,12	3,3	-4,74	7,2	29,65
9 UG/CA	0,19	-1,35	3,12	4,47	7,6	34,63
10 GC/GC	0,5	-1,65	3,22	-0,37	1,34	30,45
11 CC/GG	-0,04	-2,2	3,03	0,98	4,06	29,61
Ave.	-0,02	-1,49	3,27	-0,13	6,18	31,91

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 GC/GC	-0,74	-1,35	3,13	-1,84	0,07	33,21
2 CA/UG	0,50	-1,46	3,18	0,41	8,46	31,08
3 AC/GU	0,08	-1,36	3,24	0,99	4,19	32,37
4 CU/AG	-0,04	-1,44	3,36	-0,34	4,61	35,45
5 UA/UA	----	----	----	----	----	----
6 AG/CU	-0,44	-1,71	3,32	-1,37	5,93	31,52
7 GU/AC	-0,03	-1,31	3,29	-0,83	6,23	31,49
8 UG/CA	-0,21	-1,59	3,33	0,35	11,01	30,98
9 GC/GC	0,81	-2,31	3,06	1,86	4,70	24,22
Ave.	-0,01	-1,57	3,24	-0,10	5,65	31,29

Supplementary Table S3: Helical parameters calculated using 3DNA, based on C1'- C1' vectors for liganded structures ROG1-NCD and ROG2-NCD.



Supplementary Figure S6: Comparison of crystal structures of (A) Z-RNA form (pdb code: 4OCB) and (B) ROG2-NCD model.

Torsion angles	7G - A	8G - A	7G - B	8G - B	Z-RNA (pdb code: 4OCB)
α (O3'—P—O5'—C5')	**172.6º**	**-90.1º**	**178.5º**	**-98.6º**	**-57.3º ***
ζ (C3'—O3'—P—O5')	**-61.7º**	**-65º**	**-59.6º**	**-66º**	**-66º**

Supplementary Table S4: Torsion angles of 7G and 8G residues in liganded structure in comparison to values of torsion angles in Z-RNA structure (pdb code: 4OCB). *Torsion angle values calculated for guanosine residues in Z-RNA model.

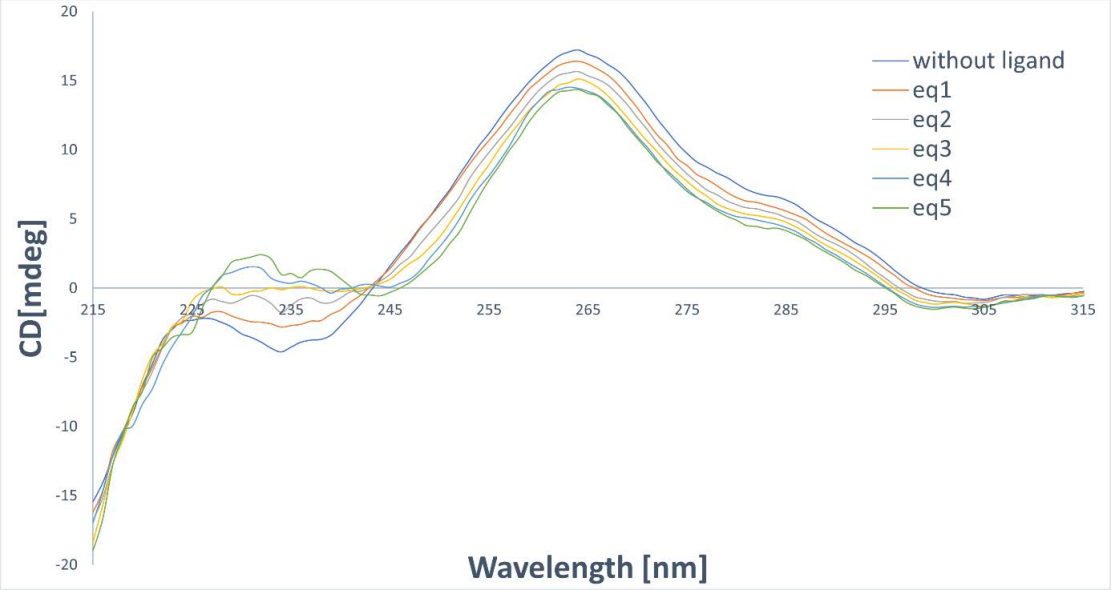
A

	ROG2	ROG2-NCD	ROG2-$Co(NH_3)_6$
Tm **(ºC)**	79	79	79.4

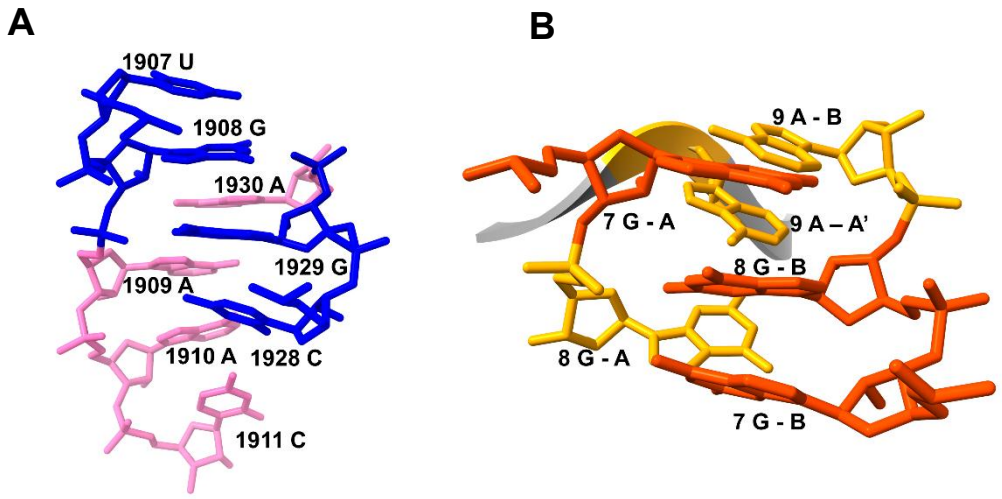
B

Tm (°C)	RNA concent.	ROG2	ROG2-NCD	ROG2-$Co(NH_3)_6$	ROG1	ROG1-NCD	ROG1-$Co(NH_3)_6$
Tm^1	100 µM	48	-	55	33	-	39
	30 µM	41	-	46	-	-	-
	10 µM	-	62	44	-	-	-
Tm^2	100 µM	80	77	80	65	70	67
	30 µM	80	77	79	67	65	65
	10 µM	79	81	80	69	62	68

Supplementary Table S5: Thermal stability of ROG1 and ROG2 duplexes determined by (A) Differential Scanning Calorimetry and (B) UV melting methods.



Supplementary Figure S7: Physicochemical evaluation of ROG2 duplex using circular dichroism. (D) The NCD ligand was titrated against ROG2 oligomer with molar equivalents (eq). The NCD concentration was in the range of 10-50 µM.



Archaebacterial rRNA from the large ribosomal subunit (LSU)

ROG2 RNA

Supplementary figure S8: Internal loops located in (A) archaebacterial rRNA from the large ribosomal subunit (LSU) and in (B) unliganded structure ROG2.

INSTITUTE OF BIOORGANIC CHEMISTRY
Polish Academy of Sciences

Poznań, 28-11-2025

Doctoral student's statement regarding contribution to the creation of the scientific work included in the doctoral dissertation

Publication title: Overview of Methods for Large-Scale RNA Synthesis

Authors: Marcin Ryczek, Martyna Pluta, Leszek Błaszczyk*, Agnieszka Kiliszek*

Journal: Applied Sciences

Date of publication: 2022

I declare that my contribution to the creation of this paper included: performing experiments, preparing research reports, analyzing and interpreting the results, and participating in the preparation of the initial version of the chapter concerning chemical synthesis of RNA.

Martyna Matejo-Pluta

(Signature)

* Corresponding authors

INSTITUTE OF BIOORGANIC CHEMISTRY
Polish Academy of Sciences

Poznań, 28-11-2025

Statement of the corresponding author regarding the doctoral student's contribution to the creation of the scientific work included in the doctoral dissertation

Publication title: Overview of Methods for Large-Scale RNA Synthesis

Authors: Marcin Ryczek, Martyna Pluta, Leszek Błaszczyk*, Agnieszka Kiliszek*

Journal: Applied Sciences

Date of publication: 2022

I declare that the contribution of MSc Martyna Mateja-Pluta to the creation of this paper included: performing experiments, preparing research reports, analyzing and interpreting the results and participating in the preparation of the initial version of the chapter concerning chemical synthesis of RNA.

Agnieszka Kiliszek

(Signature)

* Corresponding authors

INSTITUTE OF BIOORGANIC CHEMISTRY
Polish Academy of Sciences

Poznań, 28-11-2025

Doctoral student's statement regarding contribution to the creation of the scientific work included in the doctoral dissertation

Publication title: Naphthyridine carbamate dimer ligand induces formation of Z-RNA-like fold of disease-related RNA and exhibits a molecular glue characteristics in crystal lattice formation

Authors: Martyna Mateja-Pluta, Leszek Błaszczyk, Magdalena Bejger, Kazuhiko Nakatani, Agnieszka Kiliszek*

Journal: Nucleic Acid Research

Date of publication: 2025

I declare that my contribution to the creation of this paper included: performing experiments, preparing research reports, analyzing and interpreting the results, preparing the initial version of the article, providing responses to the reviewers, and preparing the final version of the manuscript.

Martyna Mateja-Pluta

(Signature)

* Corresponding author

INSTITUTE OF BIOORGANIC CHEMISTRY
Polish Academy of Sciences

Poznań, 28-11-2025

Statement of the corresponding author regarding the doctoral student's contribution to the creation of the scientific work included in the doctoral dissertation

Publication title: Naphthyridine carbamate dimer ligand induces formation of Z-RNA-like fold of disease-related RNA and exhibits a molecular glue characteristics in crystal lattice formation

Authors: Martyna Mateja-Pluta, Leszek Błaszczyk, Magdalena Bejger, Kazuhiko Nakatani, Agnieszka Kiliszek*

Journal: Nucleic Acid Research

Date of publication: 2025

I declare that the contribution of MSc Martyna Mateja-Pluta to the creation of this paper included: performing experiments, preparing research reports, analyzing and interpreting the results, preparing the initial version of the article, providing responses to the reviewers, and preparing the final version of the manuscript.

Agnieszka Kiliszek

(Signature)

* Corresponding author