Charles University

Faculty of Science

Study program: Chemistry

Bachelor's study program: Medicinal Chemistry

Karolína Pšenáková

Characterization of thermodynamically stabilizing mutations in a light-producing deoxyribozyme using NMR spectroscopy

Charakterizace termodynamicky stabilizujících mutací ve světlo-produkujícím deoxyribozymu pomocí NMR spektroskopie

Bachelor's thesis

Supervisor: RNDr. Zdeněk Tošner, Ph.D.
Advisor: Edward Arthur Curtis, Ph.D.

Prague, 2021

Prehlasujem, že som záverečnú prácu vypracovala samostatne a že som uviedla všetky použité informačné zdroje a literatúru. Táto práca, ani jej podstatná časť, nebola použitá k získaniu iného alebo rovnakého akademického titulu.

V Prahe, dňa 23.08.2021

……..

Karolína Pšenáková

Abstrakt

Rovnako ako proteíny, aj DNA molekuly sú schopné katalýzy chemických reakcií. Motivovaná potenciálom širokého využitia katalytickej DNA, skupina Dr. Curtisa identifikovala svetlo-produkujúci deoxyribozým pomocou tzv. *in vitro* evolúcie. Naším súčasným cieľom je získať štruktúru tohto deoxyribozýmu na atomárnej úrovni.

Rovanko ako v prípade proteínov, aj charakterizácia enzýmov tvorených nukleovými kyselinami je často uľahčená prítomnosťou mutácií, ktoré zvyšujú ich stabilitu pri vysokých teplotách. Na identifikáciu takýchto modifikácií bol zvolený prístup vkladania známych stabilizujúcich slučiek do vlásenky v katalytickom jadre deoxyribozýmu. Učinok týchto sekvencií na teplotu topenia deoxyribozýmu bol sledovaný pomocou NMR spektroskopie. Boli odmerané teplotne závislé ^{1}H NMR spektrá týchto variantov a pozorovalo sa topenie ich štruktúry. Takisto bola študovaná aj schopnosť deoxyribozýmov udržať si svoju katalytickú aktivitu pri vyšších teplotách, vďaka čomu boli získané ďalšie informácie ohľadom termostabilizujúcich sekvenčných prvkov.

Kľúčové slová: deoxyribozým, katalytická DNA, termostabilita, topenie, nukleárna magnetická rezonancia

Abstract

Like proteins, DNA molecules can catalyze chemical reactions. Inspired by the potential uses of catalytic DNA, an artificial evolution was used in Dr. Curtis' lab to identify a deoxyribozyme that catalyzes a chemiluminescent reaction. Our current goal is to obtain a structure of this deoxyribozyme at atomic-level resolution.

In the case of both protein and nucleic acid enzymes, high-resolution structural characterization is often facilitated by mutations that increase stability at high temperatures. A rational design approach was chosen to identify such modifications; a series of previously described stabilizing loops were inserted into a hairpin in the catalytic core of the deoxyribozyme, and NMR spectroscopy was used to characterize the effect of these loop sequences on the melting temperature of the deoxyribozyme.

Variable-temperature (VT) ^{1}H-NMR spectra of these variants were measured, and the melting transition of their structure was investigated. In addition, the ability of deoxyribozymes to retain their catalytic activity at higher temperatures was studied and provided further insight into the sequence elements that demonstrate a thermostabilizing effect.

Key words: deoxyribozyme, catalytic DNA, thermostability, melting, nuclear magnetic resonance

Table of contents

Abbreviations

A	adenine
AMP	adenosine monophosphate
ATP	adenosine triphosphate
C	cytosine
CI	confidence interval
DNA	deoxyribonucleic acid
et al.	*et alia* (and others)
FID	free induction decay
G	guanine
HPLC	high-performance liquid chromatography
mRNA	messenger RNA
NC	negative control
NMR	nuclear magnetic resonance
nt	nucleotide
PAGE	polyacrylamide gel electrophoresis
PC	positive control
PCR	polymerase chain reaction
RF	radiofrequency
RNA	ribonucleic acid
RNase	ribonuclease
rRNA	ribosomal RNA
RT	room temperature
SD	standard deviation
T	thymine
tRNA	transfer RNA
VT	variable-temperature

1 Theoretical overview

1.1 Nucleic acids with a catalytic function

1.1.1 In the beginning there was RNA

The view that the purpose of nucleic acids is solely to store and transfer genetic information in living cells was changed in the 1980s when the first self-splicing RNA molecule was discovered in the ciliate *Tetrahymena* [1]. This discovery prompted the search for other nucleic acids capable of catalyzing biochemical reactions, leading to the identification of various naturally occurring, as well as *in vitro* engineered RNA molecules with catalytic activity, termed RNA enzymes, or ribozymes. These molecules were found in the genomes of many different species, catalyzing reactions that had been previously observed only among protein enzymes.

The *Tetrahymena* preribosomal RNA was later revealed to be a member of the group I intron family, which now consists of more than 1,500 RNA sequences that interrupt coding regions of genes for rRNA, tRNA, and mRNA and catalyze their own excision in several species, including bacteria and bacteriophages, fungal mitochondria, chloroplasts of algae, eukaryotic viruses and microorganisms [2]. Following the report of the *Tetrahymena* intron was a discovery of an RNA-protein complex named RNase P [3]. The RNA component was shown to be responsible for site-specific hydrolysis of precursor tRNA substrates [4]. Homologs of this ribonucleoprotein were found in the genome of all free-living organisms, and the assessment of its evolutionary tree revealed that it might be a remnant of the ancient world [5]. One of the first discovered ribozymes is also a small and thoroughly studied Hammerhead ribozyme, which catalyzes a site-specific RNA self-cleavage. Even though it was initially found in plant satellite RNA viruses [6], the Hammerhead ribozyme was later observed across all life kingdoms [7, 8].

As the first catalytic RNAs emerged, a newly formed hypothesis named the "RNA World" followed [9–11]. This theory proposed that self-replicating RNA (or something functionally similar) preceded DNA and proteins during evolution and served both as storage of genetic information and catalysts of chemical reactions. This theory raises many questions and faces a number of problems, but although incomplete, for now, it remains the most accepted backstory to the origin of life [12, 13].

1.1.2 DNA catalysts

Catalytically active DNA molecules (DNA enzymes, deoxyribozymes) are, unlike their RNA counterparts, not known to be found in nature. This is because biological DNA forms structurally regular double-strands that cause limited spatial flexibility when it comes to forming any tertiary structures that would bind and convert substrates, and therefore there is little opportunity for such catalysts to develop *in vivo* [14]. However, numerous single-stranded deoxyribozymes and ribozymes were identified using the *in vitro* selection (also known as artificial evolution) method [15], a process of subjecting a large pool of nucleic acids with random sequence to a selection pressure until a small number of functional sequences with desired activity is obtained.

In 1994, the first report of deoxyribozyme discovered by this directed evolution technique revealed an oligonucleotide capable of accelerating cleavage of RNA substrates by ~100,000 fold compared to the non-enzymatic cleavage rate [16]. The subsequent research showed that many deoxyribozymes could be readily prepared using *in vitro* selection, catalyzing a broad range of reactions with chemical efficiency similar to those of ribozymes [17].

The difference between DNA and RNA catalysts lies in their atomic composition. DNA lacks a hydroxyl group on the 2' carbon atom of the sugar ring found in RNA, making it less susceptible to hydrolysis and therefore increasing its enzymatic stability and durability. The ability to catalyze a chemical reaction depends on the formation of diverse structural elements such as helices, pseudoknots, triplexes, and guanine quadruplexes [18].

Catalytic nucleic acids interact with potential substrates by hydrogen bonding, π-stacking, and metal-ion coordination. Despite being less chemically diverse than protein enzymes, they show a large potential to play complex functional roles in biological and chemical systems [19], mainly because of their thermal stability, ease and low cost of their synthesis, and the possibility of developing entirely new active sequences by *in vitro* evolution.

1.1.3 Applications of deoxyribozymes

Deoxyribozymes have found application in many fields, including biotechnology, biosensing, anti-virus, and materials science. Current research interest is particularly focused on the prospect of deoxyribozymes as biological sensing agents. This function is based on the capability of deoxyribozymes to target their specific cofactor and, in its presence, cause a readily observable change in the reaction system, which can be something as simple as the

production of fluorescence or a colorimetric signal. One promising result of this method was the construction of RNA-cleaving lead-specific deoxyribozyme [20, 21], which has the potential to serve as a highly sensitive and cost-effective part of a sensor system for the detection of this toxic heavy metal [22]. Other metal ions detectable by reported deoxyribozymes include Zn^{2+} [23, 24], Mg^{2+} [25, 26], Ca^{2+} [26], Hg^{2+} [27, 28], Cd^{2+} [29], Ln^{3+} [30], Ag^{+} [31] and others, and successful results in detection of Na^{+} and UO_2^{2+} ions were also acquired *in vivo* [32, 33].

The application of DNA biosensors can also be extended beyond simple metal detection. Allosteric DNA enzymes, whose activity is regulated by the binding of the target to the aptamer domain, were designed to signal the presence of specific RNA and DNA oligonucleotides [34–36], and others have been shown to detect molecules such as ATP [37, 38]. These examples are only a few of many and showcase a prospect of deoxyribozymes playing a key role in the biochemical systems of the future.

1.2 Supernova – a light-producing deoxyribozyme

Motivated by the idea of creating a chemiluminescent sensor, a light-producing deoxyribozyme called Supernova was developed in Dr. Curtis' group [39]. This catalytic DNA transfers a phosphate group from a 1,2-dioxetane substrate called CDP-Star to its 5' hydroxyl group, which causes decomposition of the substrate followed by a flash of blue light (Figure 1).

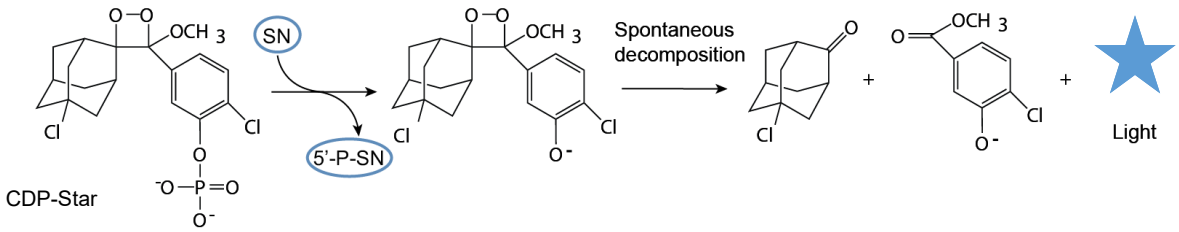

Figure 1: The mechanism of the chemiluminescent reaction of Supernova with CDP-Star. *SN*, Supernova; *5'-P-SN*, 5'-phosphorylated Supernova molecule. Adapted from [39].

1.2.1 Structure

The unusual secondary structure model of Supernova (Figure 2) was proposed based on analysis of correlated mutational changes in the deoxyribozyme by comparative sequence analysis [39]. The purine-rich active core was shown to contain 46 nucleotides, forming a purine-motif triple helix capped by an 11-nucleotide loop and another 11-nucleotide motif

adjacent to its 5′ end. Of these 46 nucleotides, 38 were conserved, and the remaining eight are linkers connecting important parts of the structure.

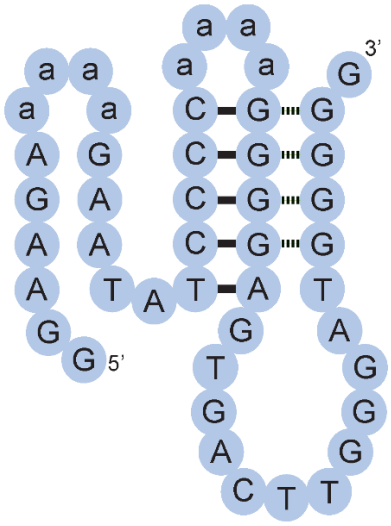

Figure 2: The proposed secondary structure of Supernova. Conserved positions are represented by capital letters and hydrogen bonds by solid and dashed horizontal lines. Adapted from [39].

1.2.2 Activity

The commercially available substrate of Supernova is called CDP-Star. It contains 1,2-dioxetane linked to a reactive group, which is in its inactive state blocked by a phosphate. It is usually used for immunodetection of alkaline phosphatase, which hydrolyses phosphate monoesters. The loss of the protecting phosphate group triggers the formation of an unstable intermediate, which decomposes into electronically excited carbonyl products followed by a flash of blue light (Figure 1) [40]. Supernova belongs to the group of self-phosphorylating deoxyribozymes [41] and triggers the chemiluminescent reaction by transferring a phosphate from the substrate to its 5' hydroxyl group.

Recent experiments have shown that an essential requirement for catalytic activity is the presence of divalent zinc. Possible functions of zinc cations include stabilizing the reaction's transition state and promoting the folding of Supernova into its active conformation. Intriguingly, alkaline phosphatase (which catalyzes a similar reaction) contains two zinc ions in its active site [42], raising the possibility that the reaction mechanisms of Supernova and alkaline phosphatase are similar.

At the optimal zinc concentration of 0.65 mM, pH 7.4, and in the presence of molecular

crowding agents, Supernova generates light corresponding to a rate enhancement of ~8,600 fold [39]. The production of light can be directly measured on the plate reader as a function of time (Figure 3).

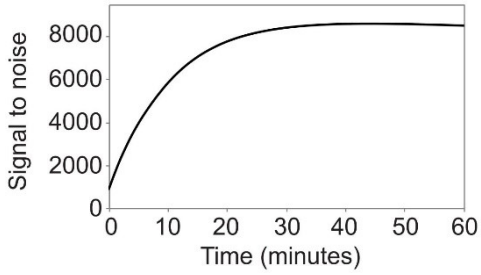

Figure 3: Light production catalyzed by Supernova under optimal conditions (amount of light generated in the presence of Supernova divided by the amount generated in the absence of Supernova) as a function of time.

The first step in the reaction is the transfer of a phosphate group from the substrate to Supernova. This can be studied independently of the substrate decomposition using a ligation assay (Figure 4 A). In this approach, after incubating Supernova with CDP-Star, T4 DNA ligase can join reacted molecules (which contain a 5' phosphate rather than a 5' hydroxyl group) with an acceptor oligonucleotide (Figure 5). The longer product can be subsequently detected on a gel (Figure 4 B), allowing for the detection of Supernova molecules that transferred a phosphate from CDP-Star to their 5'-terminus during the incubation.

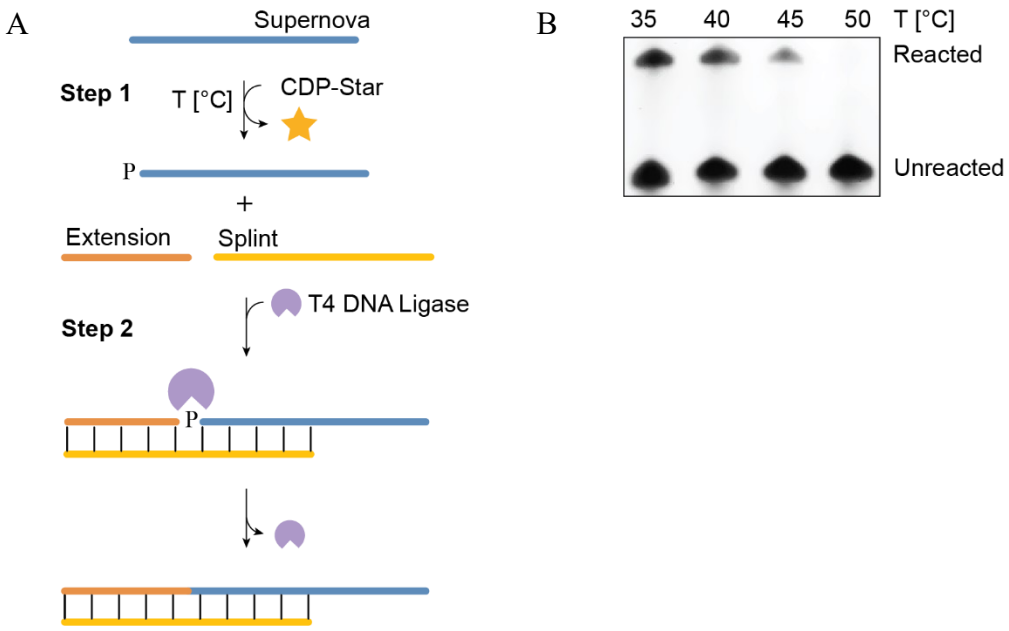

Figure 4: (A) Design of the phosphorylation analysis experiment consisting of two steps. Step 1 is the chemiluminescent reaction of Supernova with CDP-Star at various temperatures, and Step 2 is the ligation reaction of phosphorylated Supernova with Extension oligonucleotide, using Splint oligonucleotide as a template (Table 2).

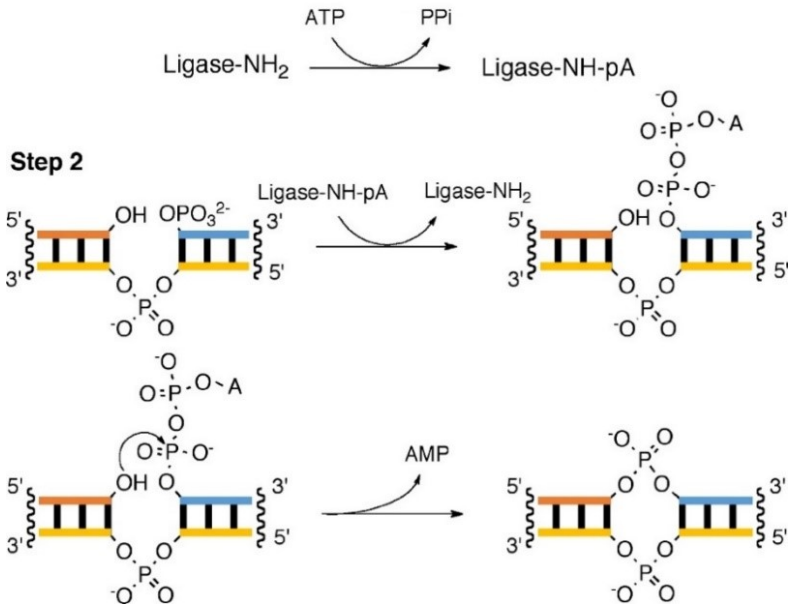

Figure 5: Mechanism of ligation of 5'-phosphate and 3'-hydroxyl termini in DNA catalyzed by T4 DNA ligase. *A*, 5'-adenosine. Reproduced from [43].

1.2.3 Applications

An allosterically regulated variant of Supernova was also created, demonstrating its potential as a biosensing agent [39]. This engineered version is only catalytically active after binding to an oligonucleotide that is complementary to its 3' end, showing that the presence of a ligand can be linked to light production. This can be utilized not only as a signaling component of biosensors, but also in DNA-based logic gates of molecular computers.

1.3 Designing a thermostable deoxyribozyme

The mechanisms of protein folding and enzymatic activity at high temperatures have been extensively studied [44]. Thermophilic proteins were established to introduce individual amino acid changes that additively contribute to the structure stabilization by either creating or optimizing electrostatic interactions. Protein enzymes were also documented to increase their affinity for metal ions or introduce new metal-binding sites to enhance their thermostability [45]. These studies resulted in the development of various methods for designing thermophilic protein variants, such as thermostable polymerases, that found their significant use in PCR.

In contrast, catalytic nucleic acids have received less attention regarding their thermostability, and comparatively less is known about the thermodynamics of these large catalytic structures [37]. As opposed to proteins that use hydrophobic interactions for their folding, nucleic acids have a hydrophilic core and therefore employ different strategies for maintaining their secondary and tertiary structures. Nucleic acids consist of hydrophobic bases and negatively charged phosphate backbone and thus are stabilized primarily by forming hydrogen bonds between nucleotides and stacking interactions. Other stabilizing forces include electrostatic attraction between the negatively charged phosphates and positive charges, as well as the coordination interactions between DNA bases and metal ions. Formation of higher-order structure is mediated by tertiary interactions such as base triples, GNRA tetraloops (N = any nucleotide, R = purine), A-minor motifs (the insertion of minor groove edges of adenines into the minor groove of neighboring helices), and binding sites for metal ions.

The thermodynamics of complex nucleic acid catalysts can be obtained from a comparison of the mesophilic and thermophilic homologs of the same nucleozymes, such as RNase P. A study has shown that the thermophilic RNA forms a less structured folding intermediate and demonstrates increased folding cooperativity that is achieved through stabilizing helices and shortening linkers between them [46]. Recent reports of the *in vitro*

development of a thermostable *Tetrahymena* ribozyme uncovered nine stabilizing mutations increasing the melting temperature of ribozyme for more than 10 °C [47]. Some of these mutations were shown to disrupt its secondary structure; however, their impact on tertiary interactions within the molecule had an overall substantial stabilizing effect. In both cases, a small number of context-dependent mutations was shown to provide significant stability. These studies have shed some light on the inner workings of these complex molecular tools and provided a background for the possible development of thermophilic variants of Supernova.

In this project, we have proposed two approaches to introduce stabilizing mutations for the light-producing deoxyribozyme – *in vitro* selection and rational design. The *in vitro* selection of Supernova variant active at high temperatures was performed; however, the aforementioned work is beyond the scope of this thesis.

The rational design focused on the hairpin tetraloop located in the triple helix of the Supernova structure (Figure 2). The sequence of this loop is not conserved, which indicates that the nucleotides in these positions are not part of the active site of the deoxyribozyme, and therefore non-essential for its catalytic activity. Therefore, the potential of this structural element playing a role in the overall thermostability of the molecule was considered and explored.

In a study by M. Nakano *et al*. (2002), the *in vitro* selection was used to identify loop sequences that stabilize the structure of a model DNA hairpin [48]. We hypothesized that these sequences would also stabilize the triple helix in Supernova, and to test that idea, a few loops with established thermodynamic parameters were inserted into the catalytic core of the chemiluminescent deoxyribozyme. With respect to the closing base pair CG, which was also shown to be an essential factor for the structural stability [48], the first tetraloop sequence, d(GCTA), was chosen based on its most favorable thermodynamic parameters, with a melting temperature of 71.8 ± 0.2 °C. To verify the practicality of this approach, the most destabilizing sequence d(GCTT) with a significantly lower melting temperature of 54.7 ± 1.0 °C was also investigated. Finally, d(GCA) sequence, known to be the most stabilizing triloop with a melting temperature of 76.1 ± 0.7 °C, was chosen to examine the effect of one missing nucleotide on the stability and activity of Supernova.

1.4 NMR spectroscopy

Nuclear magnetic resonance spectroscopy is a powerful method that observes and measures interaction of nuclear spins when placed in an external magnetic field. Its sensitivity to the chemical environment of the nucleus makes it suitable for studying the structural properties of DNA.

1.4.1 Nuclear spin

Nuclear spin angular momentum is a vector represented by the symbol $\vec{I}$ and its projection on the arbitrarily chosen z-axis (without a loss of generality) can be written as

$$I_z = m\hbar, \tag{1}$$

where $\hbar$ is a reduced Planck's constant and m is the magnetic orbital quantum number, that can be defined as

$$m = -\mathrm{I}, -\mathrm{I}+1, \ldots, \mathrm{I}-1, \mathrm{I}; \tag{2}$$

where I is a spin quantum number with values greater than zero, increasing by multiples of $\frac{1}{2}$. From the equations (1) and (2), it follows that the number of possible projections of nuclear spin on the z-axis I_z is $(2\mathrm{I}+1)$.

NMR can detect nuclei with spin other than zero, and the most routinely used isotopes have spin of $\frac{1}{2}$, such as ^{1}H with the natural abundance of 99.984%, and ^{13}C with the natural abundance of 1.108% [49].

The nucleus possesses a magnetic moment $\vec{\mu}$, which projection on the z-axis can be written as

$$\mu_z = \gamma I_z, \tag{3}$$

where γ represents the gyromagnetic ratio, a constant unique to each nucleus. The macroscopic effect of magnetic moments is represented by the bulk magnetization vector $\vec{M}$, which is obtained as the sum of magnetic moments of every nucleus in the sample:

$$M_z = \sum_i \mu_{z,i}. \tag{4}$$

1.4.2 Nuclear spin in a constant magnetic field

Placing a nucleus into a static magnetic field $\vec{B}$ causes precession of nuclear magnetization around the vector of the magnetic field. Frequency of this precession ω_L is called

Larmor frequency, and if $\vec{B} = (0; 0; B_0)$, it can be expressed (in rad s^{-1}) as

$$\omega_L = -\gamma B_0. \tag{5}$$

The application of a magnetic field to a nucleus causes a split of the $(2I + 1)$ degenerate nuclear energy levels into low and high energy states. The energy of nuclear magnetic momentum in a static magnetic field is dependent on its strength and can be expressed as

$$E = -\vec{\mu} \cdot \vec{B}. \tag{6}$$

Combining the equations (3) and (6) allows for calculating the difference between the energy states of the system:

$$\triangle E = \hbar \gamma B_0. \tag{7}$$

The distribution of spin population among energy levels is based on Boltzmann statistics. A transition between these energy levels occurs when the system is exposed to radiation with a frequency ω_0 called resonance frequency, which is equal to the Larmor frequency ω_L:

$$\omega_0 = \frac{\triangle E}{\hbar} = -\gamma B_0. \tag{8}$$

1.4.3 Interactions with an oscillating magnetic field

The nuclear spin system in a constant magnetic field can be manipulated by radiofrequency pulses that create a linearly polarized oscillating magnetic field $\vec{B}_{RF} = (2B_1 \cos \omega_{RF}\, t; 0; 0)$. The spin system, therefore, interacts with the time-varying magnetic field $\vec{B}$:

$$\vec{B}(t) = \vec{B}_0 + \vec{B}_{RF}(\mathrm{t}). \tag{9}$$

The motion of the net magnetization can be described by Bloch equations as a function of time accounting for relaxation:

$$\begin{aligned}
\frac{\mathrm{d}M_x(t)}{\mathrm{d}t} &= \gamma\big(M_y(t)B_z(t) - M_z(t)B_y(t)\big) - \frac{M_x(t)}{T_2} \\
\frac{\mathrm{d}M_y(t)}{\mathrm{d}t} &= \gamma\big(M_z(t)B_x(t) - M_x(t)B_z(t)\big) - \frac{M_y(t)}{T_2} \\
\frac{\mathrm{d}M_z(t)}{\mathrm{d}t} &= \gamma\big(M_x(t)B_y(t) - M_y(t)B_x(t)\big) - \frac{M_z(t) - M_0}{T_1},
\end{aligned} \tag{10}$$

where T_1 is the longitudinal relaxation time, T_2 is the transversal relaxation time, and M_0 is a magnetization of the system in equilibrium.

The static magnetic field B_0 causes a Larmor precession of magnetization at a frequency in hundreds of MHz and is much stronger than B_1created by radiofrequency pulses. In the laboratory frame of reference, the B_1 magnetic field is observed as a vector oscillating along the x-axis and can be thought of as two components rotating clockwise and anticlockwise around the z-axis at $+$ and $-\omega_{RF}$, respectively. To allow for a simpler observation of the effect of B_1 on the magnetization, it is convenient to transfer to a rotating frame of reference that rotates around the z-axis at a frequency of $+\omega_{RF}$. The B_1 vector rotating clockwise will in rotating frame of reference appear stationary; meanwhile, the other vector will rotate around the z-axis at $-2\omega_{RF}$, which is far from the resonance frequency of the spins and thus can be neglected. In case of the frequency of the oscillating magnetic field being equal to the Larmor frequency of nuclei, $\omega_{RF} = -\gamma B_0$, the effect of external magnetic field B_0 is negated, and therefore the effective magnetic field will have only one static component, $\vec{B}_{ROT}(t) = (B_1; 0; 0)$.

This transformation can be obtained by adjusting Bloch equations by addition of inertial force:

$$\left[\frac{\mathrm{d}\vec{M}(t)}{\mathrm{d}t}\right]_{ROT} = \left[\frac{\mathrm{d}\vec{M}(t)}{\mathrm{d}t}\right]_{LAB} + \vec{M}(t) \times \vec{\omega}, \tag{11}$$

which can be further modified to:

$$\frac{\mathrm{d}\vec{M}(t)}{\mathrm{d}t} = \vec{M}(t) \times \left[\vec{B}(t) + \frac{\vec{\omega}}{\gamma}\right], \tag{12}$$

where $\vec{\omega} = (0,0,\omega_{RF})$.

This allows for describing magnetization perpendicular to B_1, precessing around the magnetic field B_1 with so-called nutation frequency ω_1 in dozens of kHz:

$$\omega_1 = \gamma B_1. \tag{13}$$

An angle of rotation φ can be then calculated as

$$\varphi = \gamma B_1 \tau, \tag{14}$$

where τ is the length of the RF pulse.

1.4.4 Free induction decay

In a system placed in the external magnetic field, an equilibrium magnetization is parallel with the external magnetic field vector. Once a radiofrequency pulse is applied, the magnetization will tilt into the xy-plane. The development of magnetization vectors after a

single pulse can then be described in the rotating frame of reference by Bloch equations (10).

As an example, a system in the magnetic field $\vec{B} = (0; 0; B_0)$ and an application of 90°(y) RF pulse ($\varphi = \pi/2$) will be used. Supposing the RF pulse is significantly shorter than T_1 and T_2, magnetization right after the pulse is

$$M(t_0) = (M_0 \sin\varphi\,; 0; M_0 \cos\varphi) = (M_0; 0; 0). \tag{15}$$

The evolution of magnetization over time can be expressed as

$$\begin{aligned} M_x(t) &= M_0 \cos(\Omega t)\, e^{(-t/T_2)} \\ M_y(t) &= M_0 \sin(\Omega t)\, e^{(-t/T_2)}, \end{aligned} \tag{16}$$

where Ω is the offset and can be calculated as $\Omega = \omega_0 - \omega_{RF}$.

The precession of magnetization in the xy-plane induces a corresponding oscillating voltage (signal S) in a receiver coil. The voltage is digitally processed and obtained as a complex time-domain signal $S(t)$, known as FID (free induction decay) [50]:

$$S(t) = S_x + iS_y = S_0 e^{(i\Omega t)} e^{(-t/T_2)}, \tag{17}$$

where signals S_x and S_y are proportional to M_x and M_y with the constant of S_0.

Fourier transform performed of this complex time-domain signal gives a spectrum as a function of frequency $S(\omega)$ consisting of real and complex components:

$$S(\omega) = A(\omega) - iD(\omega), \tag{18}$$

where $A(\omega)$ and $D(\omega)$ are known as the absorption and dispersion mode Lorentzian lineshape functions, respectively (Figure 6). Without loss of generality, the scaling constant S_0 can be neglected and then the absorption and dispersion mode lineshapes can be defined as

$$A(\omega) = \frac{1/T_2}{(1/T_2)^2 + (\omega - \Omega)^2} \qquad D(\omega) = \frac{-(\omega - \Omega)}{(1/T_2)^2 + (\omega - \Omega)^2}. \tag{19}$$

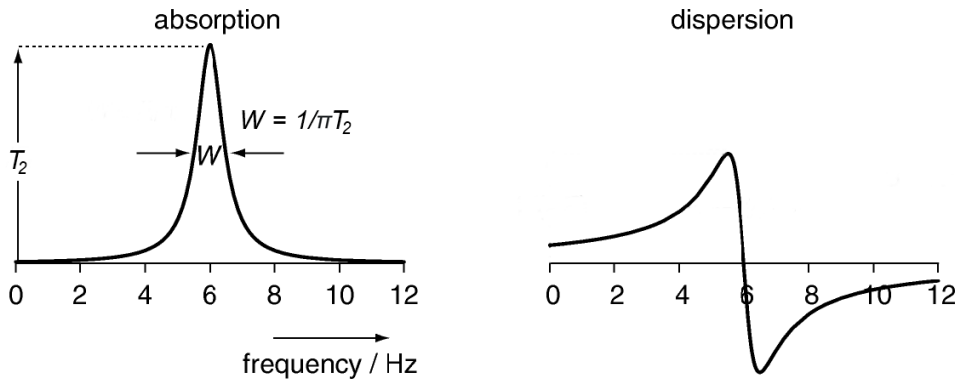

Figure 6: An absorption and dispersion Lorentzian lineshapes given in the equations (19). Adapted from [50].

Assuming $\omega - \Omega = 0$, it is easy to see from the equation (19) that the height of the peak is equal to T_2, and the width at half height of the absorption mode lineshape is $1/\pi T_2$. This means that with increase of relaxation rate (decreasing T_2) the signal is widening and its maximal intensity decreasing, and vice versa.

1.4.5 Chemical shift

From equation (8) it follows, that the magnitude of resonant frequency is directly proportional to the magnetic field B_0. In reality, the intensity of the magnetic field in the location of the nucleus B_{loc} can slightly differ from the external field B_0. This is caused by the electrons surrounding nuclei creating an opposing magnetic field to B_0, known as chemical shielding:

$$B_{loc} = B_0(1 - \sigma), \tag{20}$$

where σ is the chemical shielding constant. Change in the magnetic field at a nucleus location depends on the distribution of electrons in its vicinity; therefore, the resonant frequency differs for nuclei with different chemical environments.

Since resonant frequencies in NMR are proportional to the applied magnetic field, they can be normalized with respect to the reference signal and displayed as chemical shift δ in ppm (parts per million):

$$\delta = \frac{\omega - \omega_{ref}}{\omega_{ref}} \cdot 10^6, \tag{21}$$

where ω is a resonant frequency of the given nucleus and ω_{ref} is a reference resonant frequency of standard. A standard routinely used for measuring hydrogen spectra of samples in water buffer is, for example, 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS). For historical

reasons, the chemical shift in the NMR spectra is shown increasing from right to left.

1.4.6 Dipole-dipole interactions

Nuclei can interact with each other either directly, through space, or indirectly, in which case the interactions are mediated by electrons. In an isotropic liquid, the random motion of molecules averages the direct dipolar coupling, causing the dipole-dipole interactions to be directly unobservable. However, this leads to a process of so-called cross-relaxation and consequently causes the Nuclear Overhauser Effect (NOE), which can be used under appropriate conditions to enhance the sensitivity of rare spin observations.

Indirect dipole-dipole coupling, or J-coupling, occurs when a magnetic field produced by a spinning nucleus causes changes in the electron distribution of its bonds. These, in turn, affect the electron distribution of neighboring bonds, which affects the adjacent nuclei. This information can be transmitted by several chemical bonds and results in the often-complicated signal splitting seen in NMR spectra.

A nucleus with spin $\frac{1}{2}$ can be projected onto the z-axis as $+\frac{1}{2}\hbar$ and $-\frac{1}{2}\hbar$, and the signal of another nucleus interacting with this nucleus by J-coupling would be split into two resonance lines, the so-called doublet. In DNA, this is the case of hydrogen H6 in cytosine, which interacts with hydrogen H5 (Figure 9, p. 23) through three-bond coupling. The strength of a coupling constant depends on the distance between studied atoms; therefore, a much weaker, four-bond coupling between H6 and methyl hydrogens in thymine splits the signal into a quadruplet. This is because the three hydrogens in the methyl group have four possible spin projections on the z-axis, since combining three spins $\frac{1}{2}$ results in four magnetic quantum numbers $+\frac{3}{2}, +\frac{1}{2}, -\frac{1}{2}$ and $-\frac{3}{2}$ [51].

1.4.7 Chemical exchange

Chemical exchange is a process during which the chemical environment of a given nucleus changes, leading to a difference in its resonance frequency and, therefore, its chemical shift. In DNA, it is the case of formation and disruption of secondary structures. The effect of chemical exchange on NMR spectra depends on the ratio between the chemical exchange rate and the difference in frequency of the two nuclear states. If the exchange rate is significantly lower than the difference in frequency, two distinctive signals in the spectrum will be

observable. However, in the case of rapid two-state chemical exchange, only one signal appears at a position corresponding to the average frequency of both states of a given nucleus (Figure 7).

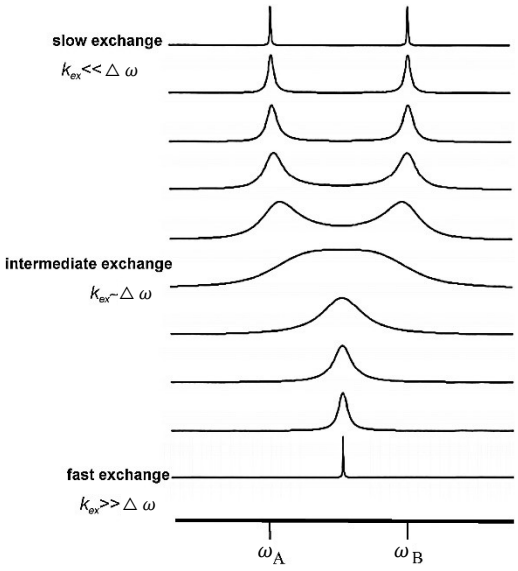

Figure 7: Chemical exchange demonstrated at different $k_{ex}/\triangle \omega$. k_{ex}, chemical exchange rate constant; ω_A, ω_B, resonance frequencies of nuclei A and B; $\triangle \omega = \omega_B - \omega_A$. Adapted from [52].

1.4.8 Pulse experiments

One of the pulse sequences used to measure one-dimensional DNA spectra is the *zgesgp* pulse program (Figure 8). This sequence contains a spin-echo segment [50] and elements for suppressing the solvent signal. It also utilizes pulse field gradients, which are used to manipulate the signal phase by spatially changing the strength of the static magnetic field for a short period of time. First, $\frac{\pi}{2}$ pulse is applied to tilt the magnetization into the transverse plane, and then a gradient pulse G_1 causes the spins to become out-of-phase. The following part of the sequence consists of a long selective π pulse at the frequency of solvent hydrogens, which flips the magnetization of the nuclei with a chemical shift corresponding to the solvent, while the magnetization of the other nuclei remains unchanged. Afterward, a non-selective π pulse flips the magnetization of all nuclei in a sample. As a result, the magnetization of the solvent, unlike other nuclei, returns to the direction in which it was immediately after the gradient pulse. Therefore, the second gradient pulse G_1 further fans out the phases of the solvent magnetization,

while the phases of other nuclei are refocused and thus return to the state before the first gradient pulse. A similar process is repeated in the second part, which contains the G_2 gradient pulses. As a result, the solvent signal is significantly suppressed, which allows for the measurement of ^{1}H spectra using solvents such as water.

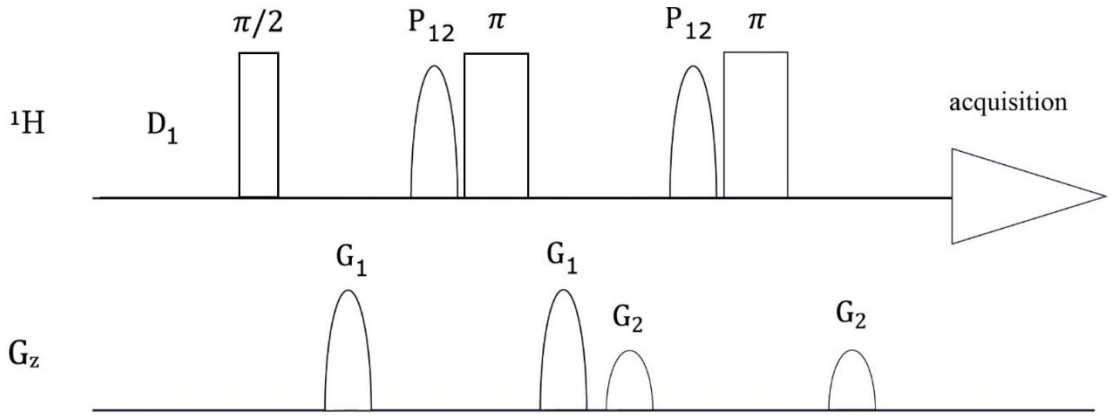

Figure 8: Pulse sequence *zgesgp* for measuring one-dimensional spectra utilizing solvent compression. *^1H*, frequency channel of the isotope; G_z, gradient channel applied in the z-axis direction; $\pi, \pi/2$, non-selective pulses; $P_{12,}$ selective π pulses at the solvent frequency. Reproduced from [53].

1.4.9 NMR spectroscopy of DNA melting

Nucleotides forming DNA are composed of four nitrogen-containing nucleobases (Figure 9), a deoxyribose sugar (Figure 10), and a negatively charged phosphate group. Chemical shift ranges of hydrogens in ^{1}H NMR spectra of DNA are illustrated in Figure 11.

By measuring ^{1}H temperature-dependent NMR spectra, a process called DNA melting can be observed. Increasing temperature causes a progressive denaturation of DNA's secondary structure until it separates into its single-stranded form. This results in gradual changes in the chemical environment of protons that can be on an NMR spectrum followed as the change in their chemical shifts and signal intensity. It allows for the examination of chemical exchange and conformational rearrangements that happen during the decomposition of native DNA structure.

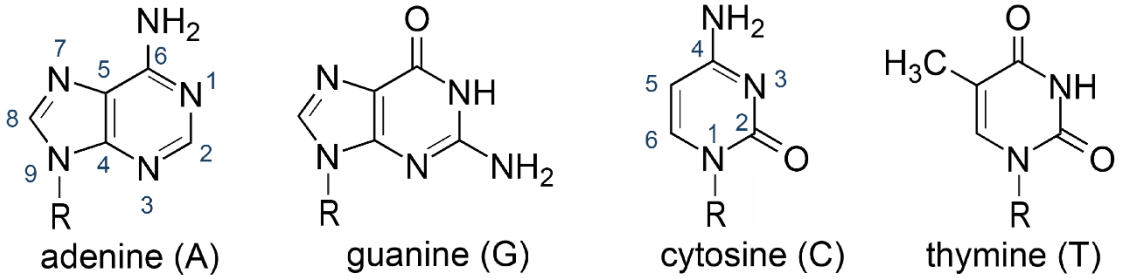

Figure 9: Chemical structures of the heterocyclic bases in DNA and standard numbering of the carbon atoms. *R*, sugar. Adapted from [54].

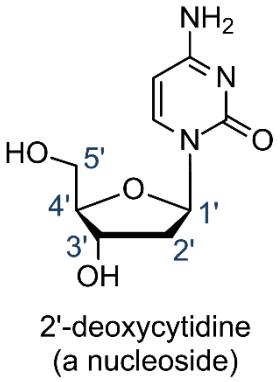

Figure 10: Chemical structure of the nucleoside 2'-deoxycytidine and standard numbering of the carbon atoms in deoxyribose. Reproduced from [54].

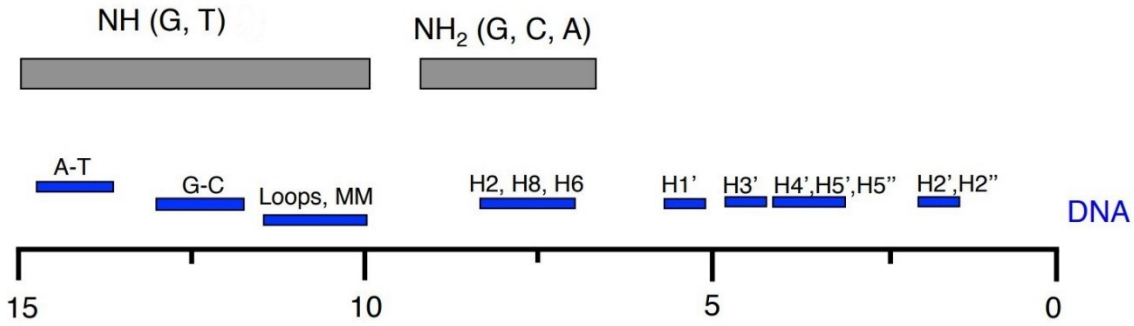

Figure 11: Chemical shift ranges of DNA protons on NMR spectrum. Hydrogens are labelled based on their position in Figure 9 and Figure 10. Reproduced from [55].

2 Objective

A light-producing deoxyribozyme named Supernova was engineered in Dr. Curtis' group, and this thesis is a part of a project designed to obtain its 3D structure at an atomic resolution. This work aims to identify thermo-stabilizing mutations in Supernova and to further characterize the temperature-induced changes in its structure by measuring and interpreting VT (variable-temperature) ^{1}H NMR spectra.

It has been proposed that thermo-stabilizing sequence mutations improve structural rigidity of proteins, as well as deoxyribozymes, which can facilitate their high-resolution characterization by limiting the conformational changes causing line-broadening in NMR spectra. Furthermore, an identification of such thermo-stable Supernova variants will allow for its possible use over a wider range of conditions and will provide an insight into general mechanisms of thermostability in deoxyribozymes.

3 Experimental part

3.1 Chemicals

DNA oligonucleotides	Sigma-Aldrich
Reagents	
CDP-Star substrate	Invitrogen
T4 DNA Ligase	Jena Bioscience
Standard Ligation Buffer (500 mM Tris-HCl pH 7.8 at 25 °C, 100 mM $MgCl_2$, 100 mM DTT, 10 mM ATP and 25 μg/ml BSA)	Jena Bioscience
T4 Polynucleotide Kinase	NEB
Other chemicals	
acrylamide	BIO-RAD
APS (ammonium persulfate)	BIO-RAD
boric acid	Sigma-Aldrich
bromophenol blue	Sigma-Aldrich
deuterium oxide	Eurisotop
DSS (sodium 2,2-dimethyl-2-silapentane-5-sulfonate)	Eurisotop
EDTA (ethylenediaminetetraacetic acid)	Invitrogen
ethanol	Penta
GelRed	Biotium
HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)	Sigma-Aldrich
potassium chloride	Sigma-Aldrich
potassium hydroxide	Sigma-Aldrich
Sigmacote® - a siliconizing reagent for glass	Sigma-Aldrich
sodium chloride	Sigma-Aldrich
TEMED (N,N,N′,N′-tetramethylethylenediamine)	Sigma-Aldrich
Tris (tris(hydroxymethyl)aminomethane)	Sigma-Aldrich
urea	Sigma-Aldrich
zinc chloride	Sigma-Aldrich

3.2 Instrumentation

Centrifuges: Eppendorf Centrifuge 5424, Germany
Hettich MIKRO 220R Centrifuge, Germany

Electrophoresis power supply: VWR Power Source 300V, Germany

Laser scanner: GE Healthcare Typhoon FLA 9500, USA

Microcentrifuge: VWR MiniStar silverline, Germany

Microplate reader: Tecan Spark Multimode Reader, Switzerland

NMR spectrometer: Bruker Avance III HD 850 MHz, USA

pH meter: VWR, pHenomenal pH 1000 L, Germany

Precision balances: VWR SE 422, Germany

Shaker: Benchmark Scientific Orbi-Shaker, USA

Spectrophotometer: Thermo Scientific NanoDrop 1000, USA

Thermal cycler: Bio-Rad T100 Thermal Cycler, USA

Vortex mixer: VWR VM-3000, USA

Other equipment

Columns: Sigma-Aldrich SigmaSpin Sequencing Reaction Clean-Up Columns, USA

Microplates: Corning 96-well Half Area White Flat Bottom NBS Microplate, USA

NMR tubes: Bruker Match tubes, Germany

Vertical gel system: Analytik Jena Biometra Model V15-17, Germany

3.3 Methods

3.3.1 General

Tested oligonucleotides were synthesized and purified by HPLC by Sigma-Aldrich and purchased as 200 μM water solutions. The sequences of Supernova variants with different loops (underlined) are shown in Table 1.

Table 1: Sequences of studied samples

Sample	Sequence (5' to 3')	Length
SN(AAAA)	GGAAGAAAAAGAATATCCCC**AAAA**GGGGAGTGACTTGGGATGGGGG	46 nt
SN(GCTA)	GGAAGAAAAAGAATATCCCC**GCTA**GGGGAGTGACTTGGGATGGGGG	46 nt
SN(GCTT)	GGAAGAAAAAGAATATCCCC**GCTT**GGGGAGTGACTTGGGATGGGGG	46 nt
SN(GCA)	GGAAGAAAAAGAATATCCCC**GCA**GGGGAGTGACTTGGGATGGGGG	45 nt

The optimized 5x buffer used for oligonucleotide folding (folding buffer) contained 100 mM KCl, 3.25 mM $ZnCl_2$ and 100 mM HEPES (pH = 7.4, adjusted by KOH).

3.3.2 NMR spectroscopy

Sample preparation

Supernova variants were diluted in water, heated at 65 °C for 2 minutes, and cooled at RT for 10 minutes. 5x folding buffer and D_2O were then added. Final conditions were 30 μM DNA, 1x folding buffer and 10 % (v/v) D_2O in a volume of 160 μl. Lastly, 0.5 μl of DSS was added as an internal NMR standard and samples were transferred to NMR tubes with a diameter of 3 mm.

NMR measurements

All spectra were measured using Bruker Avance III HD spectrometer with a magnetic field of 18.8 T, which corresponds to a hydrogen ^1H frequency of 850 MHz. Measurements were done using *zgesgp* pulse sequence (Figure 8). ^1H spectra of all samples were measured in the temperature range of 288 - 348 K in steps of 2 K, which represents 31 spectra for a sample. A 5-minute delay was inserted between the measurements to achieve thermal equilibrium of the samples. NMR measurements were done by Pavel Srb from the Structure biology group at IOCB, Prague.

Processing of spectra

Raw data were processed with the NMR software Topspin 3.2. First, the FID (time domain signal) was multiplied by squared sine-bell multiplication according to the function:

$$\mathrm{QSIN(t)} = sin\left\{\left(\pi - \frac{\pi}{\mathrm{SSB}}\right)(\mathrm{t/AQ}) + \frac{\pi}{\mathrm{SSB}}\right\}^2, \text{ where } 0 < \mathrm{t} < \mathrm{AQ} \text{ and } \mathrm{SSB} = 2 \quad (22)$$

AQ is an acquisition status parameter and SSB is a processing parameter with a value for a pure cosine function. This process resulted in improvement of signal-to-noise ratio. Afterwards, Fourier transformation was executed to obtain a frequency-domain spectrum. Finally, zero- and first-order phase corrections were performed manually, as well as adjustment of the baseline.

3.3.3 Light-Production Analysis

Chemiluminescent reaction

Supernova variants diluted in water were heated at 65 °C for 2 minutes and cooled at RT for 10 minutes. 5x folding buffer was added and samples were transferred to a white half-area 96-well plate. CDP-Star was added and the measurement of chemiluminescence using the Tecan Spark plate reader was immediately started. Light production was measured for 1 hour at 27 °C with an initial 10 s orbital shaking and luminescence attenuation with a settle time of 100 ms and 1000 ms integration.

The final conditions used for measurements were 1 µM oligonucleotide, 1x folding buffer and 100 µM CDP-Star in the total volume of 100 µl. Each measurement included a negative control (NC) consisting of 1x folding buffer and 100 µM CDP-Star (but no Supernova), and a sample consisting of 1x folding buffer.

Data processing

The amount of light detected in the 1x buffer was considered background and its values were subtracted from the measurements during data processing. Rate enhancement was calculated as the fold enrichment of the cumulative luminescence counts (relative light units) of the sample relative to the negative control.

3.3.4 Phosphorylation Analysis

Experimental design

Directly measuring light-production at high temperatures was not possible due to the

limitations of available instruments; therefore, a phosphate transfer method was used to evaluate Supernova's activity (Figure 4 A).

Chemiluminescent reaction (phosphate transfer) at various temperatures

1 µg of each Supernova variant was diluted in water, heated at 65 °C for 2 minutes, and cooled at RT for 10 minutes. Afterwards, 5x folding buffer was added and samples were incubated for 15 minutes in a thermal cycler at temperatures ranging from 15 °C to 70 °C. CDP-Star was then added to initiate the reaction. Final conditions were 1 µM Supernova, 1x folding buffer and 100 µM substrate in a total volume of 70 µl. After a 1-hour incubation in a thermal cycler at the specified temperature, the reaction was stopped with 2.8 µL of 500 mM EDTA (pH 7.4). Negative control reactions contained 1 µM oligonucleotide and 1x folding buffer (but no CDP-Star) in 70 µl.

Ethanol precipitation

Reactions were cleaned up using SigmaSpin Clean-Up Columns and ethanol precipitated afterwards. Precipitation was performed by first adding 6 µL of 4 M NaCl and then 200 µl of ethanol to each sample, which amounted to ca. 0.3 M NaCl and 70% (v/v) of EtOH in a total volume of ~ 285 µl. The suspension was then kept in a freezer at -20 °C for 45 minutes and centrifuged at 14 000 RPM at 19 °C for 40 minutes. The supernatant was then removed, and the pellet was left to dry at RT in an open tube for 30 minutes. The pellet was resuspended in water and the concentration of ssDNA was determined using a NanoDrop spectrophotometer.

Phosphorylation of the positive control

A positive control for ligation (PC) was prepared by diluting 10 µg of SN(AAAA) in water, heating the solution at 65 °C for 2 minutes, and cooling it at RT for 10 minutes. 10x T4 DNA Ligase buffer (containing 10 mM ATP) and T4 Polynucleotide Kinase were added, and the reaction was incubated at 37 °C for 30 minutes. Afterwards, the kinase was heat-inactivated by incubating the sample at 65 °C for 20 minutes. 2.7 Richardson units (2.7 µl) of kinase were used for 10 µg of DNA in a total volume of 25 µl in 1x T4 DNA Ligase buffer.

Ligation reaction

A volume containing 500 ng of DNA from the phosphorylation reaction (including NC and PC) was mixed with water and two shorter oligonucleotides (Extension and Splint, Table 2), heated at 65 °C for 2 minutes, and cooled at RT for 10 minutes. 10x T4 DNA Ligase buffer

and T4 DNA Ligase were added to bring the volume to 14 µl and the reaction was incubated in a thermal cycler at 37 °C for 30 minutes. Final conditions were 2.5 µM of each oligonucleotide, 1x T4 DNA Ligase buffer, and 0.5 Weiss Units (0.2 µL)/µg DNA of T4 DNA Ligase. The reaction was stopped by the addition of 14 µl of 2x BPB. Afterwards, phosphorylated molecules that underwent ligation with the Extension oligonucleotide were separated from non-phosphorylated molecules by 6% denaturing PAGE.

Table 2: Oligonucleotides used for phosphorylation analysis

Sample	**Sequence (5' to 3')**[a]	**Length**
Extension	ACCGCTCAGGTGTAGTATCA	20 nt
Splint	TATTCTTTTTCTTCCTGATACTACACCTGAGCGGT	35 nt
SN(XXXX)[b]	GGAAGAAAAAGAATATCCCC**XXXX**GGGGAGTGACTTGGGATGGGGG	46 nt

[a] Complementary parts of the sequences are shown in the same color
[b] XXXX = AAAA, GCTA, GCTT, or GCA

PAGE

Parts of the gel electrophoresis apparatus (two glass plates, three 1.5 mm thick spacers, and a 20-well comb) were washed by ethanol, silanized using Sigmacote to prevent the gel from sticking to the glass and securely clamped together with binder clips. 6% gels were prepared by adding 6 ml of acrylamide (40 % acrylamide/bis-acrylamide, 37.5:1 solution), 4 ml of 10x TBE buffer (900 mM Tris base, 900 mM boric acid, 20 mM EDTA), and 18 g of urea to a 50 ml tube and the final volume was adjusted to 40 ml with water. The solution was then microwaved and vortexed until all urea dissolved. 110 µl of 10% (w/v) APS and 33 µl of TEMED were added to the mixture, vortexed and immediately poured between the two glass plates. After the gel polymerized (2 - 3 hours), binder clips and the bottom spacer were removed, and the gel was secured in the vertical gel system. The top and bottom parts of the gel were submerged in 1x TBE and the comb was carefully removed. The gel was loaded with 28 µl of sample per well, and was run at 300 V for 1 hour 15 minutes (until the BPB reached the bottom of the gel). The gel was then transferred to a plastic container, immersed in 50 ml of 1x GelRed and placed on a shaker for 20 minutes. After dying the samples, the gel was scanned using a Typhoon laser scanner.

PAGE analysis and data processing

The ligation yield was determined by gel analysis using ImageQuant TL software (GE

Healthcare LifeSciences). The background was subtracted using the "Rolling ball with a radius of 200" method. Bands in the gel corresponding to ligated and unligated Supernova molecules were quantified and the phosphorylation yield was defined as a percentage of ligated molecules normalized to the PC ($PC_{ligated}$ = 100%).

4 Results and discussion

4.1 Biophysical characterization of Supernova melting

We used NMR spectroscopy to probe the thermal stability of selected Supernova variants. Four distinct regions are observed in each ^{1}H NMR spectrum of studied variants (Figure 12). The thymine methyl signals can be observed between 0.5 ppm and 1.5 ppm. Downfield from this region are found intensive signals of the HEPES buffer, which is present in significant excess compared to DNA, and a signal with a negative phase at 4.8 ppm originating from solvent suppression. Signals in this area represent overlapped deoxyribose H3, H4, H5, and H5" signals located between 2 ppm and 5 ppm. In the region between 5.0 ppm and 6.5 ppm can be found H1' proton signals from the deoxyribose. These signals are subjected to complex J-coupling related splitting and overlaps, making them difficult to distinguish. The third region between 7.0 and 9.0 ppm contains signals of labile hydrogens in NH_2 groups of C, A, and G, as well as aromatic H6 and H8. Each nucleotide of the sequence yields at least one signal, and this region displays less overlapping than deoxyribose hydrogens since these signals are comparatively less often split by J-coupling. The downfield part of the spectra shows signals between 12.5 ppm and 13.5 ppm. These are from imino-hydrogens of guanine H1 and thymine H3, and are only detectable when the studied protons are stabilized by hydrogen bonding, e.g., in Watson-Crick pairing.

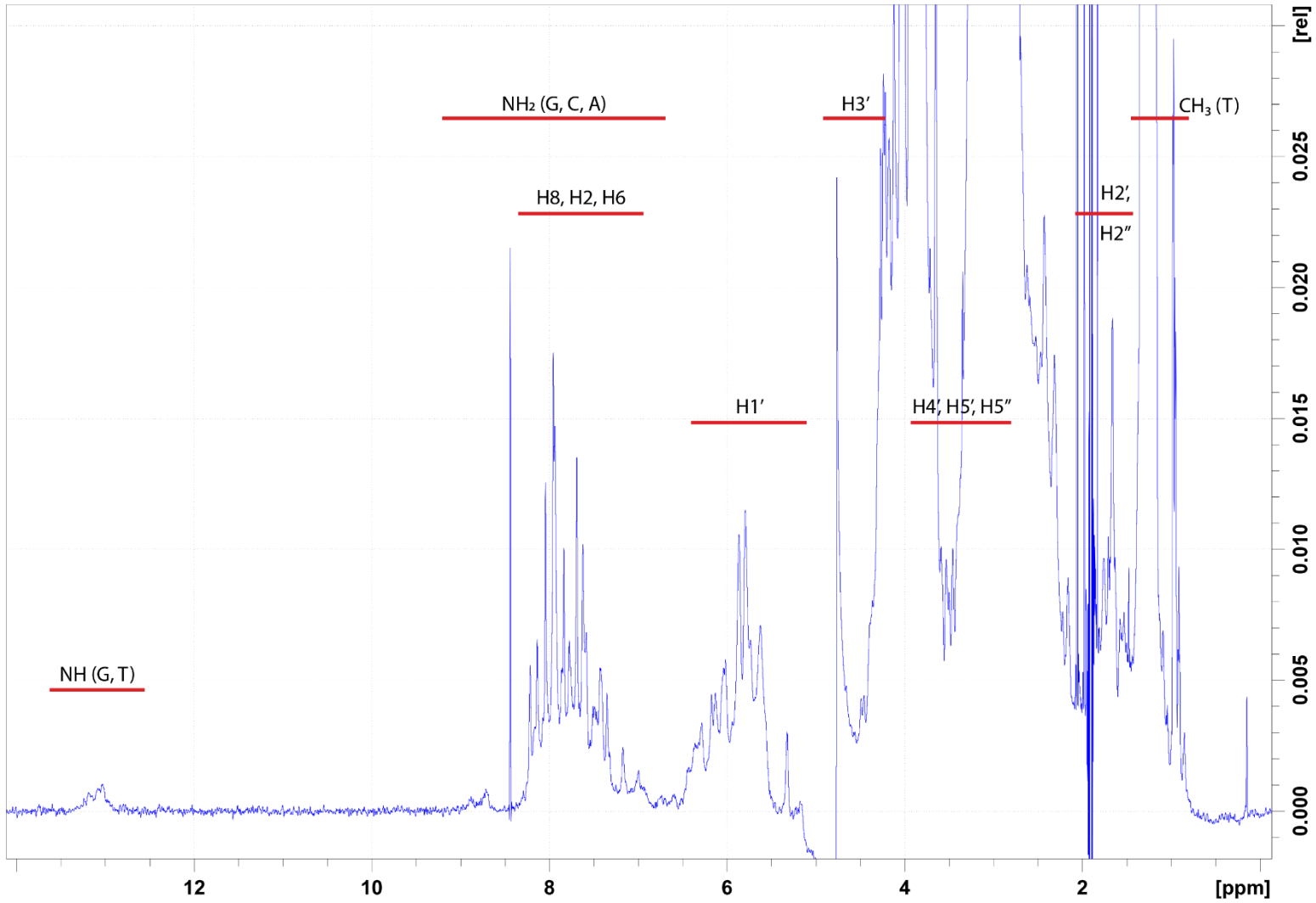

Figure 12: ^{1}H spectrum of SN(GCTA) at 298 K. Typical chemical shift ranges of hydrogens in DNA are shown according to [52].

One dimensional variable-temperature ^{1}H NMR spectra were obtained in the temperature range of 288 K - 348 K. From the temperature dependency of spectra, the changes in chemical shifts, as well as the intensity of proton signals, can be followed. Variability of these effects on similar types of atoms indicates that the corresponding hydrogens are affected by a conformational rearrangement or chemical exchange. Such effect is most apparent in the region between 7.1 ppm and 8.3 ppm of every sample (Figure 13-Figure 16). The chemical shift assignments for the signals are not yet available, therefore it is not possible to interpret the changes on a sequence-specific level. However, non-labile aromatic protons exhibit differential chemical shift changes, both in magnitude and direction, suggesting that the different parts of the molecule do not respond uniformly to temperature change. In addition, the gradual line broadening of NH_2 signals can be attributed to the hindered rotation of these groups at lower temperatures, perhaps due to their stabilization through hydrogen bonding.

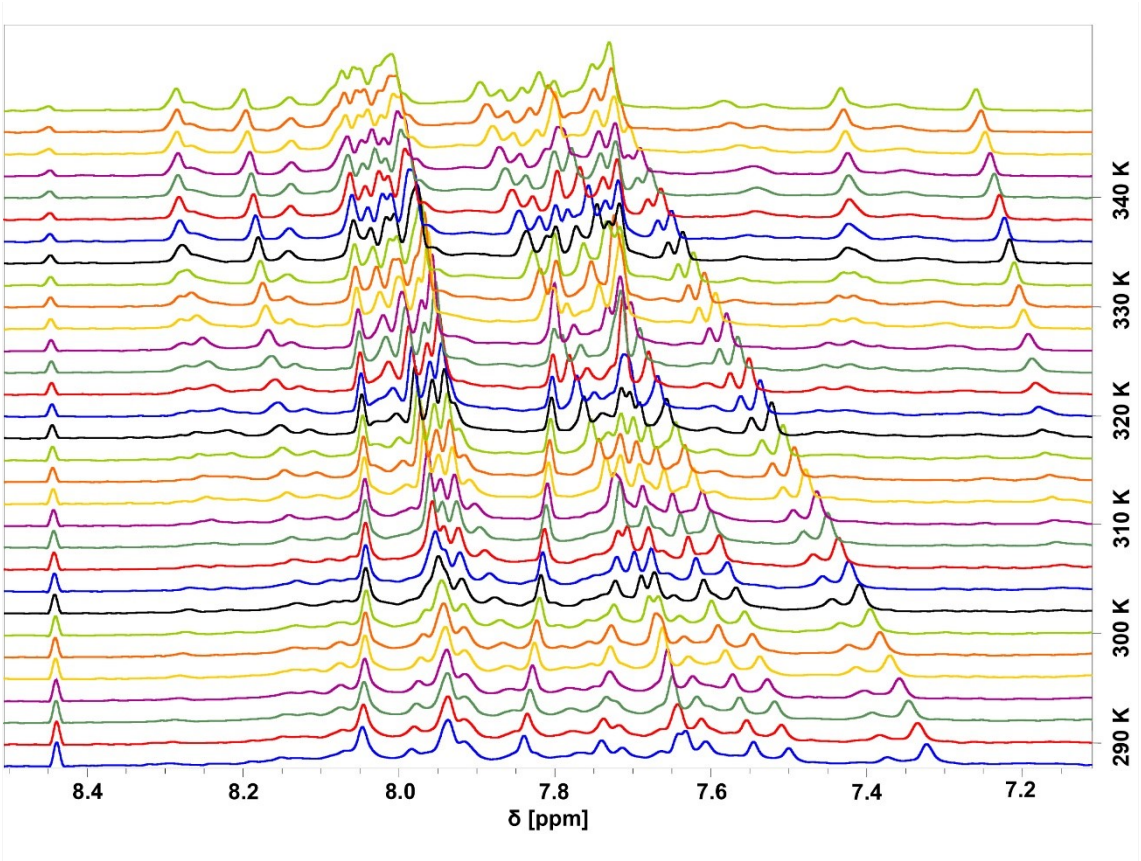

Figure 13: The temperature dependence of ^{1}H spectra in 7.1 - 8.5 ppm region of SN(AAAA).

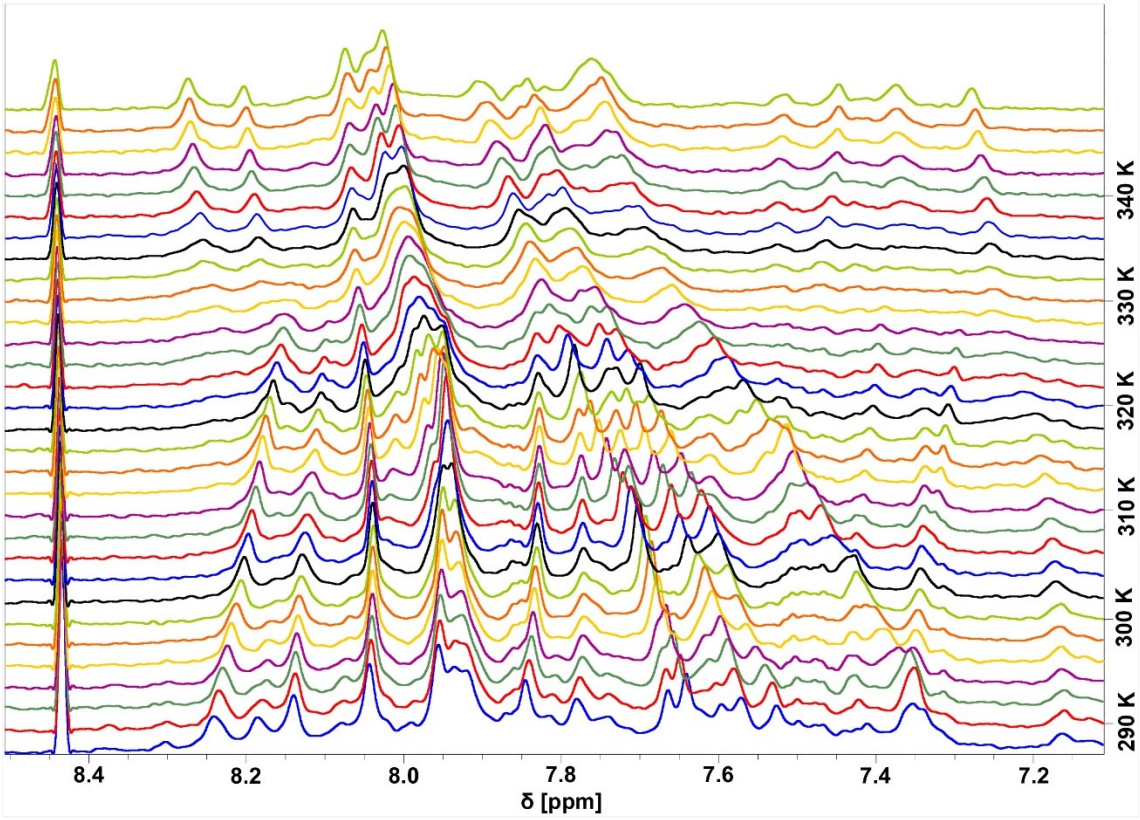

Figure 14: The temperature dependence of ^{1}H spectra in 7.1 - 8.5 ppm region of SN(GCTA).

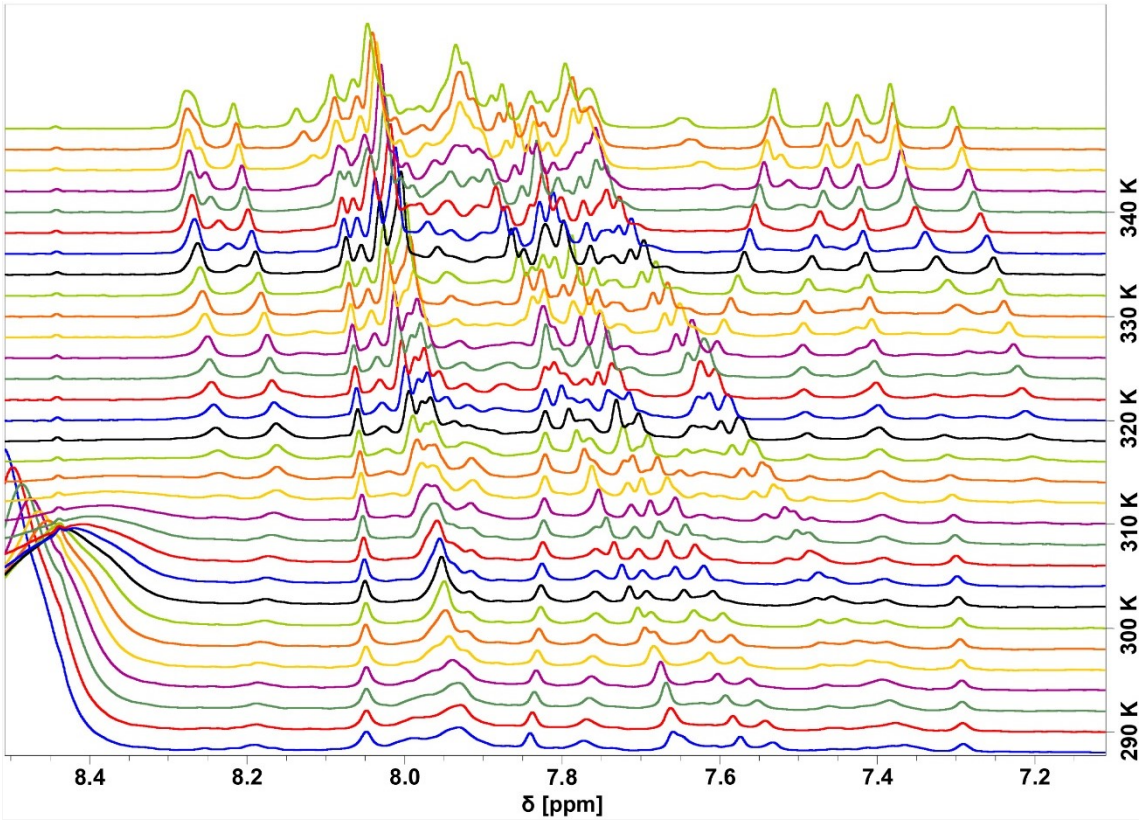

Figure 15: The temperature dependence of ^{1}H spectra in 7.1 - 8.5 ppm region of SN(GCTT).

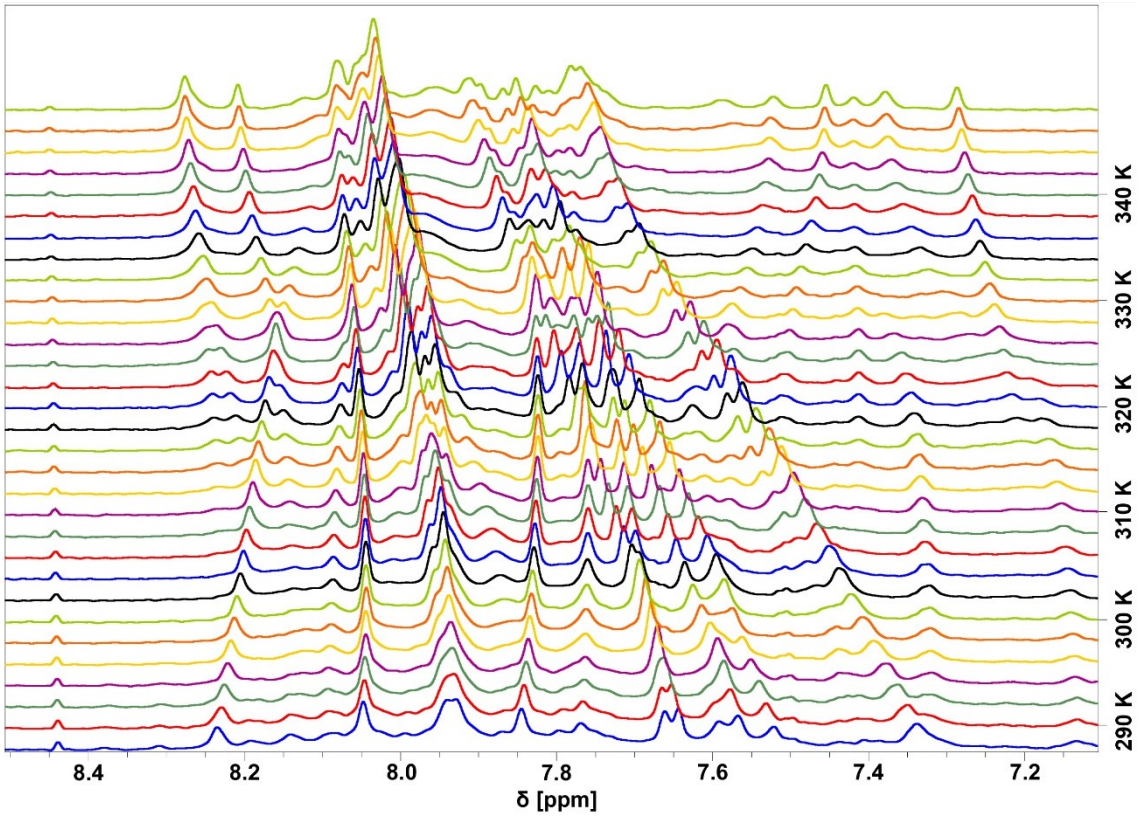

Figure 16: The temperature dependence of ^{1}H spectra in 7.1 - 8.5 ppm region of SN(GCA).

Another suitable region for observing temperature-induced changes is the imino-proton part of the spectra. The signals of imino-protons binding through hydrogen bonds in the Supernova triple helix are only detectable at low temperatures. As the temperature increases, the protons exchange with water via helix destabilization, and their resonances decrease in intensity. Such behavior is known as spectrophotometric melting [56]. The decrease of imino-protons intensities across temperature ranges of all samples can be observed in Figure 17 (A, B, C, D), and the temperature dependence of proton intensities is illustrated in Figure 18.

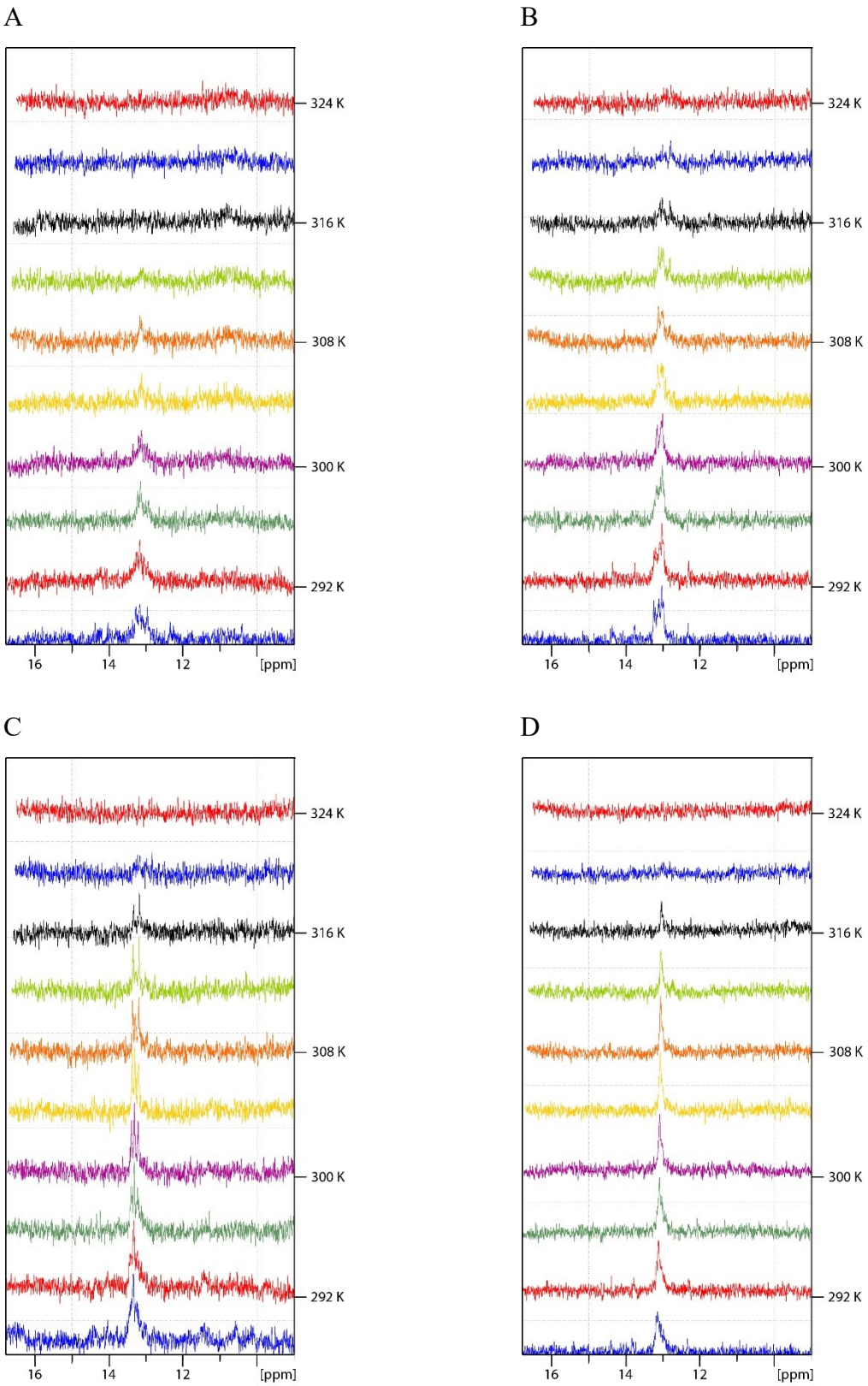

Figure 17: The cutouts of imino-proton region in temperature-dependent ^{1}H spectra of Supernova variants. (A), SN(AAAA); (B), SN(GCTA); (C), SN(GCTT); (D), SN(GCA)

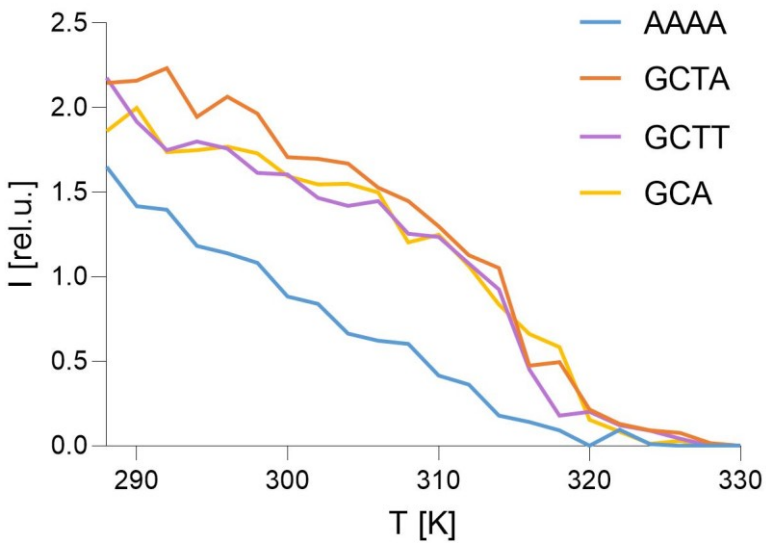

Figure 18: The temperature dependence of the imino-proton signal intensity of all Supernova variants.

The intensity of imino-protons signals is low compared to other signals in spectra, but its general temperature-dependent decrease can be observed. The signals from GCTA, GCTT, and GCTA samples display similar trends; their intensity decreases by half at ca. 315 K, whereas the weakest signal belongs to the standard variant of the Supernova SN(AAAA) and its intensity is at 50% at ca. 300 K. This difference can be associated with a destabilization of the triple helix by the connecting AAAA loop.

2D NMR spectra measurements will be performed in the future to attempt an assignment of individual hydrogens. This will facilitate the more detailed interpretation of temperature-induced changes in the Supernova structure.

4.2 Study of Supernova activity

A study by Svehlova *et al*. (2021) [39] argued that the nucleotide sequence of a loop in Supernova's core is not essential for its catalytic properties, and therefore all four deoxyribozyme variants should exhibit light-producing activity in a reaction with CDP-Star. To test this, a direct one-hour-long measurement of chemiluminescence was performed on a plate reader in standard conditions of 27 °C (Figure 19).

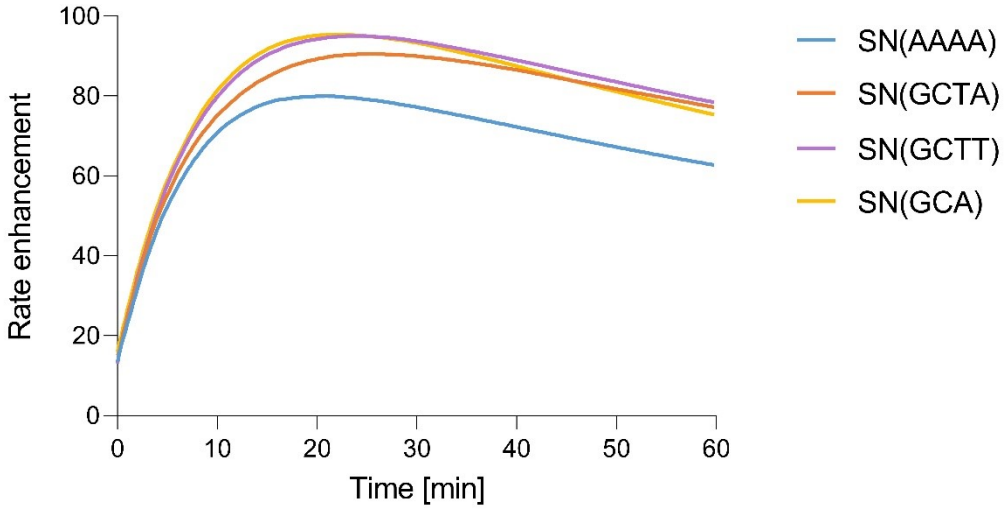

Figure 19: Light production rate enhancement of Supernova variants at 27 °C. The data represent the average of four independent measurements.

Maximal rate enhancements, which peaked at ~90-fold after 20-25 minutes, were normalized to SN(AAAA) that was considered the standard and plotted on a column graph (Figure 20). A high error rate of the light-production measurement caused that the statistical analysis assessed the differences in maximal rate enhancements as insignificant ($P > 0{,}05$). Analysis was performed by Dunnett's multiple comparisons test, which compared the mean of every column to the standard.

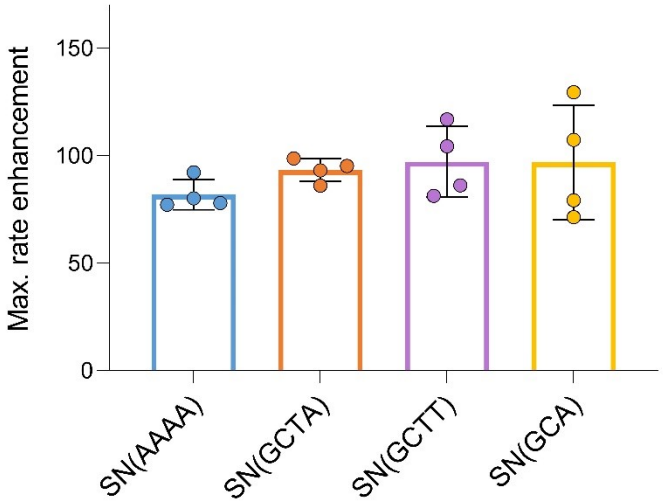

Figure 20: Maximal rate enhancements of light production of Supernova variants normalized to standard SN(AAAA). Points represent values of four independent experiments, and data are presented as mean ± SD.

Phosphate-transfer analysis of Supernova variants was performed to provide an insight into the structural thermostability of its active core. Ligation yields of all tested variants (Figure 21 A, B, C, D) were normalized to a positive control (Supernova phosphorylated by kinase) and showed similar trends regarding the temperature at which the light-producing reaction was incubated. Around 60% of Supernova molecules were phosphorylated at 15 °C, the activity reached a maximal plateau at 20 °C, and a sigmoid decay was observed with further temperature increase. Points in a range of 20 °C - 70 °C (293.15 K - 343.15 K) were fitted by a sigmoidal curve in GraphPad Prism 9.2.0 software, and its parameters are shown in Table 3.

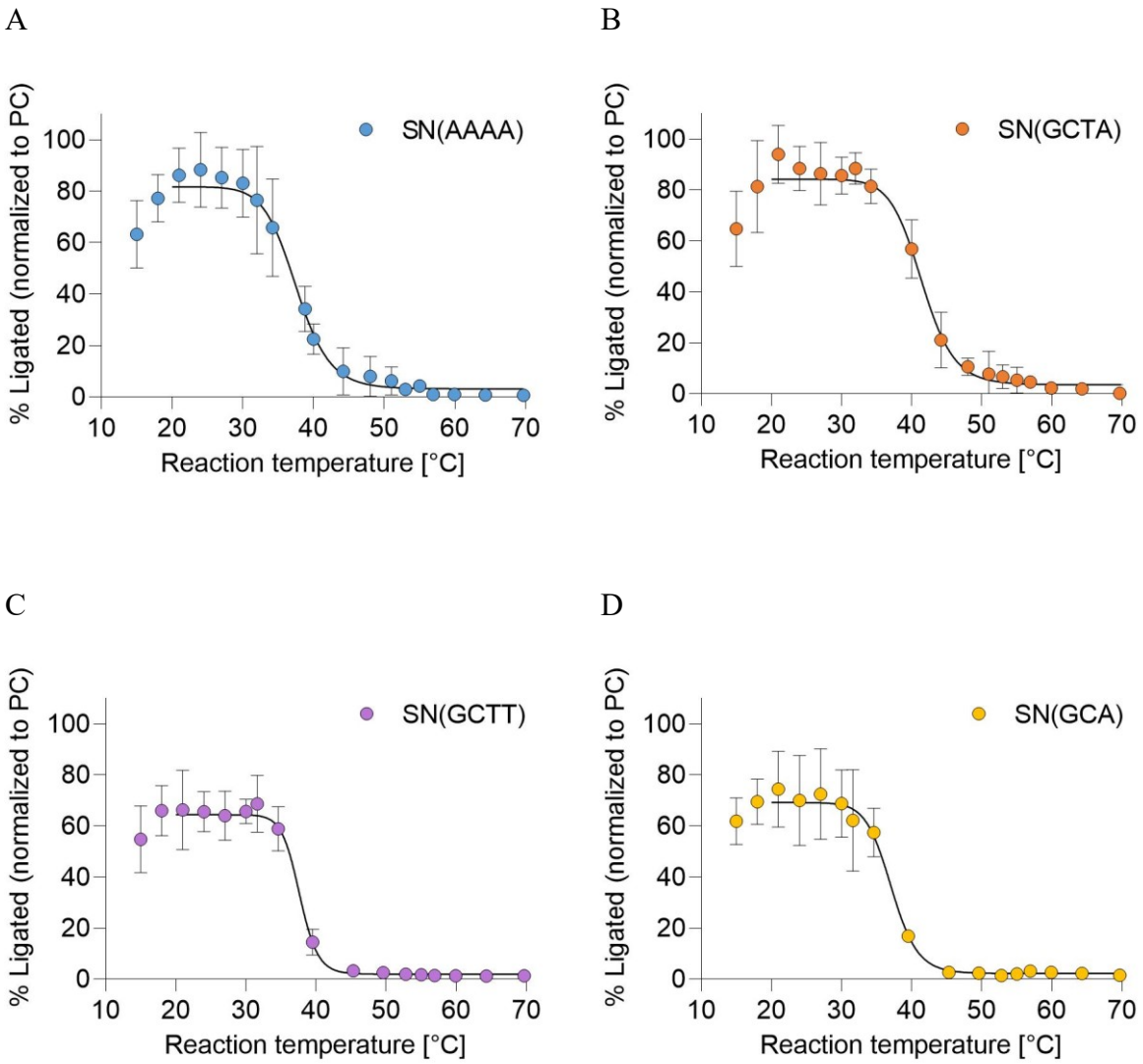

Figure 21: Phosphorylation yields of studied Supernova variants. Graphs show the dependence of formed ligation product on the reaction temperature of Supernova with CDP-Star. Points represent the mean ± SD of three independent experiments. Data are analyzed by sigmoidal, four-parameter logistic curve (4PL) with the least-squares fit.

Table 3: Parameters of the sigmoidal curve fit of phosphorylation yields.

Sample	Inflection point (95% CI) [°C]	Maximum (95% CI) [°C]	R^2
SN(AAAA)	37.57 (36.32–38.73)	81.70 (76.91–86.68)	0.9214
SN(GCTA)	41.46 (40.23–42.72)	84.24 (80.06–88.54)	0.9407
SN(GCTT)	37.66 (36.47–38.68)	64.43 (61.74–67.14)	0.9499
SN(GCA)	37.13 (35.77–38.52)	69.16 (64.71–73.86)	0.9251

The activity of Supernova was shown to decrease by half at around 37 °C for variants containing loop sequences AAAA, GCTT, and GCA; meanwhile, Supernova with GCTA tetraloop displayed the highest activity at an increased temperature and reached its half-maximum at around 41.5 °C.

Phosphorylation yields at 27 °C were plotted on a column graph (Figure 22). Similarly to light-production measurements, the statistical analysis determined that the differences in phosphorylation yields are insignificant ($P > 0{,}05$). Analysis was performed by Dunnett's multiple comparisons test, which compared the mean of every column to the standard SN(AAAA).

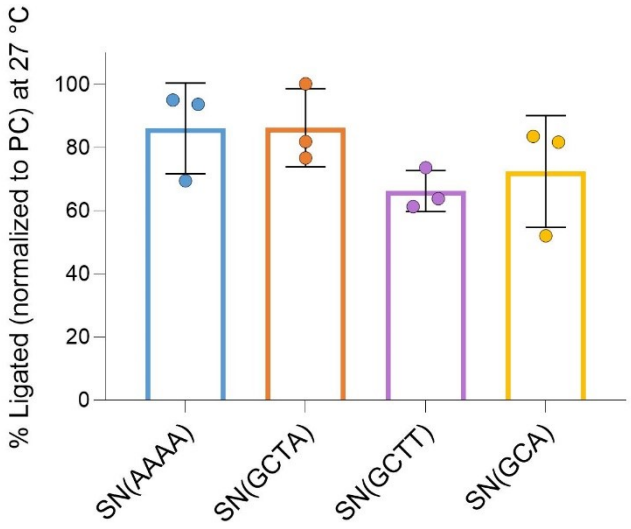

Figure 22: Phosphorylation yields of Supernova variants at 27 °C. Points represent values of three independent experiments, and data are presented as mean ± SD.

From these results, it can be concluded that at 27 ° C, all Supernova variants display

similar activity in both light-production and self-phosphorylation. This is consistent with the hypothesis that the individual loop nucleotides do not play a catalytic role in the active center of the deoxyribozyme. However, the investigation of self-phosphorylation at higher temperatures revealed that the variant with GCTA tetraloop is capable of catalytic activity at temperatures higher by ca. 5 °C, indicating that the loop sequence contributes to the thermostabilization of the active site.

5 Conclusion

The thermostability of four Supernova variants was studied by NMR spectroscopy. ^{1}H temperature-dependent NMR spectra were measured, and the melting transition of deoxyribozyme structure was investigated. Interpretation of the spectra showed that different Supernova variants do not respond uniformly to temperature changes, and the least thermostable molecule contains a tetraloop with the nucleotide sequence AAAA in its structure.

To determine the effect of different loops sequences in Supernova's core on its activity, chemiluminescent experiments were performed at 27 ° C and revealed that all variants catalyze light production at similar rates. This was also confirmed by phosphorylation analysis at the same temperature. However, the differences between these variants became apparent at higher temperatures; the deoxyribozyme with GCTA tetraloop inserted in its core was shown to retain its catalytic activity at temperatures higher by ca. 5 °C than the other variants.

The unusual activity of Supernova with GCTA loop demonstrated at higher temperatures was not observed as an increase of the thermostability of its structure in NMR spectra. Instead, it suggests that the Supernova's activity in stringent conditions depends not only on the rigidity of its structure, but rather on a combination of different factors that are yet to be understood.

Acknowledgment

I would like to thank my advisor, Ed Curtis, for his guidance throughout the first three years of my scientific journey. I am grateful for the learning opportunities and numerous inspiring discussions. My thanks belong to my supervisor, Zdeněk Tošner, for his time and advice, to my co-workers at IOCB for their support and invaluable help with my experiments, and to Václav Veverka and Pavel Srb for their assistance with NMR spectroscopy.

I thank my friends for their encouragement at every step of the way, and a special thanks belongs to my family for their help and for being my role models, both science-wise and life-wise.

References

1. 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. 31, 147 (1982).
2. Haugen, P., Simon, D.M., Bhattacharya, D.: The natural history of group I introns. Trends Genet. TIG. 21, 111 (2005).
3. 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. 35, 849 (1983).
4. Stark, B.C., Kole, R., Bowman, E.J., Altman, S.: Ribonuclease P: an enzyme with an essential RNA component. Proc. Natl. Acad. Sci. U. S. A. 75, 3717 (1978).
5. Walker, S.C., Engelke, D.R.: Ribonuclease P: The Evolution of an Ancient RNA Enzyme. Crit. Rev. Biochem. Mol. Biol. 41, 77 (2006).
6. Prody, G.A., Bakos, J.T., Buzayan, J.M., Schneider, I.R., Bruening, G.: Autolytic processing of dimeric plant virus satellite RNA. Science. **231**, 1577 (1986).
7. Jimenez, R.M., Delwart, E., Lupták, A.: Structure-based Search Reveals Hammerhead Ribozymes in the Human Microbiome. J. Biol. Chem. **286**, 7737 (2011).
8. Seehafer, C., Kalweit, A., Steger, G., Gräf, S., Hammann, C.: From alpaca to zebrafish: Hammerhead ribozymes wherever you look. RNA. **17**, 21 (2011).
9. Gilbert, W.: Origin of life: The RNA world. Nature. **319**, 618 (1986).
10. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P.: The RNA World and the Origins of Life. Mol. Biol. Cell 4th Ed. (2002).
11. Cech, T.R.: The RNA Worlds in Context. Cold Spring Harb. Perspect. Biol. **4**, a006742 (2012).
12. Bernhardt, H.S.: The RNA world hypothesis: the worst theory of the early evolution of life (except for all the others)a. Biol. Direct. **7**, 23 (2012).
13. Higgs, P.G., Lehman, N.: The RNA World: molecular cooperation at the origins of life. Nat. Rev. Genet. **16**, 7 (2015).
14. Lu, Y., Liu, J., Li, J., Bruesehoff, P.J., Pavot, C.M.-B., Brown, A.K.: New highly sensitive and selective catalytic DNA biosensors for metal ions. Biosens. Bioelectron. **18**, 529 (2003).
15. Wilson, D.S., Szostak, J.W.: In Vitro Selection of Functional Nucleic Acids. Annu. Rev. Biochem. **68**, 611 (1999).
16. Breaker, R.R., Joyce, G.F.: A DNA enzyme that cleaves RNA. Chem. Biol. **1**, 223 (1994).
17. Silverman, S.K.: In vitro selection, characterization, and application of deoxyribozymes that cleave RNA. Nucleic Acids Res. **33**, 6151 (2005).
18. McManus, S.A., Li, Y.: The Structural Diversity of Deoxyribozymes. Molecules. **15**, 6269 (2010).
19. Ma, L., Liu, J.: Catalytic Nucleic Acids: Biochemistry, Chemical Biology, Biosensors, and Nanotechnology. iScience. **23**, 100815 (2020).
20. Lan, T., Furuya, K., Lu, Y.: A highly selective lead sensor based on a classic lead DNAzyme. Chem. Commun. Camb. Engl. **46**, 3896 (2010).
21. Santoro, S.W., Joyce, G.F.: A general purpose RNA-cleaving DNA enzyme. Proc. Natl. Acad. Sci. U. S. A. **94**, 4262 (1997).

22. Hwang, K., Wu, P., Kim, T., Lei, L., Tian, S., Wang, Y., Lu, Y.: Photocaged DNAzymes as a General Method for Sensing Metal Ions in Living Cells. Angew. Chem. Int. Ed Engl. **53**, 13798 (2014).
23. Santoro, S.W., Joyce, G.F., Sakthivel, K., Gramatikova, S., Barbas CF 3rd, null: RNA cleavage by a DNA enzyme with extended chemical functionality. J. Am. Chem. Soc. **122**, 2433 (2000).
24. Li, J., Zheng, W., Kwon, A.H., Lu, Y.: In vitro selection and characterization of a highly efficient Zn(II)-dependent RNA-cleaving deoxyribozyme. Nucleic Acids Res. **28**, 481 (2000).
25. Breaker, R.R., Joyce, G.F.: A DNA enzyme with Mg(2+)-dependent RNA phosphoesterase activity. Chem. Biol. **2**, 655 (1995).
26. Zhou, W., Zhang, Y., Ding, J., Liu, J.: In Vitro Selection in Serum: RNA-Cleaving DNAzymes for Measuring Ca2+ and Mg2+. ACS Sens. **1**, 600 (2016).
27. Hollenstein, M., Hipolito, C., Lam, C., Dietrich, D., Perrin, D.M.: A highly selective DNAzyme sensor for mercuric ions. Angew. Chem. Int. Ed Engl. **47**, 4346 (2008).
28. Lin, Y.-W., Liu, C.-W., Chang, H.-T.: Fluorescence detection of mercury(II) and lead(II) ions using aptamer/reporter conjugates. Talanta. **84**, 324 (2011).
29. Huang, P.-J.J., Liu, J.: Rational evolution of Cd2+-specific DNAzymes with phosphorothioate modified cleavage junction and Cd2+ sensing. Nucleic Acids Res. **43**, 6125 (2015).
30. Huang, P.-J.J., Lin, J., Cao, J., Vazin, M., Liu, J.: Ultrasensitive DNAzyme beacon for lanthanides and metal speciation. Anal. Chem. **86**, 1816 (2014).
31. Saran, R., Liu, J.: A Silver DNAzyme. Anal. Chem. **88**, 4014 (2016).
32. Torabi, S.-F., Wu, P., McGhee, C.E., Chen, L., Hwang, K., Zheng, N., Cheng, J., Lu, Y.: In vitro selection of a sodium-specific DNAzyme and its application in intracellular sensing. Proc. Natl. Acad. Sci. U. S. A. **112**, 5903 (2015).
33. Wu, P., Hwang, K., Lan, T., Lu, Y.: A DNAzyme-Gold Nanoparticle Probe for Uranyl Ion in Living Cells. J. Am. Chem. Soc. **135**, 5254 (2013).
34. Achenbach, J.C., Nutiu, R., Li, Y.: Structure-switching allosteric deoxyribozymes. Anal. Chim. Acta. **534**, 41 (2005).
35. Mao, X., Simon, A.J., Pei, H., Shi, J., Li, J., Huang, Q., Plaxco, K.W., Fan, C.: Activity modulation and allosteric control of a scaffolded DNAzyme using a dynamic DNA nanostructure. Chem. Sci. **7**, 1200 (2016).
36. Zheng, X., Yang, J., Zhou, C., Zhang, C., Zhang, Q., Wei, X.: Allosteric DNAzyme-based DNA logic circuit: operations and dynamic analysis. Nucleic Acids Res. **47**, 1097 (2019).
37. Guo, Y., Liu, J., Yang, G., Sun, X., Chen, H.-Y., Xu, J.-J.: Multiple turnovers of DNAzyme for amplified detection of ATP and reduced thiol in cell homogenates. Chem. Commun. **51**, 862 (2014).
38. Liu, F., Zhang, J., Chen, R., Chen, L., Deng, L.: Highly effective colorimetric and visual detection of ATP by a DNAzyme-aptamer sensor. Chem. Biodivers. **8**, 311 (2011).
39. Svehlova, K., Lukšan, O., Jakubec, M., Curtis, E.A.: Supernova: a deoxyribozyme that catalyzes a chemiluminescent reaction. Manuscript submitted for publication.
40. Olesen, C.E.M., Mosier, J., Martin, C.S., Voyta, J.C., Bronstein, I.: 1, 2-Dioxetane chemiluminescent detection of proteins and nucleic acids. SEIBUTSU BUTSURI KAGAKU. **42**, 265 (1998).

41. Li, Y., Breaker, R.R.: Phosphorylating DNA with DNA. Proc. Natl. Acad. Sci. **96**, 2746 (1999).
42. Coleman, J.E.: Structure and mechanism of alkaline phosphatase. Annu. Rev. Biophys. Biomol. Struct. **21**, 441 (1992).
43. Lohman, G.J.S., Tabor, S., Nichols, N.M.: DNA Ligases. Curr. Protoc. Mol. Biol. **94**, 3.14.1 (2011).
44. Robertson, A.D., Murphy, K.P.: Protein Structure and the Energetics of Protein Stability. Chem. Rev. **97**, 1251 (1997).
45. Petsko, G.A.: Structural basis of thermostability in hyperthermophilic proteins, or 'there's more than one way to skin a cat'. Methods Enzymol. **334**, 469 (2001).
46. Brown, J.W., Haas, E.S., Pace, N.R.: Characterization of ribonuclease P RNAs from thermophilic bacteria. Nucleic Acids Res. **21**, 671 (1993).
47. Guo, F., Cech, T.R.: Evolution of Tetrahymena ribozyme mutants with increased structural stability. Nat. Struct. Biol. **9**, 855 (2002).
48. Nakano, M., Moody, E.M., Liang, J., Bevilacqua, P.C.: Selection for thermodynamically stable DNA tetraloops using temperature gradient gel electrophoresis reveals four motifs: d(cGNNAg), d(cGNABg),d(cCNNGg), and d(gCNNGc). Biochemistry. **41**, 14281 (2002).
49. Günther, H.: NMR spectroscopy. John Wiley & Sons Ltd, Chichester 1983
50. Keeler, J.: Understanding NMR spectroscopy. John Wiley & Sons Ltd, Chichester 2005
51. Bransden, B.H., Joachain, C.J.: Physics of Atoms and Molecules. Pearson Education 2003, p. 1132.
52. Piccioli, M., Turano, P.: Transient iron coordination sites in proteins: Exploiting the dual nature of paramagnetic NMR. Coord. Chem. Rev. **284**, 313 (2015).
53. SGALLOVÁ, R. Termodynamika tvorby DNA vlásenek. Praha (2019). Diplomová práce. Univerzita Karlova, Matematicko-fyzikální fakulta, Katedra fyziky nízkých teplot. Vedoucí práce Římal, V.
54. Brown, T., Brown, T.: Nucleic Acids Book. https://www.atdbio.com/nucleic-acids-book.
55. Germann, M.W.: Nucleic Acids NMR Spectroscopy. Georgia State University.
56. Jaroszewski, J., Clausen, V., Cohen, J., Dahl, O.: NMR Investigations of Duplex Stability of Phosphorothioate and Phosphorodithioate DNA Analogues Modified in Both Strands. Nucleic Acids Res. **24**, 829 (1996).