Heinrich Heine
Universität
Düsseldorf

Understanding metal-mediated catalysis:
structural and functional analysis of an RNA-cleaving DNAzyme and Fe–S cluster containing proteins

Inaugural-Dissertation

zur Erlangung des Doktorgrades der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

vorgelegt von
Hannah Rosenbach
aus Duisburg

Düsseldorf, Juni 2020

aus dem Institut für Physikalische Biologie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Jun.-Prof. Dr. Ingrid Span
Korreferent: Prof. Dr. Peter Gilch

Tag der mündlichen Prüfung: 28.08.2020

Contents

Zusammenfassung

Die vorliegende Dissertation befasst sich mit der metallabhängigen Reaktivität von Biomolekülen und besteht aus zwei Kapiteln. Das erste Kapitel befasst sich mit dem Mg^{2+}-abhängigen, artifiziellen 10-23 Desoxyribozym (DNAzym), einem einzelsträngigen DNA-Molekül, welches die Spaltung von RNA-Strängen katalysiert. DNAzyme bestehen aus einem katalytischen Kern und zwei flankierenden Erkennungssequenzen, die mit dem RNA-Substrat Doppelstränge ausbilden. In Anwesenheit von divalenten Kationen wie Mg^{2+} oder Mn^{2+} können sie an ein spezifisches RNA-Substrat binden und dieses hydrolysieren, woraufhin die RNA-Produkte anschließend dissoziieren. Obwohl das 10-23 DNAzym bereits vor über zwei Jahrzehnten mittels *In-vitro*-Selektion isoliert wurde, ist der Mechanismus, mit dem das DNAzym die Spaltung katalysiert, weitestgehend unbekannt. Der Grund hierfür liegt im Fehlen hochauflösender, struktureller Daten. Die Sammlung solcher strukturellen Daten auf atomarer Ebene würde die Grundlage schaffen, um die Funktionsweise dieser Biokatalysatoren zu verstehen und den Mechanismus der Reaktion aufzuklären.

Artikel eins (Rosenbach *et al.* (2020) Review article on the 10-23 DNAzyme) bietet in Form eines Reviews Einblicke in den Stand der Forschung im Hinblick auf Struktur, Funktion und Metallabhängigkeit des 10-23 DNAzyms. Hierbei werden alle Modifikationen (Mutationen und Deletionen) an unterschiedlichen Positionen im DNAzym systematisch zusammengefasst. Weiterhin werden die Effekte unterschiedlicher Metallionen auf die DNAzym-Kinetik zusammengefasst. Die kritische Analyse dieser Daten unter Berücksichtigung der Reaktionsbedingungen liefert neue Erkenntnisse für zukünftige Ansätze zur Strukturaufklärung des 10-23 DNAzyms.

In Artikel zwei (Rosenbach *et al.* (2020) Metal ion dependency of the 10-23 DNAzyme) wird der Einfluss unterschiedlicher mono- und divalenter Metallionen (Na^{+}, K^{+}, Mg^{2+} und Mn^{2+}) auf die Reaktionsgeschwindigkeit der RNA-Hydrolyse, die durch das 10-23 DNAzym katalysiert wird, untersucht. Mithilfe FRET-basierter Aktivitätsmessungen, ITC-Messungen sowie NMR- und EPR-Spektroskopie konnte gezeigt werden, dass monovalente Kationen das Bindungsverhalten von Mg^{2+} und Mn^{2+} an den DNAzym:RNA Komplex unterschiedlich beeinflussen: In Anwesenheit von steigenden Konzentrationen an K^{+}- und Na^{+}-Ionen wird die Mg^{2+}-vermittelte Spaltung verlangsamt, während die Mn^{2+}-vermittelte Reaktion beschleunigt wird. Die unterschiedlichen Korrelationen deuten darauf hin, dass der Mechanismus der Katalyse in Anwesenheit von Mg^{2+} und Mn^{2+} unterschiedlich ist. Während Mg^{2+} und Na^{+} um die jeweiligen Bindungsstellen konkurrieren, wird die Bindung von Mn^{2+} an den DNAzym:RNA Komplex durch die Na^{+}-Ionen unterstützt. Unsere spektroskopischen Messungen geben detaillierte Einblicke in die dynamische Assoziation der Metallionen mit dem DNAzym.

Die Strukturaufklärung von Nukleinsäuren mittels Röntgenstrukturanalyse stellt eine große Herausforderung dar. Zum einen wird die Oberfläche von DNA- und RNA-Molekülen durch negativ geladene Phosphatgruppen dominiert, was häufig zu Kristallen mit einer geringen Fernordnung führt. Zum anderen stellt das Phasenproblem eine weitere große Hürde dar.

Artikel drei (Rosenbach *et al.* (2020) U1A as crystallization tool) befasst sich daher mit der Etablierung eines neuen und effizienten Protokolls für die Kristallisation von Nukleinsäu-

ren mit Hilfe des bereits häufig als Kristallisationshelfer genutzten humanen RNA-bindenden Proteins U1A. Der Artikel präsentiert Kristallstrukturen von drei U1A-Varianten, die im Gegensatz zur herkömmlich genutzten Sequenz ein Tryptophan enthalten. Diese U1A-Varianten erweitern nicht nur das Repertoire an Kristallisationshelfern, sondern ermöglichen zusätzlich eine einfache Unterscheidung zwischen Protein- und Salzkristallen mittels Detektion der intrinsichen Fluoreszenz von Tryptophan. Untersuchungen eines ternären Komplexes aus U1A, 10-23 DNAzym und RNA-Substrat mittels NMR-Spektroskopie und Kleinwinkelröntgenstreuung zeigen, dass das U1A-Protein keinen Einfluss auf die Struktur des katalytischen Zentrums des 10-23 DNAzyms hat und somit hohes Potenzial als Kristallisationshelfer für die Strukturaufklärung des ternären Komplexes besitzt. Zusätzlich konnte ein schnelles und effizientes Protokoll zur Herstellung von Protein-RNA-Kristallen etabliert werden. Hierbei werden U1A-Protein-Kristalle mit RNA-Molekülen inkubiert. Diese bilden mit dem vorgeformten Proteinkristall einen Komplex aus, welcher dann über fluoreszenz-basierte Methoden die Detektion der RNA-Bindung *in crystallo* ermöglicht.

In Artikel vier (Rosenbach *et al.* (2020) Crystallization of the 10-23 DNAzyme) werden unterschiedliche Strategien zur Kristallisation des 10-23 DNAzyms im Komplex mit seinem RNA-Substrat vorgestellt, die alle darauf abzielen, die Struktur des DNAzym:RNA-Komplexes in einer katalytisch relevanten Konformation aufzuklären. Hier erwiesen sich vor allem die Co-Kristallisation des DNAzym:RNA-Komplexes mit dem RNA-bindenden Protein U1A als Kristallisationshelfer, als auch kombinatorische Kristallisationsexperimente mit binären Komplexen aus DNAzymen und RNA-Substraten unterschiedlicher Länge als besonders vielversprechend.

Das zweite Kapitel der vorliegenden Dissertation befasst sich mit der spektroskopischen Untersuchung von verschiedenen Eisen-Schwefel-Proteinen. Artikel fünf (Rosenbach *et al.* (2020) Maturation of Fe–S cluster containing proteins) beschreibt die systematische Untersuchung der Auswirkungen unterschiedlicher Expressionssysteme auf Expressionslevel, Proteinausbeuten und den Gehalt an Fe–S-Zentren im Protein. Die Biosynthese der Fe–S-Proteine wurde anschließend mit der Herstellung der Holoproteine durch chemische oder semi-enzymatische Rekonstitution verglichen. Unsere Daten zeigen, dass die spezifischen Expressionssysteme sowohl die Proteinausbeute als auch die Spezifität des Einbaus verbessern. Zudem zeigte sich, dass nicht ein einziges System für alle Zielproteine die besten Resultate lieferte. Die neuen Erkenntnisse aus den Studien drei etablierter [4Fe–4S]-Proteine konnten dann auf das Radikal-SAM-Enzym ThnB angewendet werden, welches bislang nur durch chemische Rekonstitution maturiert wurde.

Artikel sechs (Rosenbach *et al.* (2020) Spectroscopic characterization of Asp1$^{365\text{-}920}$) befasst sich mit der Pyrophosphatase-Domäne des bifunktionalen Enzyms Asp1 aus *Schizosaccharomyces pombe*. Das Asp1-Protein gehört zu der hoch konservierten Familie der Diphosphoinositol-Pentakisphosphat-Kinasen PPIP5K/Vip1. Die Fähigkeit des Proteins ein [2Fe–2S]-Zentrum zu binden, wurde erst kürzlich entdeckt und die biologische Relevanz des Zentrums wurde bislang noch nicht aufgeklärt. In dieser Arbeit konnte gezeigt werden, dass das Protein nach rekombinanter Herstellung in einem speziell für die Produktion von Fe–S-haltigen Proteinen optimierten *Escherichia-coli*-Expressionsstamms mit einem [2Fe–2S]-Zentrum isoliert

werden kann. Mittels verschiedener spektroskopischer Techniken konnte gezeigt werden, dass das [2Fe–2S]-Zentrum von vier Cystein-Liganden an den Positionen 607, 663, 864 und 879 in der Aminosäuresequenz koordiniert wird. Unsere Daten deuten darauf hin, dass das Fe–S-Zentrum O_2-empfindlich und höchstwahrscheinlich nicht an Redox-Reaktionen beteiligt ist. Im Gegensatz zu vorangegangenen Studien an einer aerob isolierten und chemisch rekonstituierten Proteinprobe konnte mithilfe von *in-vivo-* und *in-vitro*-Experimenten gezeigt werden, dass das [2Fe–2S]-Zentrum keinen Einfluss auf die Pyrophosphatase-Aktivität hat. Diese Ergebnisse deuten darauf hin, dass die Pyrophosphatase-Domäne von Asp1 eine weitere biologische Funktion aufweist, die durch die Anwesenheit eines [2Fe–2S]-Zentrums reguliert wird.

Summary

The present thesis deals with the metal-dependent reactivity of biomolecules and is divided into two chapters. The first chapter deals with the Mg^{2+}-dependent, artificial 10-23 Deoxyribozyme (DNAzyme), a single-stranded DNA molecule that catalyzes the cleavage of RNA strands. The DNAzyme consists of a catalytic core that is flanked by two substrate recognition arms that are capable of forming double helices with RNA substrates. In the presence of Mg^{2+} or Mn^{2+}, the DNAzyme binds its specific RNA substrate and hydrolyzes the RNA between two nucleotides at the cleavage site, then the RNA products dissociate. Although the 10-23 DNAzyme has been isolated during *in vitro* selection processes more than 20 years ago, the mechanism by which the DNAzyme catalyzes the cleavage reaction is still unknown, due to the lack of high-resolution structural data. Collection of such information would provide the basis to obtain a deeper understanding of the function of these biocatalysts and to reveal the mechanism by which the DNAzyme performs the RNA cleavage.

As a review, article 1 (Rosenbach *et al.* (2020) Review article on the 10-23 DNAzyme) provides an overview on the actual state of research with regard to structure, function and metal ion dependency of the 10-23 DNAzyme. For this purpose, all modifications (mutations and deletions) at different positions within the DNAzyme sequence are systematically summarized. Additionally, the effects of different metal ions on the DNAzyme activity are compiled. The critical analysis of the data with regard to the conditions under which the reactions have been performed provide new insights for future attempts to solve the structure of the 10-23 DNAzyme.

In article 2 (Rosenbach *et al.* (2020) Metal ion dependency of the 10-23 DNAzyme) the influence of different mono- and divalent metal ions (Na^{+}, K^{+}, Mg^{2+} and Mn^{2+}) on the reaction rates of RNA hydrolysis performed by the 10-23 DNAzyme is analyzed. Using FRET-based activity assays, isothermal titration calorimetry (ITC) measurements, as well as nuclear magnetic resonance (NMR) and electron paramagnetic resonance (EPR) spectroscopy it was shown that monovalent cations influence the binding properties of Mg^{2+} and Mn^{2+} to the DNAzyme:RNA complex in different ways: In the presence of K^{+} and Na^{+} ions, the speed of the Mg^{2+}-mediated cleavage reaction is reduced, while the speed of the Mn^{2+}-mediated reaction is enhanced. The different correlations suggest that the mechanism of the cleavage reaction is different for both divalent metal ions. While Mg^{2+} and Na^{+} compete for the same binding site within the DNAzyme:RNA complex, Mn^{2+} ions bind to a Na^{+}-favored structure of the DNAzym:RNA complex. Our spectroscopic analysis allow for deep insights into the dynamic association of metal ions with the DNAzyme.

Structure elucidation of nucleic acids using X-ray crystallography is highly challenging due to the negatively charged and regularly ordered phosphate backbone, which often leads to the formation of crystals with a poor long-range order. Additionally, the phasing is a major hurdle.

Article 3 (Rosenbach *et al.* (2020) U1A as crystallization tool) deals with the establishment of a new and efficient protocol for the crystallization of nucleic acids using the human RNA-binding protein U1A that has been previously used as a crystallization helper for the crystallization of nucleic acids. In this article, three crystal structures of U1A variants are presented that in con-

trast to the commonly used U1A sequence host a tryptophan residue. These U1A variants do not only expand the repertoire of crystallization modules, they also allow for an easy distinction between salt and protein crystals due to their intrinsic tryptophan fluorescence. Analyses of a ternary complex consisting of U1A protein, 10-23 DNAzyme, and RNA substrate using small angle X-ray scattering (SAXS) and NMR spectroscopy reveal that the U1A protein does not influence the structure of the catalytic core of the 10-23 DNAzyme and that therefore, the U1A protein has a high potential for structure elucidation of the ternary complex. In addition, a fast and efficient protocol for the production of protein-RNA crystals has been established. Here, pre-formed U1A protein crystals are incubated with RNA molecules that form a complex with the protein crystals. This binding can be visualized by fluorescence-based methods *in crystallo*.

Article 4 (Rosenbach *et al.* (2020) Crystallization of the 10-23 DNAzyme) presents different strategies for the crystallization of the 10-23 DNAzyme in complex with its RNA substrate, which all aim at solving the structure of the DNA:RNA complex in a catalytically relevant conformation. Here, co-crystallization of the DNAzyme:RNA complex with the RNA-binding protein U1A as a crystallization helper, as well as combinatorial crystallization experiments with a binary complex consisting of DNAzyme and RNA substrate with varying length were most promising.

The second chapter of this thesis deals with the spectroscopic analysis of different iron-sulfur proteins. Article 5 (Rosenbach *et al.* (2020) Maturation of Fe–S cluster containing proteins) describes the systematic analysis of the impact of different expression systems on the expression level, protein yield and the occupancy with Fe–S cluster. The biosynthesis of the Fe–S proteins was then compared to the production of the holo protein using chemical and semi-enzymatic reconstitution. Our data reveal that the use of specific expression systems enhances both protein yield as well as the specificity of the cluster incorporation. Furthermore, it was shown that no general-purpose system exists that provides the best results for every protein. The knowledge from these systematical studies on three well characterized [4Fe–4S] cluster have then been transferred to the production of the radical SAM enzyme ThnB, which has previously only been matured using chemical reconstitution.

Article 6 (Rosenbach *et al.* (2020) Spectroscopic characterization of Asp1$^{365\text{-}920}$) deals with the pyrophosphatase domain of the bifunctional enzyme Asp1 from *S. pombe*. This protein belongs to the highly conserved family of the diphosphoinositol-pentakisphosphate kinasen PPIP5K/Vip1. The capability of the protein to bind a [2Fe–2S] cluster has recently been reported. However, the biological relevance of the cluster is still under debate. In this thesis, the phosphatase domain of Asp1 was isolated with a [2Fe–2S] cluster from an *E. coli* strain that is especially modified for the production of Fe–S cluster proteins. Using different spectroscopic methods, it was shown that the [2Fe–2S] cluster is coordinated by four cysteine residues at the position 607, 663, 864, and 879. Our data suggest that the Fe–S cluster is sensitive towards oxygen and that it is most likely not involved in redox reactions. In contrast to a previous study using aerobically isolated and chemically reconstituted Asp1, using *in vivo* and *in vitro* experiments we could show that the [2Fe–2S] cluster does not influence the pyrophosphatase activity of the protein. Our results point towards an additional biologically relevant function of the domain that is regulated by the presence of [2Fe–2S] cluster.

1 General introduction to DNA-mediated catalysis

1.1 Nucleic acids as catalysts

Until the discovery of ribozymes in the early 1980s, biomolecule-mediated catalysis in nature was exclusively attributed to proteins. The notion that RNA molecules can have intrinsic enzyme-like activity and that they can catalyze chemical reactions in the complete absence of proteins with impressive rate enhancements lent credence to the hypothesis that RNA molecules could have played a central role in the early stages of life. In such an RNA world, RNA could served both as a genetic blueprint as well as a macromulecule with enzymatic function (Higgs & Lehman, 2015).

The first RNA molecule reported to be capable of catalyzing a chemical reaction has been the self-splicing pre-ribosomal RNA of *Tetrahymena* (Kruger *et al.*, 1982). One year later, the first example of a *trans*-acting RNA molecule was reported to be the RNA component of ribonuclease (RNase) P (Guerrier-Takada *et al.*, 1983). Since then, many ribozymes have been discovered that are capable of catalyzing different important reactions *in vivo* such as the site-specific cleavage of RNA (Forster & Symons, 1987) or the formation of peptide bonds in the ribosome (Leung *et al.*, 2011). The introduction of *in vitro* selection methods (Robertson & Joyce, 1990; Tuerk & Gold, 1990) facilitated the identification of synthetic ribozymes with desired properties (Joyce, 2007).

The ability of RNA molecules to catalyze complex chemical reactions with rate enhancements that are compatible to those of proteins is due to their ability to fold into complex three-dimensional structures and their use of metal ions (Shan *et al.*, 1999) or small molecules (Winkler *et al.*, 2004) as cofactors.

The discovery of riboyzmes also fueled the speculation whether single-stranded DNA catalysts exist (Kruger *et al.*, 1982), but until now no such DNA catalyst has been found in nature. RNA and DNA molecules show very similar chemical structures, but the roles of RNA and DNA in nature differ significantly from one another: while double-stranded DNA is a long-term storage of genetic information, RNA plays a more dynamic role. In the form of a messenger molecule, mRNA passes genetic information on to the protein biosynthesis machinery. Therefore, mRNA is an indispensable link between genes and proteins as the main gene products. But in addition, non-coding RNAs carry out other biological functions. Among others, non-coding RNAs include transfer RNAs (tRNAs) (Hoagland *et al.*, 1958) and ribosomal RNAs (rRNAs) (Littlefield *et al.*, 1955; McQuillen *et al.*, 1959; Roberts, 1958) both involved in protein synthesis, as well as small RNAs such as microRNAs (miRNAs) (Lee *et al.*, 1993; Ruvkun, 2001), small-interfering RNAs (siRNAs) (Hamilton & Baulcombe, 1999), both involved in RNA degradation and post-transcriptional gene silencing (PTGS), and small nuclear RNAs (snRNAs) (Hadjiolov *et al.*, 1966) which are involved in the processing of pre-messenger RNA.

1.1.1 Advantages of nucleic acids over proteins as *de novo* catalysts

Compared to proteins which are mostly composed of 20 proteinogenic amino acids, the composition of nucleic acids is less diverse. While RNA sequences comprise ribonucleotides with the purine bases adenine (A) and guanine (G) and the pyrimidine bases cytosine (C) and uracil (U) as their basic building blocks, in DNA sequences, the sugar moiety lacks the 2'-hydroxyl group and the pyrimidine base uracil is replaced by thymine (T). Due to the large variety of amino acids with polar and non-polar, hydrophobic and hydrophilic as well as charged and uncharged residues, proteins exhibit diverse surface properties. In contrast, nucleic acids feature a negatively charged and—in double-stranded nucleic acids—regularly ordered phosphate backbone that dominates the surface properties of these macromolecules. In addition, the diversity of functional groups is significantly reduced in nucleic acids, since—besides the phosphate backbone—they only host hydroxyl and keto groups, as well as primary, secondary, and tertiary amines. Therefore, it may appear counterintuitive that nucleic acids are capable of catalyzing complex chemical reactions with significant rate enhancements (Famulok & Jenne, 1999; Jäschke, 2001; Silverman, 2016). The key to the catalytic activity of nucleic acids is the ability of single-stranded DNA and RNA molecules to fold into complex three-dimensional structures, including binding sites and catalytic centers, and therefore provide an environment in which a number of different reactions can be facilitated.

The diversity in the reaction scope of catalytic nucleic acids has further been prompted by the achievements in *in vitro* evolution methods (Joyce, 2007) that started in the late 1960s with a publication by Spiegelman and co-workers (Mills *et al.*, 1967). This report on "An Extracellular Darwinian Evolution Experiment with a Self-Duplicating Nucleic Acid Molecule" has been recognized as the beginning of a very fruitful area of research. A breakthrough for the *in vitro* evolution of DNA and RNA came with the development of the polymerase chain reaction (PCR) for the amplification of nucleic acids (Saiki *et al.*, 1988, 1985). This general-purpose amplification method for nucleic acids led to the first description of the *in vitro* selection of nucleic acids with defined properties in 1990, called ***S**ystematic **E**volution of **L**igands by **EX**ponential Enrichment* (Tuerk & Gold, 1990) (SELEX, see Section 1.2.3).

Selection of new catalytic nucleic acid sequences with desired properties is possible without any knowledge of the sequence requirements. This is contrary to selection methods for protein enzymes. Here, *in vitro* selection from random sequences, in general, is not feasible due to two main reasons: firstly, since proteins can contain up to 20 different amino acid building blocks, the number of possible interactions between the residues is immense and, secondly, secondary structure elements of proteins such as helices, turns and sheets are unstable, if the structure elements are not incorporated in a complex tertiary structure (Gruenewald *et al.*, 1979). That is why proteins with random amino acid sequences usually are not capable of adopting complex tertiary structures, since protein folding is a cooperative process (Privalov, 1979) and the energies involved in the formation of secondary structures are comparable to those involved in the formation of tertiary structures. However, secondary structure elements of nucleic acids such as hairpin loops extended by a double helix are stable in the absence

of more complex tertiary structures, allowing a hierarchical and sequential folding of RNA and DNA sequences (Tinoco & Bustamante, 1999).

Another advantage of choosing nucleic acids over proteins for the *in vitro* selection of catalysts is the ease with which nucleic acids can be synthesized, amplified and functionalized. DNA and RNA polymerases as well as reverse transcriptases are widely established tools for the amplification of nucleic acid sequences. Such an enzymatic approach for the amplification of peptide sequences is not available. Chemical synthesis of relatively short fragments of nucleic acids with defined sequences is implemented as solid-phase synthesis using phosphoramidite chemistry and phosphoramidite building blocks derived from protected 2'-deoxynucleosides, ribonucleosides, or chemically modified nucleosides, among others including locked-nucleic acids (LNAs), bridged nucleic acids (BNA) or nucleotides with chemically modified functional groups. This process is completely automated and involves a sequential coupling of the nucleotide building blocks to the growing oligonucleotide chain. The product is then released from the solid phase into the solution, deprotected, and separated from the reactions agents. For the synthesis of randomized oligonucleotide libraries, the automated synthesizer uses a mixture of all four nucleotides instead of coupling one specific nucleotide at a defined position.

Another major advantage of nucleic acid catalysts over protein catalysts, which is of paramount interest for this thesis, is the simplicity by which DNAzymes or ribozymes can be produced that cleave nucleic acid substrates in a sequence specific manner. This fact makes nucleic acids suitable for application as post-transcriptional gene silencing (PTGS) agents. Single-stranded DNA as a *de novo* catalyst for *in vivo* applications has several advantages compared to RNA, including a higher stability in biological fluids (Akhtar *et al.*, 2000) and a more cost-effective production process. In contrast, intracellular production is easily achievable for ribozymes, while DNAzymes require delivery systems. Table 1.1 provides a comparative overview on the advantages and disadvantages of protein enzymes, ribozymes and DNAzymes.

Table 1.1: Comparison of advantages and disadvantages of protein enzymes, ribozymes and DNAzymes.

Catalytic Molecule	Advantages	Disadvantages
Protein	Stabilized transition state	Often require complex cofactors
	Diversity of functional groups	Unstable secondary structure
		In vitro selection from random sequences is unachievable
Ribozymes	*In vitro* selection from random sequences	*In vivo* stability
	Straightforward intracellular production	Reduced chemical diversity of functional groups
DNAzymes	Inexpensive	*In vivo* activity
	In vitro selection from random sequences	Reduced chemical diversity of functional groups
	Easy synthesis	Limited structural information
	Easy to modify	Requires delivery mechanism

1.2 Deoxyribozymes

While natural catalysts include both protein and ribozymes, DNA catalysts have not been discovered in nature (Silverman, 2016). However, the chemical similarities between RNA and DNA led to the identification of synthetic DNA molecules as catalysts. Deoxyribozymes or DNAzymes are single-stranded DNA molecules that are capable of catalyzing chemical reactions due to their ability to adopt three-dimensional structures, which is the prerequisite for their catalytic activity. In this regard, they resemble ribozymes. Compared to RNA, DNA is more stable, slightly easier to synthesize — the chemical synthesis of RNA requires an additional protection group for the 2'-hydroxyl group — and more cost-effective. In 1994, the first single-stranded DNAzyme was identified by *in vitro* selection methods (Breaker & Joyce, 1994). The DNAzyme catalyzes the Pb^{2+}-dependent cleavage of an RNA phosphodiester linkage within its oligonucleotide sequence (*cis*-cleavage), similar to the activity of a ribonuclease (Cuchillo *et al.*, 2011). Since this first report of a DNA catalyst, many laboratories have been working on the identification of DNAzymes with new functionalities and several review articles have been published on this topic within the past decades (for reviews see Breaker, 1997; Hollenstein, 2015; Ma & Liu, 2020; Morrison *et al.*, 2018; Silverman, 2009, 2015, 2016).

1.2.1 RNA-cleaving DNAzymes

This thesis focuses on the most exstensively studied class of deoxyribozymes: the RNA-cleaving DNAzymes. These DNAzymes find application in the laboratory for the manipulation of large RNAs (Pyle *et al.*, 2000), in the field of therapeutics for the treatment of diseases that can be managed by lowering gene expression (Baum & Silverman, 2008), and as biosensors for the detection of metal ions (Ma & Liu, 2020). It is assumed that all known RNA-cleaving DNAzymes catalyze the cleavage of their substrate in a way similiar to the mechanism of the hammerhead

riboyzme: a Brønsted base abstracts a proton from the 2'-OH group of the ribonucleotide that is located in 3' direction of the scissile bond and thereby enables the nucleophilic attack of the thus created oxyanion on the neighboring phosphodiester bond. This nucleophile then performs an in-line attack on the phosphorus atom at the center of the phosphodiester bond, resulting in a pentacoordinated phosphorane intermediate which hydrolyzes into two RNA fragments that are terminated by a 2',3'-cyclic phosphate and a free 5'-hydroxyl group (Breaker *et al.*, 2003a; Santoro & Joyce, 1998) (for further details see Chapter 2 and Figure 2.2). However, in the absence of structural data, the exact mechanism of the cleavage reaction remains unknown. The best studied DNAzymes are the RNA-cleaving 8-17 DNAzyme and 10-23 DNAzyme (Santoro & Joyce, 1997). Their names derive from their *in vitro* selection, since the 8-17 DNAzyme was derived from the 17^{th} clone that was isolated from the pool after eight selection cycles, while the 10-23 DNAzyme was derived from the 23^{rd} clone that was isolated after ten cycles. Both of these DNAzymes have been widely used in the above-mentioned fields including RNA manipulation in the test tube, therapeutics and biosensors. Table 1.2 provides an exemplary summary of RNA-cleaving DNAzymes, their preferred cleavage sites and cofactors as well as values for cleavage rates and catalytic efficiency (k_{cat}/K_m).

Table 1.2: RNA-cleaving DNAzymes in comparison with the hammerhead ribozyme (HHRz), the hairpin ribozyme (HPRz), and ribonuclease (RNase) A. The maximal cleavage rate is either defined as a k_{cat} or k_{obs} value, depending on the value that is given in the corresponding reference. Since cleavage rates of DNAzymes highly depend on buffer and pH conditions, temperature, cofactor concentration, and the dissociation constants of the substrates, these values should therefore only serve as a general estimate.
n.r., not reported.

Name	Cleavage site (5'–3')	Cofactor	rate constant (k)	k_{cat}/K_m min^{-1}/M^{-1}	References
Pb^{2+}-dep.	rAdG	Pb^{2+}	1 min^{-1} at 1 mM Pb^{2+}; pH 7.0	5×10^5	Breaker & Joyce (1994)
8-17	rArG	Mg^{2+}	0.019 min^{-1} at 3 mM Mg^{2+}; pH 7.4	4×10^5	Bonaccio *et al.* (2004); Santoro & Joyce (1997)
10-23	rRrY	Mg^{2+}	3.4 min^{-1} at 50 mM Mg^{2+}; pH 8.0	4.5×10^9	Santoro & Joyce (1997)
Bipartite	rArA	Mg^{2+}	1.4 min^{-1} at 30 mM Mg^{2+}; pH 7.4	6×10^{12}	Feldman & Sen (2001)
Ag10c	rAdG	Ag^{+}	0.41 min^{-1} at 10 µM Ag^{+}; pH 7.5	n.r.	Saran & Liu (2016)
E6	rAdG	Mg^{2+}	0.039 min^{-1} at 10 mM Mg^{2+}; pH 7.4	3×10^3	Breaker & Joyce (1995)
Na8	rAdG	Na^{+}	0.007 min^{-1} at 0.5 M NaCl; pH 7.0	n.r.	Geyer & Sen (1997)
Mg5	rAdG	Ca^{2+}	0.1 min^{-1} at 10 mM Ca^{2+}; pH 7.0	1.6×10^4	Faulhammer & Famulok (1996)
HHRz	rCrG	Mg^{2+}	1.4 min^{-1} at 10 mM Mg^{2+}; pH 7.5	2.9×10^7	Fedor & Uhlenbeck (1992)
HPRz	rGrC	Mg^{2+}	2.1 min^{-1} at 10 mM Mg^{2+}; pH 7.5	7.5×10^7	Hampel & Tritz (1989)
RNase A	poly(C)	Na^{+}	3.1×10^4 min^{-1} at 0.1 M Na^{+}; pH 6.0	9×10^8	delCardayre & Raines (1994)

1.2.2 The 10-23 DNAzyme

The most important and most extensively studied RNA-cleaving DNAzyme is the so-called 10-23 DNAzyme (Santoro & Joyce, 1997). The 10-23 DNAzyme comprises a fixed catalytic core (loop) with 15 nucleotides that is flanked on either site by variable sequences. The variable stretch can be designed to bind any RNA of interest via Watson-Crick base pairing leaving just a single ribonucleotide unpaired (Figure 1.1A). In the presence of divalent metal ions, preferencially Mg^{2+} or Mn^{2+}, the DNAzymes cleaves its specific RNA substrate between a central unpaired purine and paired pyrimidine nucleotide. While the cleavage reaction requires the presence of divalent metal ions, it is not known, whether these divalent metal ions are involved in the actual cleavage reaction of the RNA substrate or whether they play a role in simply stabilizing the complex and therefore prompt the folding of the DNAyzme into an active structure (Joyce, 2001). With regard to the catalytic reaction itself, it is assumed that the catalytic core adopts a three-dimensional structure that allows the cleavage of the phosphodiester bond at the specific cleavage site of the RNA substrate. Several groups have used mutational or deletion studies to investigate the involvement of specific nucleobases, ribose sugars or phosphates in catalysis (Nawrot *et al.*, 2007; Wang *et al.*, 2010a; Zaborowska *et al.*, 2002, 2005). Although these studies provide insights into the importance of distinct functional groups, they cannot compensate for the lack of a high-resolution structure. Two decades after the identification of the 10-23 DNAzyme only one high-resolution X-ray structure with the catalyst in a biologically irrelevant duplex conformation, which is adopted due to a self-complementary sequence within the catalytic core, is available (Nowakowski *et al.*, 1999b). Attempts to obtain the solution structure of the DNAzyme:RNA complex by NMR spectroscopy have also not been successful (Choe *et al.*, 2000).

To disrupt the above mentioned self-complementary sequence and to prevent dimer formation in structural studies, variants of the 10-23 DNAzyme with point mutations within the catalytic core were used (Figure 1.1A). These variants still retain partial activity (Zaborowska *et al.*, 2002). The 10-23 DNAzyme is capable of cleaving RNA, which makes it particulary attractive as a therapeutic agent. Currently, two 10-23 DNAzyme variants are in clinical trials: one DNAzyme for the treatment of basal cell carcinoma (Cho *et al.*, 2013) and the other one for the treatment of asthma (Krug *et al.*, 2015). These DNAzymes are designed to bind and cleave mRNA substrates and thereby reduce the expression level of the related proteins (Figure 1.1C). A detailed review on the 10-23 DNAzyme is given in the following chapter (Chapter 2).

In this thesis, we focus on a variant of the 10-23 DNAzyme that specifically binds to the mRNA encoding the human prion protein (PrP) (Victor *et al.*, 2018). The name prion is derived from **pr**oteinacous **in**fectious particle (Prusiner *et al.*, 1982). Prion diseases are mostly neurodegenerative conditions that can affect both humans and animals. They include chronic wasting disease (CWD) in elk and mule deer (Williams & Young, 1980), scrapie in sheep (Aguzzi, 2006), and bovine spongiform encephalopathy (BSE) in cattle (Wells *et al.*, 1987). In humans, PrP is causal to Creutzfeldt-Jakob disease (CJD) (Creutzfeldt, 1920; Jakob, 1921), a rare, degenerative, invariably fatal brain disorder, which involves a constant autocatalytic conversion of the

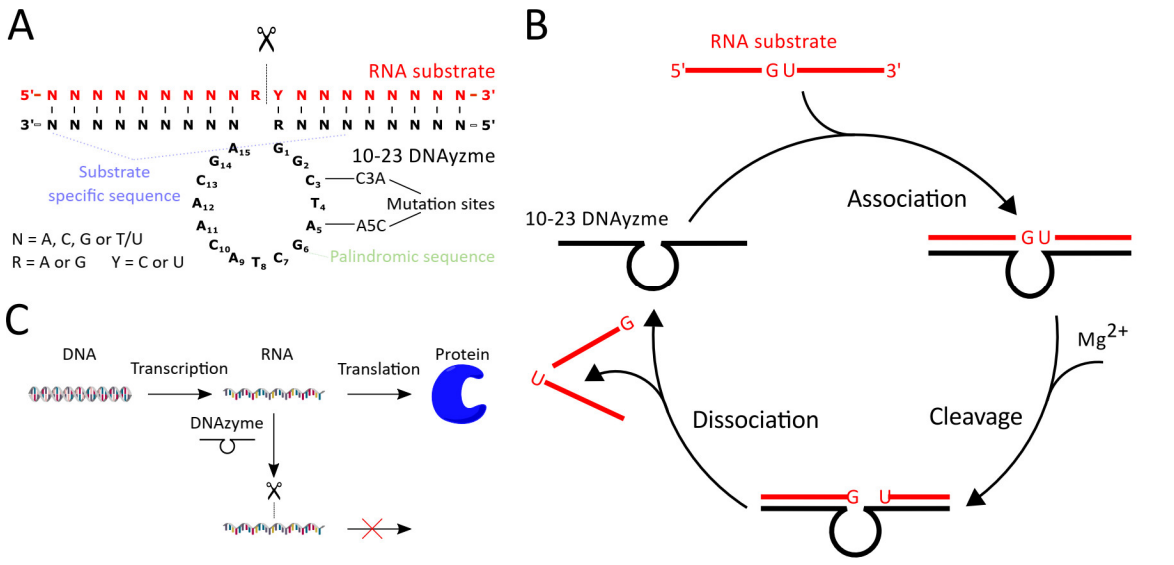


Figure 1.1: Secondary structure and function of the RNA-cleaving 10-23 DNAzyme. (A) Schematic representation of the 10-23 DNAzyme (black) in complex with its RNA substrate (red). The 10-23 DNAzyme comprises as fixed catalytic loop, which is flanked by two variable substrate recognition arms (blue). The self-complementary sequence within the catalytic loop that led to the formation of a catalytically irrelevant crystallization artifact (Nowakowski *et al.*, 1999b) is highlighted in green. The mutations C3A and A5C lead to a partially active DNAzyme (Zaborowska *et al.*, 2002). (B) Reaction scheme of the 10-23 DNAzyme. The RNA substrate hybridizes with the DNAzyme via Watson-Crick base pairing. Cleavage of the RNA is induced by Mg^{2+}. The ternary complex dissociates and the DNAzyme is released to bind another target molecule. Adapted from Victor *et al.* (2018). (C) Proposed effect of the 10-23 DNAzyme on translation of the mRNA and protein levels.

cellular prion protein (PrP^C) to the pathogenic conformational isomer PrP^{Sc} (Sc for scrapie), resulting in the formation of protein aggregates. The accumulation of insoluble aggregates depends on the level of cellular prion protein and therefore on the gene expression and mRNA level in the cell. Thus, decreasing the level of mRNA encoding PrP is considered a promising therapeutic strategy for the treatment of CJD (Creutzfeldt, 1920; Jakob, 1921) and other prion-related diseases (Gajdusek & Zigas, 1957; Gambetti *et al.*, 2008; Mastrianni *et al.*, 1999). This approach was supported by studies with mice devoid of PrP^C, which revealed that the absence of endogenous expression protects the animals from prion infection without showing major side effects due to the lack of PrP^C. After infection with scrapie prions mice devoid of PrP^C showed no symptoms for at least 13 months, whereas the wild-type controls all died within six months. These results support the potential of gene silencing methods for treatment or management of CJD in humans (Büeler *et al.*, 1993). The DNAzyme, which was studied in this work, was designed to cleave the prion protein mRNA at a stretch which is highly accessible for the hybridization of short nucleic acid molecules such as DNAzymes. Those highly accessible cleavage target sites were predicted *in silico* with the use of a sequential folding algorithm as described by Victor *et al.* (2018), based on a previous work by Tafer *et al.* (2008).

To use the 10-23 DNAzyme as a true catalyst, the binding arms need to be designed properly, allowing the association of the RNA substrate to the DNAzyme prior to the M^{2+}-initiated

cleavage reaction via intermolecular Watson-Crick base pairing. After the cleavage reaction, the ternary complex needs to dissociate, so that the DNAzyme can bind to another substrate molecule (Figure 1.1B). Therefore, association and dissociation of the DNAzyme:RNA complex are of great importance for multiple turnover reactions. To be capable of performing multiple turnover reactions at 37 °C, the binding free energy ($\Delta G^0_{37\,^\circ C}$) of each binding arm should have a value of about -8 to -10 kcal/mol (Fokina *et al.*, 2012).

1.2.3 *In vitro* selection of RNA-cleaving DNAzymes

The *in vitro* selection process for RNA-cleaving DNAzymes and artificial ribozymes is very similar to the process for aptamers (Wilson & Szostak, 1999). Aptamers are single-stranded RNA or DNA molecules that can bind proteins or other ligands with a high binding specificity and affinity (Ellington & Szostak, 1990; Tuerk & Gold, 1990). Like aptamers, DNAzymes are selected from a vast library of different molecules with random sequences. For a randomized sequence of length n, the number of possible sequences in the library is 4^n. For methematical reasons, such a library contains 10^{14} to 10^{15} sequences, which is narrowed down to a few sequences with the desired function in a stepwise manner (Silverman, 2016; Tuerk & Gold, 1990).

The principle of *in vitro* selection of nucleic acids with desired function is based on the assumption that a three-dimensional structure, which is required for function, can be formed by many different single-stranded sequences. In concrete terms: in theory, a random pool of sequences with a length of 30 nucleotides (N_{30}) comprises $4^{30} = 10^{18}$ different sequences, covering all the possible structures within the boundaries of the sequence length. One key step in the selection process is the separation of nucleic acids with a desired functionality from those who lack this property and therefore should be excluded from the pool in the next selection round. To facilitate this separation step for the selection of RNA-cleaving DNAzymes, the DNA sequences with the catalytic property can detach themselves from a solid phase.

For the selection of the 10-23 DNAzyme, a library with 10^{14} sequences was used as a starting pool (Santoro & Joyce, 1997, see Figure 1.2). At the 5'-end, each sequence was attached to a biotin label, followed by a short linker sequence, a 12 nucleotides RNA sequence and a randomized DNA sequence of 50 nucleotides. The molecules were attached to a streptavidin-coated solid phase. Cleavage of an RNA phosphodiester bond was triggered by eluting with 10 mM $MgCl_2$ at pH 7.5 and 37 °C. The detached 3' cleavage products, which presumably contain DNA sequences capable of RNA cleavage, were then recovered and amplified by PCR. The recovered sequences were immobilized on the streptavidin solid support and a new selection round was started (Figure 1.2). Because only the DNA sequences with the desired catalytical function reach a new selection round, the library is gradually enriched with active sequences.

Today, several chemical modification for nucleotides are available that alter the function or stability of catalytically active nucleic acids. Various modifications of nucleotides have been described that are compatible with the enzymatic steps of the selection process. These modifications include substitutions at the phosphate/ribose backbone or at the nucleobases (Keefe

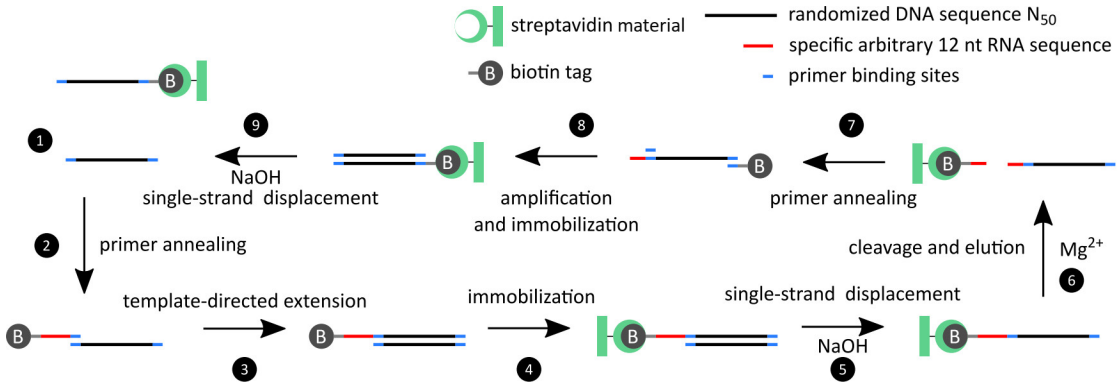


Figure 1.2: *In vitro* selection process for the RNA-cleaving 10-23 DNAzyme. (1) The selection starts from a library of 10^{14} randomized sequences with a length of 50 nucleotides (black). The DNA stretch is flanked by defined primer sites (blue) on each site. (2) The DNA sequences are annealed to a primer containing an RNA region of 12 nucleotides (red) as well as a biotin-tag (gray) on the 5'-end. (3) Template-directed extension is performed enzymatically resulting in double-stranded products. The products are (4) immobilized on a streptavidin column before (5) treatment with NaOH to remove the strand that lacks the biotin-tag. (6) The column is then washed with the reaction buffer containing Mg^{2+}. DNA sequences that are capable of adopting the correct folding for the catalysis of an RNA-cleavage reaction detach themselves from the column. (7) A new set of primers is added to the eluted sequences for amplification. (8) The amplification products are immobilized on a streptavidin column, before (9) treatment with NaOH to release the non-biotinylated strand. The released strand is then used for the next selection round. Adapted from Silverman (2005) and Kumar *et al.* (2019), based on Santoro & Joyce (1997).

& Cload, 2008). For the incorporation of modified nucleotides during the selection process Click-SELEX (Tolle *et al.*, 2015) has been described as a versatile method for the generation of modified nucleic acid libraries by employing copper(I)-catalyzed alkyne–azide cycloaddition (CuAAC) (Liang & Astruc, 2011). CuAAC reactions belong to the class of so-called click reactions, a class of biocompatible small molecule reactions (Kolb *et al.*, 2001). Although the method of Click-SELEX is described for the selection of aptamers, it has high potential for the selection of DNAzymes.

1.3 Applications of RNA-cleaving DNAzymes

Since the first identification of a catalytic DNA by Breaker & Joyce in 1994, DNAzymes became a hot topic in research, far beyond academic interests in the mechanisms of DNA-mediated catalysis. RNA-cleaving DNAzymes are used in different scientific fields including applications in sensing of metal ions in environmental samples (Chen *et al.*, 2016; Saran & Liu, 2016; Zhou *et al.*, 2016), serum (Zhou *et al.*, 2015), and cells (Torabi *et al.*, 2015); intracellular sensing and imaging of RNAs (He *et al.*, 2017); bacterial (Aguirre *et al.*, 2013) and cancer cells (He *et al.*, 2015); and adenosine 5'-triphosphate (Achenbach *et al.*, 2005), gene regulation (Fan *et al.*, 2015); intracellular enzyme activities (Chen *et al.*, 2017a,b); and the manipulation of cell behavior (Li *et al.*, 2018). Since a detailed discussion on all of these applications is beyond the scope of this thesis, only two applications will be discussed in more detail which are central to the presented here: the use of RNA-cleaving DNAzymes as therapeutics for the treatment of

diseases that can be managed by attenuating gene expression on the RNA level (Section 1.3.1) and DNAzymes for the detection of metal ions (Section 1.3.2).

1.3.1 DNAzymes targeting mRNAs

Until today, various oligonucleotide-based pharmatheutics have been tested for gene therapy, including anti-sense oligonucleotides (ASOs) (Paz-Ares *et al.*, 2006; Yacyshyn *et al.*, 2002), siRNAs (Kim *et al.*, 2004; Reich *et al.*, 2003; Tolentino *et al.*, 2004), ribozymes (Weng *et al.*, 2005; Wong-Staal *et al.*, 1998) and DNAzymes (Cho *et al.*, 2013; Krug *et al.*, 2015). For all of these approaches, the main targets are mRNA sequences, and the key steps involve mRNA recognition and in several cases, mRNA cleavage. ASOs are single stranded DNA sequences with a length of 15 to 25 nucleotides that are designed to bind to a specific target mRNA sequence via base pairing. Thereby the translation into protein is blocked due to hindrance of ribosomal movement along the transcript, or by activation of endogenous ribonuclease (RNase) H which degrades the mRNA (Bhindi *et al.*, 2007; Kher *et al.*, 2011). Although the concept is quite simple, it is not very efficient. Unmodified ASOs are prone to degradation by nucleases and their delivery into cells is not trivial due to their negative charge (Bhindi *et al.*, 2007).

Another approach to down-regulate mRNA levels in cells is the use of siRNAs. The siRNAs are double-stranded RNA molecules with two nucleotides overhangs at each 3'-end that are incorporated into a multiprotein complexes, known as RNA-induced silencing complex (RISC). In the presence of ATP, the siRNAs undergo unwinding by RNA helicases, while being processed into the RISC. During the formation of the RISC, the sense strand of the double-stranded siRNA is degraded (Leuschner *et al.*, 2006), and the antisense strand guides the RISC to the complementary target mRNA, which is finally cleaved by Argonaute enzymes associated with the RISC (Kher *et al.*, 2011).

The use of ribozymes and DNAzymes may offer several advantages over other knockdown techniques. Ribozymes as well as DNAzymes are self-sufficient biocatalysts that do not rely on the presence of other biomolecules like Argonaute proteins and RNase H. In addition, DNA is inherently more stable, cost-effective and easier to synthesize or to label compared to RNA-based agents. Although a plethora of reports on the use of DNAzymes as therapeutic agents is available, so far only two DNAzymes are reported to have entered clinical trials: (i) a DNAzyme for the treatment of basal cell carcinoma (Cho *et al.*, 2013) and (ii) a DNAzyme for the treatment of asthma (Krug *et al.*, 2015). Since the discovery of the RNA-cleaving DNAzymes, attempts have been made to exploit them for therapeutic purposes. The first motivation to develop RNA-cleaving DNAzymes was to target viral RNA. That is why the 8-17 as well as the 10-23 DNAzyme were designed to target human immunodeficiency virus (HIV)-derived RNA (Santoro & Joyce, 1997). Today, DNAzymes have been designed for a broad variety of different substrates, including bacterial (Chen *et al.*, 2004; Li *et al.*, 2005) and viral (Trepanier *et al.*, 2008; Unwalla & Banerjea, 2001), cancer-related (Wu *et al.*, 1999), allergy/asthma-related (Krug *et al.*, 2015), cardiovascular-related (Xiang *et al.*, 2005) and prion-related (Victor *et al.*, 2018) RNAs. Since the number of reports on DNAzymes targeting different disease-associated

RNAs is immense, this thesis does not aim at reviewing them one-by-one. Interested readers are guided to a number of review articles dealing with therapeutic applications of RNA-cleaving DNAzymes (Bhindi *et al.*, 2007; Fokina *et al.*, 2015; Fu & Sun, 2015; Zhou *et al.*, 2017a). Instead, this thesis will point out the discrepancy between the many promising reported applications and the absence of DNAzyme-based drugs on the actual pharmaceutical market.

When it comes to the design of an RNA-cleaving DNAzyme, first the selection of an appropriate cleavage site has to be considered. DNAzymes cleave their RNA substrates at specific cleavage sites. For example, the 10-23 DNAzyme catalyzes the cleavage of its RNA substrate between an unpaired purine ribonucleotide and a paired pyrimidine ribonucleotide. If only GU sequences are taken into account, the transcript variant 1 of the human PrP mRNA (the 10-23 DNAzyme variant studied in this thesis is specific for the mRNA of the human PrP) contains 143 possible cleavage sites (Victor *et al.*, 2018). However, due to secondary structures of the mRNA transcript only a few of them will be actually accessible for the hybridization of the DNAzyme. This is why Victor *et al.*, who designed the 10-23 DNAzyme studied in this thesis, used a sequential folding algorithm (Bernhart *et al.*, 2006a; Tafer *et al.*, 2008) to rank the potential GU cleavage sites within the target mRNA based on the probability that regions with a length of 10 nucleotides with a centered GU sequence are unpaired. In addition to the accessibility of the cleavage site, the stability of the DNAzyme:RNA complex during the catalytic cycle (Figure 1.1B) needs to be taken into account, since the association and dissociation are of great importance for multiple-turnover kinetics. These steps depend both on the length and the nucleotide composition of the target recognition arms: if the stability of the complex is too low, association of the RNA and the DNAzyme is hampered, whereas a too high stability prevents the dissociation of the cleavage products and therefore blocks the DNAzyme for binding to another substrate RNA. According to Fokina *et al.* (2012), for DNAzymes that perform catalysis inside the human body, each binding arm should naturally have a melting temperature (T_m) of 37 °C, which is in accordance with a binding free energy of $\Delta G^0_{37\,^\circ C} = -8$ to -10 kcal/mol for each arm. Using this sequencial folding algorithm by Tafer *et al.* (2008) and taking into account the stability of the pre- and post-catalytic DNAzyme:RNA complex Victor *et al.* identified a variant of the 10-23 DNAzyme specific for the mRNA of the human PrP that efficiently cleaves structured *in vitro* transcripts. This DNAzyme is termed Dz839 (Victor *et al.*, 2018).

Strategies to enhance complex stability and to compete with internal structures for mRNA binding involve the modification of bases within the substrate recognition arms of the DNAzyme with 2'-O-methyl groups or LNAs (Schubert *et al.*, 2004; Vester *et al.*, 2002). LNA bases comprise a methylene bridge that connects the 2'-O with the 4-C, resulting in a locked C3'-endo conformation (Koshkin & Wengel, 1998; Wengel, 1999), reducing the conformational flexibility of the ribose which leads to an increased local organization and therefore in increasing the affinity for complememtary sequences (Braasch & Corey, 2001).

Besides the delivery of DNAzymes to the cell or its target tissue—a challenge that will not be discussed in this thesis—the achievement of *in vivo* stability of DNAzymes is of great importance for the generation of an efficient therapeutic. Due to the lack of the 2'-hydroxyl group of the ribose sugar DNA is stable over a large pH range. However, single-stranded DNA is

prone to degradation from 3'→5' exonucleases. Several chemical modifications have been reported to stabilize DNAzymes or single-stranded anti-sense DNA molecules against nucleolytic degradation, including 2'-O-methyl (Fokina *et al.*, 2012), LNAs (Kurreck *et al.*, 2002b), inverted thymidines at the 3'-end (Dass *et al.*, 2002), and phosphorothioate modifications of the backbone (Lu *et al.*, 2005) (Figure 1.3). Introduction of a 3'-3'-linked inverted thymidine nucleotide to the 3' end of the DNAzyme creates a sequence with two 5' ends which is protected from nucleolytic degradation by 3'→5' exonucleases (Ortigão *et al.*, 1992). Phosphorothioate modifications are widely used, although they are known to decrease substrate affinity, have toxic side effects and provoke an immunological response (Wahlestedt *et al.*, 2000).

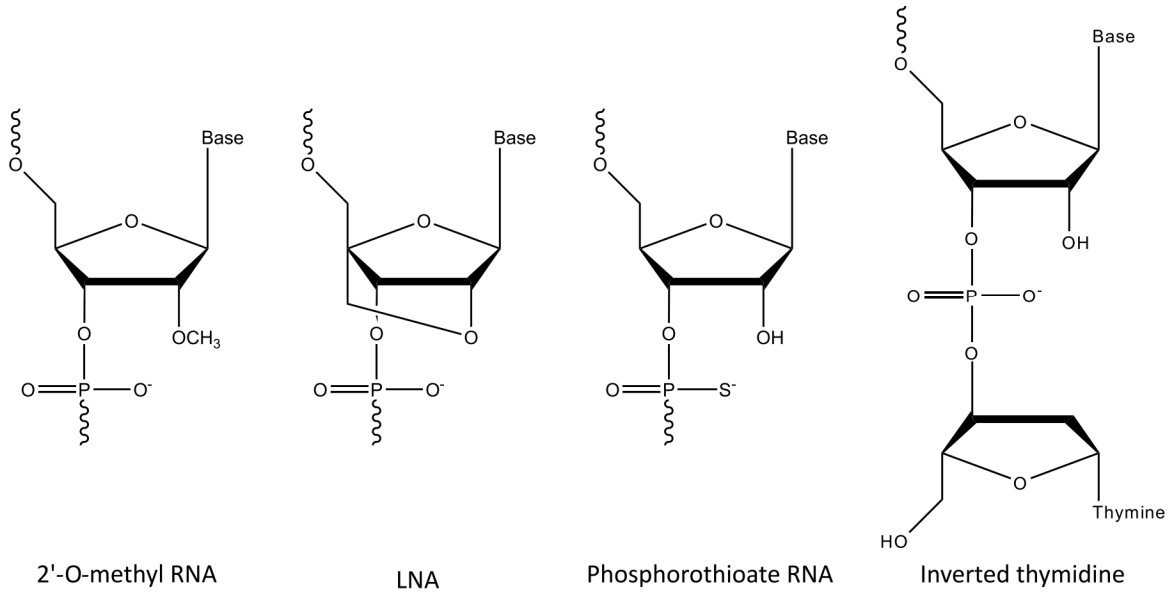


Figure 1.3: RNA modifications reported to enhance DNAzyme stability against nucleolytic degradation.

The major concern that arises with RNA-cleaving DNAzymes for therapeutic applications is their *in vivo* activity. Most of the identified DNAzymes require high concentrations of metal ions for their activity, which exceed intracellular concentrations. For example, standard protocols for activity assays on the 10-23 DNAzyme involve concentrations between 10 mM and 50 mM $MgCl_2$, while the intracellular concentration of free Mg^{2+} is about 1 mM (Pechlaner & Sigel, 2012). Under these conditions, activity of the 10-23 DNAzyme is very low. Young *et al.* (2010), using caged DNAzymes for photochemical regulation of DNAzyme activity and gene expression, concluded that the observed effects on gene expression are the result of antisense effects due to the hybridization of the DNAzyme with its complementary RNA counterpart and not a result of actual catalytic DNAzyme activity. However, medical literature reporting on successful usage of DNAzymes for mRNA cleavage to treat asthma (Krug *et al.*, 2015) and cancer

(Cho *et al.*, 2013), raised the prospect that the 10-23 DNAzyme could serve as a therapeutic agent, although control experiments excluding that the observed effects are due to antisense mechanism have not often been described. Control experiments may involve usage of inactive DNAzymes or cells lacking RNase H. Wang *et al.* (2015a) studied both antisense and catalytic effects of the 8-17 DNAzyme on mRNA extracts by comparing the effects of active and inactive DNAzymes on mRNA levels. They concluded that the role of the DNAzyme depends on its relative activity: in the presence of Mg^{2+} the antisense effect dominates, while in the presence of Zn^{2+} cleavage activity dominates. This observation was attributed to the preferential use of Zn^{2+} by the 8-17 DNAzyme. However, these studies were not performed in cells and the *in vivo* relevance is questionable due to the low abundance of free Zn^{2+} in cells, i. e. the concentration of free Zn^{2+} is in the pM range (Maret, 2017). Nevertheless, these results point towards the complex regulation of DNAzyme activity. A very elegant approach to improve the *in vivo* activity of DNAzymes has been reported by Fan *et al.* (2015), who report a DNAzyme-MnO_2 nanosystem for effective gene-silencing. Chlorin-e6-labeled 10-23 DNAzymes are adsorbed to MnO_2 nanosheets which protect the DNAzymes from enzymatic degradation and facilitate efficient uptake by the cells. Inside the cells, MnO_2 is reduced to Mn^{2+} ions by intracellular glutathione. Mn^{2+} then serves as a cofactor for the DNAzyme.

Despite the low abundance of divalent metal ions in cells, complexity of intracellular concentrations of both mono- and divalent metal ions, molecular crowding effects (Nakano *et al.*, 2014) and interference with proteins should be taken into account for the evaluation of *in vivo* competence of DNAzymes, since they are usually selected and studied under diluted buffer conditions with defined metal ion concentrations. In addition, complementary methods can be used to quantify an effect on gene expression: (i) Western blot analysis to compare protein levels in the presence and absence of DNAzymes, (ii) quantitative polymerase chain reaction (qPCR) to determine mRNA levels, (iii) reporter assays, for example with luciferase fusion-proteins to obtain a luminescence readout, (iv) the use of DNAzymes targeting the mRNA of the green fluorescent protein (GFP) to obtain a fluorescence readout, or (v) analysis of the cells in the presence and absence of the DNAzyme based on their phenotype.

In order to make RNA-cleaving DNAzymes amenable for the application as therapeutic agents an enhanced *in vivo* activity is required. Therefore, we need a fundamental understanding of the reaction mechanism. Structural information would contribute to understanding the function of the DNAzyme and the metal ion cofator and allow for designing optimized catalysts with modifications that compensate for the low intracellular Mg^{2+} concentration. Since the DNAzyme changes its structure due to the RNA clavage process in order to perform its task, as a consequence also metal ion coordination changes during this process. This makes an understanding of the role of the metal ion even more challenging.

1.3.2 DNAzymes as biosensors for metal ions

Metal ions play a major role in many biological and chemical processes. That is why understanding their distribution and fluctuation in living cells is of broad interest in the fields of cell

signaling, enzyme catalysis or medical research. The detection of metal ions in environmetal samples becomes more and more important, especially regarding serious man-made environmental and health problems caused by mining, industrial emissions, or waste dumping (Järup, 2003). However, highly sensitive analytical standard methods for the detection of metal ions such as mass spectrometry or atomic absorption/emission spectroscopy require costly and large instruments and a pre-treatment of the samples. Here, metal-detecting biosensors provide a practical and more cost-effective alternative which allows a user to perform real-time and direct *in situ* measurements. Furthermore, biosensors can compete with chemical metal sensors, which are usually obtained by the rational design of fluorescent chelators (Jiang & Guo, 2004; Kim *et al.*, 2012a). Regarding the field of biosensors, nucleic acid-based sensors provide many advantages over antibody-, protein-, or peptide-based approaches. Peptide-based sensors (Cheng *et al.*, 1996; Wegner *et al.*, 2007) as well as antibodies usually work best under physiological conditions, while environmental samples may provide conditions that are not within this optimal range and may even cause irreversible degradation of the detection module. Usually, nucleic acid-based biosensors can be denatured and renatured many times without loosing their binding affinity. Furthermore, antibodies require the chelation of metal ions prior to binding, since metal ions are too small for a direct interaction (Reardan *et al.*, 1985).

While research in the field of DNAzymes started with the aim of cleaving RNA, during the last years the work on using DNAzymes as biosensors has advanced significantly. Given their high affinity, stability, and enhanced selectivity, the use of DNAzymes enables the detection of numerous analytes such as metal ions (reviewed in Gong *et al.*, 2015; Zhang *et al.*, 2011b; Zhou *et al.*, 2017c). The development of biosensors involves the integration of signal transducers or reporter molecules that transform the recognition event to a measurable signal. In the last two decades, RNA-cleaving DNAzymes, so-called catalytic beacons, have been studied for the development of different signaling mechanisms involving labeled (Torabi *et al.*, 2015) or non-labeled fluorescent (Fan *et al.*, 2012), colorimetric (Tian *et al.*, 2015), or electrochemical (Shen *et al.*, 2008) sensors. Furthermore, a lead-detecting biosensor has been reported that combines the specifity of DNAzymes with the sensitivity of surface-enhanced Raman scattering (SERS) (Wang & Irudayaraj, 2011).

Among the different optical sensing methods, fluorescence-based techniques are the most widely used, since they provide high sensitivity, selectivity, and reproducibility, as well as a broad linear range of detection, and an easy operational procedure (Kumar *et al.*, 2019). Fluorescence is monitored as a function of time, and the rate increase is proportional to the metal (analyte) concentration. However, incorporation of fluorophores into the DNAzyme is essential to obtain the desired properties, since nucleic acids lack intrinsic fluorescence. Fluorescence-based biosensors can be divided into labeled and non-labeled sensors. Labeled fluorescence-based sensors have the fluorophore molecule covalently attached to the DNAzyme. For example, the initially reported Pb^{2+}-detecting DNAzyme-based biosensor had a fluorophore attached to one end of the substrate and a corresponding fluorescence quencher molecule attached to the DNAzyme (Li *et al.*, 2000) (Figure 1.4A). In the absence of the analyte, i. e. metal ion, the substrate is bound to the DNAzyme and the fluorescence signal is quenched due to the close prox-

imity of the fluorophore and the quencher molecule. This principle is known as fluorescence resonance engery transfer (FRET). Thus, substrate cleavage in the presence of the analyte (Pb^{2+}) can be measured by an increasing fluorescence signal. However, although the above described arrangement of fluorophore and quencher molecule results in a highly sensitive and selective sensor at 4 °C, a high background fluorescence signal is observed at higher temperatures. To avoid this disadvantage the position of the fluorophore and quencher molecule can be varied (Liu & Lu, 2006). To reduce background fluorescence derived from non-hybridized substrate molecules, the attachment of an additional quencher molecule at the 5' end of the substrate has been successfully introduced (Liu & Lu, 2003b) (Figure 1.4B). Furthermore, it is possible to link both the fluorophore and quencher molecule to the two ends of the substrate (Zhang *et al.*, 2010b) (Figure 1.4C) or next to the cleavage site (Chiuman & Li, 2007) (Figure 1.4D). The latter case is of interest especially for monitoring fast kinetics. FRET-based biosensors can also be used for the real-time detection of analytes, for example to visualize the dynamics of metal ion distributions in cells (Torabi *et al.*, 2015).

The principle of non-labeled fluorescence-based sensors relies on the principle that DNA intercalating dyes such as SYBR Green I (Zhang *et al.*, 2013) or picogreen (Zhang *et al.*, 2011a) bind double-stranded nucleic acid helices with a high fluorescence intensity. Upon substrate cleavage in the presence of the analyte, the double helices formed by the substrate recognition arms of the DNAzyme and the substrate unwind due to dissociation of the cleavage products. This event results in an decrease in the fluorescence signal. Compared to labeled fluorescence-based catalytic beacons, non-labeled or label-free sensors are generally more straightforward and cost-effective, since they do not require modified DNA sequences (Kumar *et al.*, 2019).

The principle of colorimetric sensors that function as catalytic beacons relies on the detection of a change in color that is induced by the metal-dependent cleavage event. For example, gold nanoparticles (AuNPs) can be used to achieve colorimetric detection. A Pb^{2+}-specific biosensor was reported which is based on DNAzyme-directed assembly of gold nanoparticles. The biosensor consists of a DNAzyme and its substrate that can hybridize to a 5'-thio-modified DNA attached to gold nanoparticles. Upon hybridization, the gold nanoparticles are brought together, resulting in a blue color of the AuNPs. In the presence of Pb^{2+} the specific substrate is cleaved by the DNAzyme, which prevents the formation of the AuNP assembly, resulting in red nanoparticles (Liu & Lu, 2003a).

An electrochemical sensor for the detection of Pb^{2+} has been achieved by labeling a DNA catalyst with methylene blue (MB), hybridizing the DNAzyme to its substrate, and immobilizing the DNAzyme on an electrode surface via a thiol group on the 5' end of the DNAzyme. Due to the rigid conformation of the DNAzyme:substrate complex, MB is kept at a certain distance from the electrode so that no electron (e^-) transfer can occur. After Pb^{2+}-induced cleavage, the cleavage products dissociate from the DNAzyme and the DNAzyme gains the flexibility that is required for e^- transfer. Sensor activity in the presence of Pb^{2+} is conducted using altering current voltammetry (Xiao *et al.*, 2007) with a signal intensity that is dependent of the present metal ion concentration (Figure 1.5A). A biosensor with SERS activity based on an RNA-cleaving DNAzyme has been reported for the detection of Pb^{2+} (Wang & Irudayaraj, 2011).

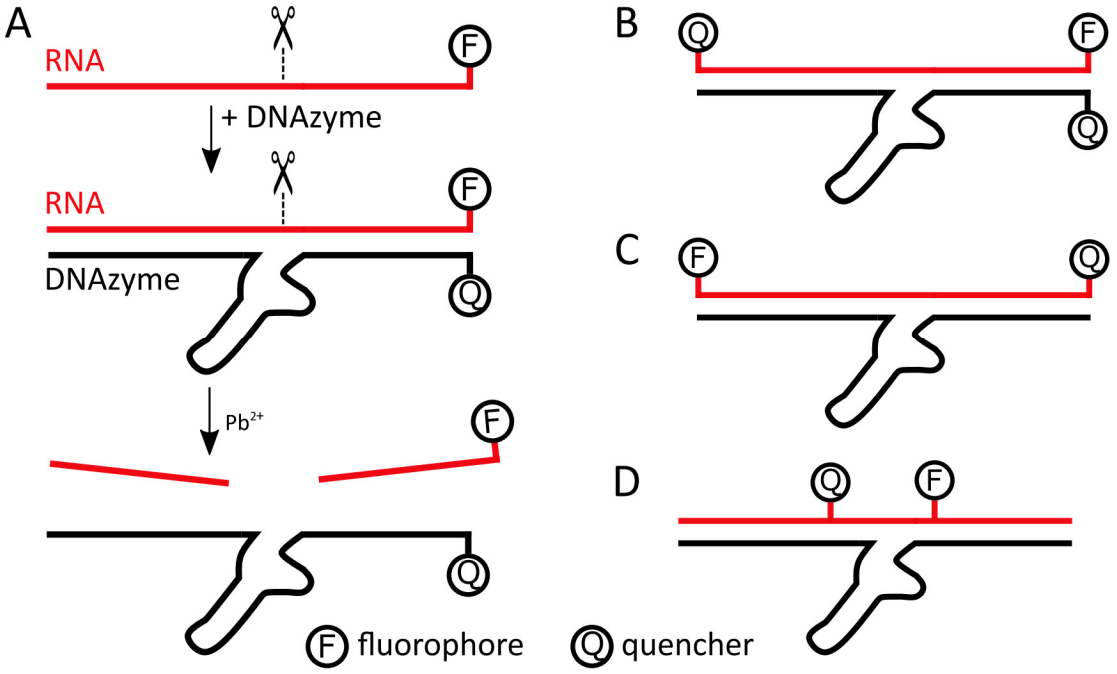


Figure 1.4: Schematic representation of fluorescence-labeled biosensors for lead detection based on the RNA-cleaving 8-17 DNAzyme. (A) The initially reported Pb^{2+}-detecting DNAzyme-based biosensor has a fluorophore attached to one end of the substrate (red). A fluorescence quencher molecule is attached to the corresponding site of the DNAzyme (black) (Li *et al.*, 2000). Pb^{2+}-induced cleavage leads to an increase in the fluorescence signal, since fluorophore and quencher are separated due to the dissociation of the cleavage products. (B) To reduce background fluorescence derived from non-hybridized substrate molecules, an additional quencher molecule can be attached at the 5' end of the substrate (Liu & Lu, 2003b). Linking both the fluorophore and the quencher molecule to the substrate is possible either (C) on both ends (Zhang *et al.*, 2010b) or (D) next to the cleavage site to measure Pb^{2+}-induced RNA cleavage (Chiuman & Li, 2007).

While the above described biosensors mostly find application in environmental samples for the detection of toxic metal ions, RNA-cleaving DNAzyme-based biosensors can also be used to measure intracellular metal ion concentrations. Within this context, so-called caged DNAzymes that have a photolabile group covalently attached to prevent its activity have been proven to be beneficial, since they allow for a controlled activation of the DNAzyme. The caging group can be detached by irradiation with light of a defined wavelength (Young *et al.*, 2010). Caging groups can be introduced into RNA-cleaving DNAzymes either at conserved nucleobases within the catalytic loop (Ting *et al.*, 2004), within nucleotides in the substrate (Lusic *et al.*, 2007), or at the 2'-OH group at the cleavage site to prevent cleavage of the sensor until decaging has taken place (Hwang *et al.*, 2014). Such a caged substrate can be used for intracellular Na^+ sensing (Torabi *et al.*, 2015) (Figure 1.5B). In addition to intracellular sensing, DNAzymes have been used for the detection of metal ions in blood and serum samples. Variants of the 8-17 DNAzyme have been selected and tested in serum for their ability to detect Ca^{2+} and Mg^{2+} (Zhou *et al.*, 2015, 2016).

In 2000, the 17E DNAzyme which was previously selected for the Zn^{2+}-dependent cleavage of RNA (Li *et al.*, 2000) was the first DNAzyme reported for the detection of Pb^{2+} (Li & Lu,

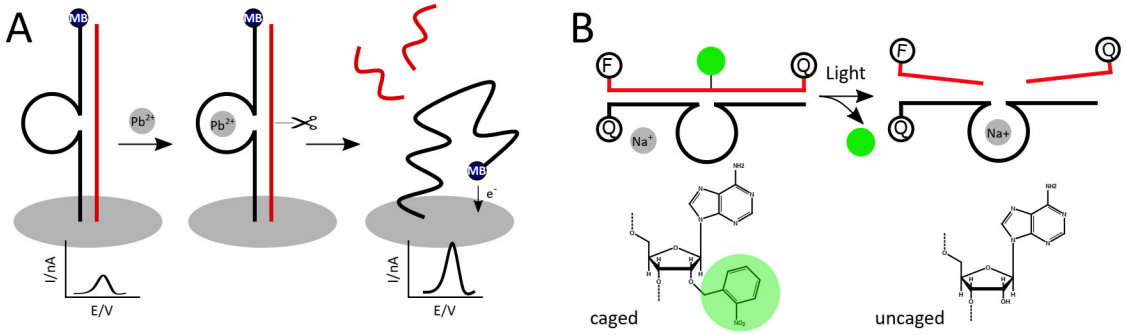


Figure 1.5: DNAzyme-based biosensors. (A) The electrochemical DNAzyme-based biosensor developed by Xiao *et al.* (2007) for the detection of Pb^{2+}. The sensor consists of a methylene-blue (MB) labeled DNAzyme (black) that is hybridized to its substrate (red). The DNAzyme:substrate complex is immobilized on a gold electrode via a thiol group on the 5'-end of the DNAzyme. Due to the rigid conformation of the complex, MB is kept at a certain distance from the electrode so that no electron (e^-) transfer can occur. Pb^{2+} induces DNAzyme-mediated cleavage of the substrate. The cleaved substrate dissociates from the DNAzyme and allows MB to transfer electrons. Sensor measurements are conducted using alternating current voltammetry. (B) Schematic representation of a decaging process for the DNAzyme for intracellular Na^+ detection (Torabi *et al.*, 2015). The chimeric DNA/RNA substrate (red) of the Na^+-specific DNAzyme (black) is labeled with a fluorophore at the 5' end and a quencher molecule at the 3' end. An additional quencher molecule is attached to the 3' end of the DNAzyme. To prevent cleavage of the substrate during the delivery of the DNAzyme into the cell and to allow controlled activation, the 2' OH group at the cleavage site is modified with a photolabile o-nitrobenzyl group (highlighted in green). Irradiation at 365 nm removes the caging group, leading to a cleavable substrate and an active biosensor.

2000). Since this first report, numerous DNAzymes have been selected for the detection of various metal ions. Until today, biosensors based on RNA-cleavage have been developed for Na^+ (Torabi *et al.*, 2015), Pb^{2+} (Li & Lu, 2000), Ca^{2+} (Zhou *et al.*, 2017b), Cd^{2+} (Huang & Liu, 2015), Zn^{2+} (Li *et al.*, 2015; Wang *et al.*, 2016), UO_2^{2+} (Liu *et al.*, 2007), Hg^{2+} (Hollenstein *et al.*, 2008), Ag^+ (Saran & Liu, 2016), Cu^{2+} (Huang & Liu, 2016), Tl^{3+} (Huang *et al.*, 2015), Cr^{3+} (Zhou *et al.*, 2016), and different lanthanoides (Huang *et al.*, 2014, 2016). However, most of these DNAzymes are only relevant for the analysis of environmental (water) samples, while their physiological relevance is limited, due to the low abundance of most of these above mentioned metal ions in living systems.

Although selection of RNA-cleaving DNAzymes is usually carried out in the presence of metal ions to achieve specificity, in the absence of structural data, the role of the metal ions during the catalytic process remains unknown. It is assumed that the metal ions perform their catalytic activity by transient interactions with the DNAzyme rather than binding to an aptamer-like binding pocket, making selectivity for one specific metal ion difficult to achieve. For example, the 10-23 DNAzyme has been selected in the presence of Mg^{2+}, but it shows highest activity in the presence of Mn^{2+} and is still active in the presence of Pb^{2+}, Ca^{2+}, Cd^{2+}, Sr^{2+}, Ba^{2+}, Zn^{2+}, and Co^{2+} (Santoro & Joyce, 1998). This broad tolerance towards different metal ions is also true for other RNA-cleaving DNAzymes including DNAzymes sharing the 8-17 DNAzyme motif (Faulhammer & Famulok, 1996; Li & Lu, 2000; Santoro & Joyce, 1997). Furthermore, inhibitory effects of other components in an analytical sample need to be considered (see Chapter 3). Thus, DNAzyme-based biosensors that rely on the cleavage of an RNA substrate, until now,

have only been established for a few elements of the periodic table. To obtain DNAzyme-based biosensors with higher affinity and specificity an deeper understanding of the metal ion within the cleavage reaction is of paramount importance.

1.4 Reaction scope of DNA-mediated catalysis

As mentioned before, the chemical repertoire of DNAzymes is not restricted to RNA cleavage reactions. Since the first report of a DNA catalyst, a myriad of DNAzymes with different catalytic activities have been reported. In addition,new techniques for the the identification of catalytically active DNA sequences have been developed and improved. These new DNAzymes include both catalysts with oligonucleotide and non-oligonucleotide substrates. DNAzymes with oligonucleotide substrates for example perform the ligation (Cuenoud & Szostak, 1995) or cleavage (Chandra *et al.*, 2009; Gu *et al.*, 2013) of DNA or the ligation of RNA strands (Flynn-Charlebois *et al.*, 2003). DNAzymes with non-oligonucleotide substrates include catalysts which are capable of performing the photoreversion of thmyine dimers (Chinnapen & Sen, 2004), peroxide reduction mimicking peroxidase activity (Travascio *et al.*, 1998), and peptide side chain and backbone modifications (Chandrasekar & Silverman, 2013; Chandrasekar *et al.*, 2015; Chu *et al.*, 2014; Pradeepkumar *et al.*, 2008; Silverman, 2015; Walsh *et al.*, 2015, 2013; Wong *et al.*, 2011).

With regard to applications as biosensors, the horseradish peroxidase (HRP) metallo-enzyme mimicking DNAzyme with a G-rich sequence as the recognition element for the detection of potassium ions is worth mentioning (Fan *et al.*, 2012). Upon binding of K^+, the G-rich element folds into a G-quadruplex structure, allowing a hemin cofactor to bind, resulting in the formation of the HRP-DNAzyme. In the presence of the DNAzyme, 3-(p-Hydroxyphenyl)-propanoic acid (HPPA) can be oxidized by H_2O_2, resulting in oxidative fluorescent product. The fluorescence signal increases with increasing K^+ concentrations. Further G-quadruplex-based biosensors are for example reported for the detection of miRNA in the presence of hemin as a cofactor (Li *et al.*, 2017b), ATP in the presence of peptide/Au nanocomposites, hemin and K^+ (Li *et al.*, 2014a), or choline in the presence of 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS) (Nikzad & Karami, 2018). A review article on the reaction scope and repertoire of DNAzymes was published by Silverman (2016).

1.5 Aim of this thesis

RNA-cleaving DNAzymes have been extensively studied for their potential to serve as therapeutics for the treatment of diseases which are associated with high RNA levels (Bhindi *et al.*, 2007) or for their suitability to serve as biosensors for metal ions or other small analytes (Gong *et al.*, 2015). Among the large number of RNA-cleaving DNAzymes, the 10-23 DNAzyme is the most extensively studied RNA-cleaving DNAzyme, which also entered clinical trials for the treatment of asthma (Krug *et al.*, 2015) and cancer (Cho *et al.*, 2013). However, one major drawback of the 10-23 DNAzyme for *in vivo* application is its requirement for a Mg^{2+} concentration that exceed the level of free Mg^{2+} ions in cells (Victor *et al.*, 2018). *In vitro*, the 10-23 DNAzyme shows activity in the presence of different divalent metal ions, including Mg^{2+}, Mn^{2+}, or Pb^{2+} (Santoro & Joyce, 1998) which hampers its application as a metal-detecting biosensor.

So far, the lack of structural information on DNAzymes in general, and the 10-23 DNAzyme in particular, has prevented a detailed understanding of the mechanism behind the RNA-cleavage reaction. More than two decades after the discovery of DNAzymes only three high-resolution X-ray structures are available: one crystal structure was obtained of the 10-23 DNAzyme in a biologically irrelevant duplex conformation, which is adopted due to a self-complementary sequence within the catalytic core (Nowakowski *et al.*, 1999b). Seventeen years later, the first crystal structure of a DNA catalyst in a biologically relevant conformation was reported, however, the structure of the RNA-ligating DNAzyme 9DB1 represents a post-catalytic state (Ponce-Salvatierra *et al.*, 2016). In 2017, the crystal structure of the RNA-cleaving 8-17 DNAzyme in the presence of a non-native chimeric DNA/RNA substrate was published (Liu *et al.*, 2017).

The objective of the following work is firstly to report on the numerous functional studies on the RNA-cleaving 10-23 DNAzyme and to provide a comprehensive overview of the nucleotide modifications within the catalytic loop that have an impact on DNAzyme function as well as of the impact of different metal ions on the DNAzyme activity. The compiled data are intended to serve as a starting point for future attempts at elucidating the mechanism of DNAzyme-mediated RNA cleavage. This review aims at highlighting the molecular features that a structural model should address and provides information that could facilitate the identification and localization of different metal ion binding sites.

The second objective of this thesis is to explore the properties of the 10-23 DNAzyme with special regard to its metal ion dependency in solution using Förster resonance energy transfer (FRET)-based cleavage assays, isothermal titration calorimetry (ITC), nuclear magnetic resonance (NMR) spectroscopy, and electron paramagnetic resonance (EPR) spectroscopy. Here, the influence of both mono- and divalent metal ions on metal binding and DNAzyme activity will be explored with the aim to shed light on the reasons for the poor performance the 10-23 DNAzyme *in vivo* and provide insights that are of great interest with regard to the use of RNA-cleaving DNAzymes as biosensors for metal ions and as therapeutics.

High-resolution structural data are crucial for understanding the mechanism by which the 10-23 DNAzyme performs RNA-cleavage. However, the preparation of well-ordered nucleic acid crystals that are suitable for structure determination by X-ray crystallography is particu-

larly challenging, since the surfaces of RNA and DNA molecules are dominated by negatively charged residues and regularly ordered phosphate groups, which frequently results in crystals with a poor long-range order (Ferré-D'Amaré & Doudna, 2000a). That is, as a third objective of this work, a protocol will be established that enables the fast and efficient crystallization of nucleic acids by soaking. Optimized variants of the RNA-binding protein U1A will be used, which enable the detection of RNA-binding *in crystallo* based on a change in fluorescence.

In addition, different strategies for the crystallization of the 10-23 DNAzyme in complex with its native RNA substrate will be presented that all aim at obtaining high-quality crystals. These strategies involve (i) co-crystallization with the RNA-binding protein U1A, (ii) a combinatorial screen of the DNAzyme:RNA complex with varying length of the RNA substrate and the binding arms of DNAzyme resulting in duplexes of varying length with and without overhangs, (iii) use of a reduced DNAzyme with nucleotide deletions within the flexible catalytic core region, and (iv) tetrahedral nanostructures consisting of three DNAzyme molecules and their RNA substrates.

All of these approaches are not restricted to the nucleic acid sequences used for this thesis, but they offer a prospect for functional studies on and the structure elucidation of various nucleic acid sequences, including ribozymes, riboswitches, DNAzymes and DNAzyme:RNA complexes.

2 Molecular features and metal ions that influence 10-23 DNAzyme activity

This chapter reflects content of the following publication.

2.1 Publication information

Hannah Rosenbach, Julian Victor, Manuel Etzkorn, Gerhard Steger, Detlev Riesner, and Ingrid Span

Submitted to: **Molecules**, 31.05.2020

2.2 Abstract

Deoxyribozymes (DNAzymes) with RNA hydrolysis activity have a tremendous potential as gene suppression agents for therapeutic applications. The most extensively studied representative is the 10-23 DNAzyme consisting of a catalytic loop and two substrate binding arms that can be designed to bind and cleave the RNA sequence of interest. The RNA substrate is cleaved between central purine and pyrimidine nucleotides. The activity of this DNAzyme *in vitro* is considerably higher than *in vivo*, which was suggested to be related to its divalent cation dependency. Understanding the mechanism of DNAzyme catalysis is hindered by the absence of structural information, however, numerous biological studies provide comprehensive insights into the role of particular deoxynucleotides and functional groups in DNAzymes. Here we provide an overview of the thermodynamic properties, the impact of nucleobase modifications within the catalytic loop, and the role of different metal ions in catalysis. We highlight the critical molecular features that will enable to develop improved strategies for structure determination and to understand the mechanism of the 10-23 DNAzyme. These insights provide the basis for improving activity in cells and pave the way for developing DNAzyme applications.

2.3 Introduction

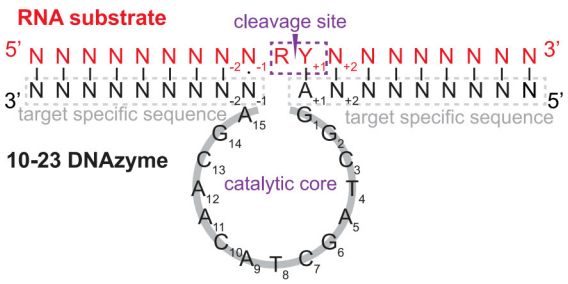
Deoxyribozymes or DNAzymes are single-stranded DNA molecules that are capable of catalyzing a variety of chemical reactions, including RNA cleavage (Santoro & Joyce, 1997) and ligation (Flynn-Charlebois *et al.*, 2003), DNA cleavage (Chandra *et al.*, 2009) and ligation (Cuenoud & Szostak, 1995), the photoreversion of thmyine dimers (Chinnapen & Sen, 2004), as well as peptide side chain and backbone modifications (Chandrasekar & Silverman, 2013; Chu *et al.*, 2014; Pradeepkumar *et al.*, 2008; Silverman, 2015; Wong *et al.*, 2011). The 10-23 DNAzyme is the most extensively studied RNA-cleaving DNAzyme. It was obtained via *in vitro* selection from a pool of randomized DNA sequences. The 10-23 DNAzyme consists of a catalytic loop of 15 nucleotides that is flanked by two substrate binding arms (Figure 2.1A). Those arms can be varied in length and sequence in order to allow specific binding to virtually any RNA of interest. After binding, the catalytic loop facilitates cleavage of the RNA substrate between a 5' central purine and its 3' neigbouring pyrimidine nucleotide (Figure 2.1B). The 10-23 DNAzyme has been considered a promising tool to reduce the expression of therapeutically relevant genes on the RNA level (Cairns *et al.*, 1999; Cho *et al.*, 2013; Fokina *et al.*, 2012; Krug *et al.*, 2015; Kurreck *et al.*, 2002a; Santiago *et al.*, 1999b; Santoro & Joyce, 1997; Sel *et al.*, 2008; Victor *et al.*, 2018; Yen *et al.*, 1999; Zhang *et al.*, 2004, 2002). However, a major obstacle for its *in vivo* application is the high dependency of the DNAzyme on divalent metal ions. Indeed, recent studies suggest that the 10-23 DNAzyme is catalytically inactive under the conditions inside cells and that visible knockdown effects could be attributed to antisense effects (Victor *et al.*, 2018; Young *et al.*, 2010). Despite the efforts of several groups to improve catalytic performance by designing DNAzymes that function at lower metal ion concentration, with different metal ion specificity, or even in the complete absence of metal ion cofactors, the activity of the DNAzyme under conditions resembling the cell remains low (Hollenstein *et al.*, 2009; Kasprowicz *et al.*, 2017; Lermer *et al.*, 2002; Sidorov *et al.*, 2004).

In this review, we aim at providing an overview of the structural and functional data on the RNA-cleaving 10-23 DNAzyme. First, we highlight the potential of the 10-23 DNAzyme for therapeutic applications. We then summarize efforts to gain structural insights on DNAzymes utilizing various methods and point out the challenges for this endeavor. We further describe the reaction mechanism with its relevant parameters. In the following sections, we take a closer look at the functional data available on the impact of nucleotide substitutions within the binding arms as well as substitutions and modifications within the catalytic loop. Finally, we outline the influence of various metal ions on activity.

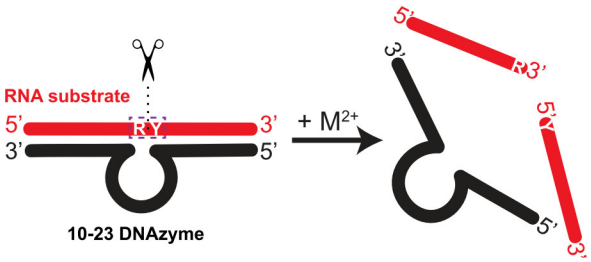
Therapeutic potential of the 10-23 DNAzyme

The ability of the 10-23 DNAzyme to cleave disease-related (messenger or non-coding) RNA makes it particularly attractive as a therapeutic agent. It is highly versatile yet specific. In comparison to the currently more established approaches for post-transcriptional gene silencing, such as the use of RNA interference (RNAi) or antisense deoxyribonucleotides (AS-ODNs),

A



B



C

$$\mathrm{Dz} + \mathrm{RNA} + \mathrm{N}^{+} \rightleftharpoons \mathrm{Dz{:}RNA{:}N}^{+} \tag{2.1}$$

$$\mathrm{Dz{:}RNA{:}N}^{+} + \mathrm{M}^{2+} \rightleftharpoons \mathrm{Dz{:}RNA{:}N}^{+}\mathrm{{:}M}^{2+} \rightleftharpoons \mathrm{Dz{:}RNA_{cleaved}{:}N}^{+} + \mathrm{M}^{2+} \tag{2.2}$$

$$\mathrm{Dz{:}RNA_{cleaved}{:}N}^{+} \rightleftharpoons \mathrm{Dz} + \mathrm{RNA_{cleaved}} + \mathrm{N}^{+} \tag{2.3}$$

Figure 2.1: Reaction and sequence of the 10-23 DNAzyme. (A) RNA cleavage performed by the 10-23 DNAzyme. The DNAzyme (black) cleaves its specific RNA substrate (red) between two central purine (R) and pyrimidine (Y) nucleotides in a manner dependent on divalent cation(s) M^{2+}. (B) Schematic representation of the 10-23 DNAzyme (indicated in black) with the 15-nt comprising catalytic core bound to the RNA substrate (indicated in red). While in general the binding of the DNAzyme with the RNA substrate occurs via Watson-Crick base pairing (|), nucleotides forming wobble base pairs (·) at position −1 lead to enhanced cleavage. (C) Chemical reactions describing the cleavage activity of 10-23 DNAzyme Dz; (2.1) the DNAzyme associates with the target RNA to a complex that binds for charge neutralization more cations than the single strands; (2.2) the DNAzyme:RNA complex binds divalent cation(s) to form the catalytically active enzyme:substrate complex that is able to cleave the RNA; (2.3) the enzyme:product complex of DNAzyme and cleaved RNA dissociates into its components; for further details see text.

the DNAzyme has several remarkable advantages. These advantages include the fact that it—supposedly—is a self-sufficient biocatalyst that does not rely on the presence of other biomolecules the way that RNAi relies on argonaut proteins and AS-ODNs on the action of RNase H. These features make DNAzymes also perfect candidates to target RNA-based viruses such as human immunodeficiency viruses (HIV) (Basu *et al.*, 2000; Goila & Banerjea, 1998; Jakobsen *et al.*, 2007; Zhang *et al.*, 1999) or corona viruses (CoV), entering the organism but before entering the cell. In addition, compared to RNA, single-stranded DNA is much more cost-effective and inherently more stable in biological fluids. A growing number of 10-23 DNAzyme variants is currently in preclinical model studies and a small selection in clinical trials focusing on the treatment of basal cell carcinoma and Th2-driven asthma(Cho *et al.*, 2013; Krug *et al.*,

2015) (also see e.g. Khachigian for more detailed review of clinically relevant aspects). While the DNAzyme approach offers a very attractive therapeutic strategy, contradicting observations are found in *in vivo* experiments, including studies that report on limited catalytic activity under physiological conditions (Victor *et al.*, 2018; Young *et al.*, 2010). The current situation implies that there is an urgent need to better understand the fundamental processes of DNA-mediated catalysis to enable the design of DNAzymes with improved *in vivo* activity and unravel the full therapeutic potential of DNAzymes.

Structural information on DNAzymes

Structural information on DNAzymes is scarce and the absence of detailed information on the spatial arrangement and the metal coordination sites prevent a deeper understanding of the molecular mechanism of DNAzyme catalysis. Based on its similarity to the hammerhead ribozyme, the RNA-cleaving reaction is believed to involve a transesterification (Figure 2.2) (Breaker *et al.*, 2003b; Santoro & Joyce, 1998), however, in the absence of structural data on a molecular level the exact mechanism by which the 10-23 DNAzyme performes its reaction is still under debate. The first crystal structure reported by Nowakowski *et al.* with the 10-23 DNAzyme in complex with its RNA target showed the formation of a four-way junction. This arrangement composed of two DNAzyme strands and two RNA substrate molecules is unlikely to represent the catalytically active conformation of the complex. In a follow-up study, Nowakowski *et al.* presented a combinatorial approach using 81 different DNAzyme:RNA complexes for crystallization screenings. The different biological samples involved different combinations of DNAzymes and RNA substrates with different length, with or without overhangs. Their strategy led to the formation of 40 crystals and data sets with a resolution of up to 2.8 Å. However, the structure solved from these diffraction experiments did not lead to a crystal structure in a catalytically active conformation.

A study by Kondo & Takénaka found that heat-treatment is effective to prevent formation of the inactive quaternary complex between the two DNAzymes and the two substrates, however, they did not obtain diffracting crystals. Another attempt to obtain high resolution structural information on DNAzymes was reported by Dolot *et al.*. They have crystalized a 52−nucleotide DNA/2'-OMe-RNA oligomer mimicking 10-23 DNAzyme in complex with its substrate and were able to collect data to 1.2 Å resolution. Despite tremendous efforts to solve the phase problem, including direct methods, molecular replacement, and single-wavelength anomalous diffraction of phosphorus atoms, they were not able to derive the phases and obtain an electron density.

All attempts described so far have focused on determining the structure using X-ray crystallography. However, the formation of highly ordered nucleic acid crystals is often challenging due to the surface properties of nucleic acids. While the surfaces of proteins consist of hydrophobic patches that favour intermolecular interactions, the surfaces of nucleic acids is dominated by negatively charged and regularly ordered phosphate groups. Uniform surfaces prevent the formation of distinct interactions between molecules in the crystal lattice and lead to the formation of crystals with poor long-range order. In addition to the challenge of obtaining high

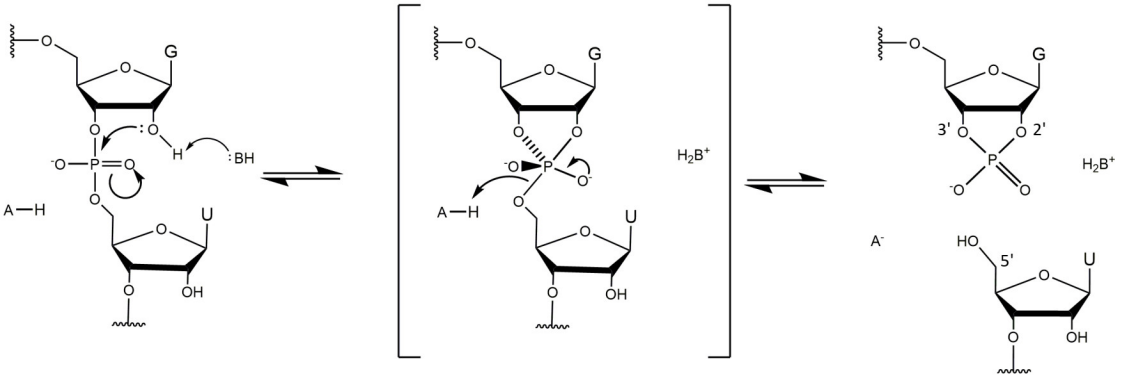


Figure 2.2: Proposed reaction mechanism of the 10-23 DNAzyme. A proton is abstracted from the 2'-hydroxyl group on the ribose of the unpaired purine nucleotide by an unknown Brønsted base B. The resulting oxyanion functions as a nucleophile that attacks the phosphorus center of the phosphodiesterbond thus generating a penta-coordinated phosphorane intermediate. The intermediate degrades into two RNA fragments: one fragment that terminates in a 2'-3'-cyclic phosphate and second fragment that terminates in a 5' hydroxyl group. G = guanine; U = uracil. Based on Breaker *et al.*.

quality crystals, solving the phase problem is an additional obstacle in nucleic acid crystallography. The phase problem arises from the loss of phase information during the diffraction experiment. When collecting X-ray diffraction data from a crystal, the intensities of the diffracted waves scattered from a series of planes are measured (Taylor, 2010). From these intensities the amplitudes of the scattered waves are derived. At this point, we lose the phase information, which describes the offset of these waves when we add them together to reconstruct an image of our molecule. For small molecule crystallography determining phases by *ab initio* approaches is quite common. In protein crystallography phases are derived either by using the atomic coordinates of a structurally similar protein (molecular replacement) or by finding the positions of heavy atoms that are intrinsic to the protein or that have been added (methods such as MIR, MIRAS, SIR, SIRAS, MAD, SAD or combinations of these). Strategies to obtain heavy atom derivatives of RNA crystals (Keel *et al.*, 2007) have been described as well as co-crystallization approaches with nucleic acid-binding proteins (Ferré-D'Amaré *et al.*, 1998; Liu *et al.*, 2017), but solving the phase problem for nucleic acid data remains nontrivial. The only available high-resolution structural information on DNA catalysts have been obtained for the RNA-ligating DNAzyme 9DB1 and the RNA-cleaving 8-17 DNAzyme. The crystal structure of the DNAzyme 9DB1 has been solved in complex with its two RNA substrates at a resolution of 2.8 Å (Ponce-Salvatierra *et al.*, 2016). The only crystal structure of an RNA-cleaving DNAzyme has been solved in the presence of a DNA instead of the RNA substrate at a resolution of 2.55 Å (Liu *et al.*, 2017). However, since the investigated DNA catalyst belongs to the family of the 8-17 DNAzymes the information obtained from this structural study do not allow for any conclusion about the mechanistic aspects of RNA-cleavage performed by the 10-23 DNAzyme. Information about size and shape of the 10-23 DNAzyme in complex with its RNA substrate and the RNA-binding protein U1A has been obtained by small-angle X-ray

scattering (SAXS) (Rosenbach *et al.*, 2020). Another structural model of the 10-23 DNAzyme in complex with its RNA target has been obtained by molecular dynamics (MD) simulation (Plashkevych & Chattopadhyaya, 2011). However, the latter study was not confirmed by any experimental validation of the presented structure. Attempts to obtain the solution structure of the 10-23 DNAzyme:RNA complex by NMR spectroscopy have not yet been successful (Choe *et al.*, 2000), but offer a promising alternative strategy. To our knowledge no efforts to solve the structure of DNAzymes utilizing cryogenic electron microscopy (cryo-EM) have been reported so far. The absence of a high-resolution structure of a RNA-cleaving DNAzyme bound to its RNA substrate has hampered the efforts to modify the 10-23 DNAzyme for improved *in vivo* performance.

Kinetics of DNAzyme-mediated catalysis

The chemical reactions shown in Figure 2.1C depict a minimal kinetic scheme for catalysis by the 10-23 DNAzyme. In the association step (2.1), the DNAzyme's binding arms form base-pairs with the target-specific RNA (Figure 2.1B). The thermodynamic stability of the complex consisting of DNAzyme and RNA depends on reactants' structures, which have to be dissolved at least partially prior to association, on the lengths of the formed helices, the basepair stacking in the helices, and the ionic strength. A well-designed DNAzyme sequence should form neither intramolecular nor intermolecular basepairs with itself; the RNA or at least the target region should also have a low degree of structure (Tafer *et al.*, 2008; Turner *et al.*, 1990; Victor *et al.*, 2018). The stability of the complex increases with increasing length of the helical arms and strongly stacking basepairs (Steger, 1994; Sugimoto *et al.*, 1995), but a too high stability of the helices inhibits dissociation of the DNAzyme:product complex (see Figure 2.1C(2.3)) and thus catalytic turnover. Increasing ionic strength overcomes the electrostatic repulsion of the poly-electrolytic nucleic acid single strands, leads to higher stability and denaturation temperature of the complex, but may be varied only *in vitro* while ionic conditions are given in a selected *in vivo* system. While reaction (2.1) is independent from the type of cation, the catalytic reaction depends on divalent cations M^{2+}; most experimental information is available for Mg^{2+} or Mn^{2+} (see below). Notably, divalent cations may substitute for the mentioned monovalent ions in reaction (2.1), especially in *in vitro* experiments; on the other hand, high concentrations of divalent cations—especially in combination with slightly basic pH values and/or elevated temperatures—may lead to degradation of RNA (Forconi & Herschlag, 2009). Such high concentrations of divalent cations are quite often used to analyze catalysis of mutated DNAzymes (see below), however, controls for the cation-induced RNA degradation are rarely performed.

The combined reactions (2.2)+(2.3) (Figure 2.1C) can be summarized and interpreted by a standard Michaelis-Menten equation; that is,

$$\mathrm{Dz} + \mathrm{RNA} + \mathrm{N^+} + \mathrm{M^{2+}} \underset{k_\mathrm{r}}{\overset{k_\mathrm{f}}{\rightleftharpoons}} \mathrm{Dz{:}RNA{:}N^+{:}M^{2+}} \xrightarrow{k_\mathrm{cat}} \mathrm{Dz} + \mathrm{RNA_{cleaved}} + \mathrm{N^+} + \mathrm{M^{2+}}$$

with rate constants k_f fo the forward reaction, k_r of the reverse reaction, and k_{cat} of the catalytic reaction. Under the conditions that DNAzyme concentration is much below RNA concentration, dissociation of DNAzyme and cleaved RNA is fast, M^{2+} concentration sufficiently high, and the rate of RNA ligation low (Santoro & Joyce, 1998), the rate of product formation is given by

$$\frac{\mathrm{d}\,[\mathrm{RNA_{cleaved}}]}{\mathrm{d}\,t} = V_{max}\frac{[\mathrm{RNA}]}{K_M + [\mathrm{RNA}]}$$

with maximum rate

$$V_{max} = k_{cat} \cdot [\mathrm{Dz_{total}}]$$

and Michaelis constant

$$K_M = \frac{k_r + k_{cat}}{k_f}\ .$$

Under single turnover conditions, i. e. DNAzyme is in excess over RNA or the reaction is started by addition of M^{2+} to the preformed Dz:RNA complex, as done in many of the analyses described below, a simple first-order reaction

$$\mathrm{Dz{:}RNA} \xrightarrow[M^{2+}]{k_{obs}} \mathrm{Dz} + \mathrm{RNA_{cleaved}}$$

is obtained. From the concentration dependence upon time t, the rate constant

$$[\mathrm{RNA_{cleaved}}]_t = [\mathrm{Dz{:}RNA}] \cdot \left(1 - \exp(-k_{obs}t)\right)$$

and yield

$$Y_t = \frac{[\mathrm{RNA_{cleaved}}]_t}{[\mathrm{RNA}]_{t=0}} \tag{2.4}$$

are obtained. k_{obs} is the observed rate constant.

2.4 Impact of nucleotide substitutions within the binding arms

DNAzyme activity depends upon nucleotides neighbouring the cleavage site

The 10-23 DNAzyme cleaves RNA substrates with high nucleotide selectivity at purine-pyrimidine junction sites (rRrY) in the presence of divalent metal ions. The purine nucleotide of the RNA cleavage site remains unpaired while the pyrimidine nucleotide forms a base pair with one of the deoxyribonucleotides of the DNAzyme. The pyrimidine nucleotide of the RNA cleav-

age site as well as the paired base of the DNAzyme will be referred to as being in position +1 in the following because they are located “downstream” of the actual site of cleavage, i. e. between the purine and pyrimidine nucleotide (see Figure 2.1A and 2.1B). In this paragraph we will focus on the (i) impact of base pair mutations at the scissile bond, (ii) the influence of mutations of stacking interactions between the ribonucleotides neighbouring the cleavage site and the deoxynucleotides at position 1 and 15 of the DNAzyme’s catalytic loop, as well as (iii) the impact of the introduction of modified deoxynucleotides at position 1 of the catalytic loop on the cleavage rate. In 2003, Cairns *et al.* found that the cleavage rates of the DNAzyme for RNA substrates increase in the following order of the cleavage site sequence: $rGrU_{+1} = rArU_{+1} \geqslant rGrC_{+1} \gg rArC_{+1}$. When uridine (rU) is found in the +1 position the nature of the unpaired purine base has no impact on the cleavage rate. However, if cytosine (C) is found in the +1 position the type of the unpaired purine also affects activity. The significantly higher activity of the 10-23 DNAzyme against rRrU substrates over rRrC substrates could be explained by a higher binding strength of the rC_{+1}:dG_{+1} pairing at position +1 with three hydrogen bonds compared to the relatively weak rU_{+1}:dA_{+1} pairing with only two hydrogen bonds and altered stacking features. Also, the cleavage rates for substrates with an rRrY junction at the cleavage site are dramatically increased when using DNAzymes with a deoxyinosine (dI) nucleotide (see Figure 2.3 for structures of the different nucleotide modifications, which are reviewed in the following; the nucleobase in deoxyinosine is hypoxanthine) instead of the canonical dR counterpart at the +1 position. The substitution results in two weaker hydrogen bonds $rY_{+1} \cdot dI_{+1}$ pairing compared to the canonical Watson-Crick base pairing. This effect is particularly striking for substrates with an rArC core sequence, resulting in a $rC_{+1} \cdot dI_{+1}$ interaction instead of the rC:dG bonding. The opposite effect was observed when using DNAzymes with a diaminopurine (DAP) substitution at position +1 for rRrU substrates. The three hydrogen bond between rU_{+1} and $dDAP_{+1}$ pair significantly decreases the cleavage rate of the reaction compared to the weaker rU_{+1}:dA_{+1} pair.

However, this rate enhancement has only been observed when using substrates with an rArU and not a rGrU cleavage site indicating an influence of the stacking interaction between AU or GU, respectively. This observation goes hand in hand with the one made for substrates with an rRrC core sequence where the cleavage rate was extensively enhanced by a lower binding strength when the unpaired rR nucleotide was an adenosine. In addition to being unpaired, rArY or rGrY cleavage sites require an extended degree of conformational flexibility or lower stacking interaction on the 3’ side (N_{+1}) for efficient cleavage rates. This is provided by standard nucleotides when followed by uridine (rU_{+1}:dA_{+1} pair), whereas a paired dG needs to be replaced by a dI leading to a wobble base pair to enhance the activity.

Substitution of the dA in an rU_{+1}:dA_{+1} pair at the cleavage site by 7-deaza (ΔN^7)-deoxyadenine and 8-aza-7-deaza ($N^8\Delta N^7$)-deoxyadenine reveal that the nitrogen atom is of equal importance for the cleavage reaction at positions 7 and 8 in the ring system of dA_{+1}, although these atoms do not contribute to thermal stability by forming hydrogen bonds. Deletion of the exocyclic amino group at position 6 or its substitution by bulky groups lead to decreased stability and a lower reaction rate (Zhu *et al.*, 2015). Table 2.1 provides an overview of the effect of nu-

Table 2.1: Effect of nucleotide substitutions at position +1 of the DNAzyme (dN_{+1}) and the sequence of the nucleotides at the cleavage site of the RNA substrate (paired as well as unpaired N_{+1}) on the RNA cleavage rate. DAP_{+1}, 2,6-diaminodeoxypurine; $d\Delta N^7A_{+1}$, 7-deazadeoxyadenine; $d(N^8\Delta N^7)A_{+1}$, 8-aza-7-deazadeoxyadenine; for formulas see Figure 3.

$A^{\downarrow}U_{+1}$	=	$G^{\downarrow}U_{+1}$	$\geqslant$	$G^{\downarrow}C_{+1}$	$\gg$	$A^{\downarrow}C_{+1}$
$R^{\downarrow}U_{+1}$		$R^{\downarrow}C_{+1}$				
‖	>	‖				
dA_{+1}		dG_{+1}				
$R^{\downarrow}Y_{+1}$		$R^{\downarrow}Y_{+1}$		$A^{\downarrow}C_{+1}$		$A^{\downarrow}C_{+1}$
‖	>	‖		‖	$\gg$	‖
dI_{+1}		dR_{+1}		dI_{+1}		dG_{+1}
$A^{\downarrow}U_{+1}$		$A^{\downarrow}U_{+1}$		$G^{\downarrow}U_{+1}$		$G^{\downarrow}U_{+1}$
‖	>	‖		‖	$\approx$	‖
dA_{+1}		DAP_{+1}		dA_{+1}		DAP_{+1}
$R^{\downarrow}U_{+1}$		$R^{\downarrow}U_{+1}$		$R^{\downarrow}U_{+1}$		
‖	$\approx$	‖	$\approx$	‖		
dA_{+1}		$d\Delta N^7A_{+1}$		$d(N^8\Delta N^7)A_{+1}$		

cleotide substitutions at position +1 of the DNAzyme (dN_{+1}) and the sequence of nucleotides at the cleavage site of the RNA substrate (paired and unpaired rN_{+1}, respectively) on the reaction rate.

The neighbouring nucleotide rN_{-1} on the 5' side of the unpaired purine next to the cleavage site prefers formation of a non-canonical or wobble interaction with the corresponding deoxynucleotide within the DNAzyme. A suitable substitution of this deoxynucleotide in the DNAzyme sequence leads to an increase in the RNA cleavage rate (Cairns *et al.*, 2000, 2003). It has been reported that the activity of the DNAzyme against some RNA substrate sequences was improved by base substitutions with reduced interaction strength between the DNAzyme and the substrate. It was shown that the formation of wobble and mispairs immediately 5' rather than 3' of the cleavage site are responsible for an enhanced DNAyzme activity (Cairns *et al.*, 2003).

Introduction of an intercalator between the catalytic loop deoxynucleotide dA_{15} and dN_{-1} of the DNAzyme binding arm greatly improved the cleavage activity towards an RNA substrate with a rCrG cleavage site compared to the unmodified DNAzyme. The intercalator was attached via an amide bond to a D-treoninol linker that is inserted to the DNA backbone using phosphoramidite chemistry (Asanuma *et al.*, 2006). It was shown, that the catalytic activity was only improved, when the intercalator was introduced to the DNAzyme via D-treoninol, while

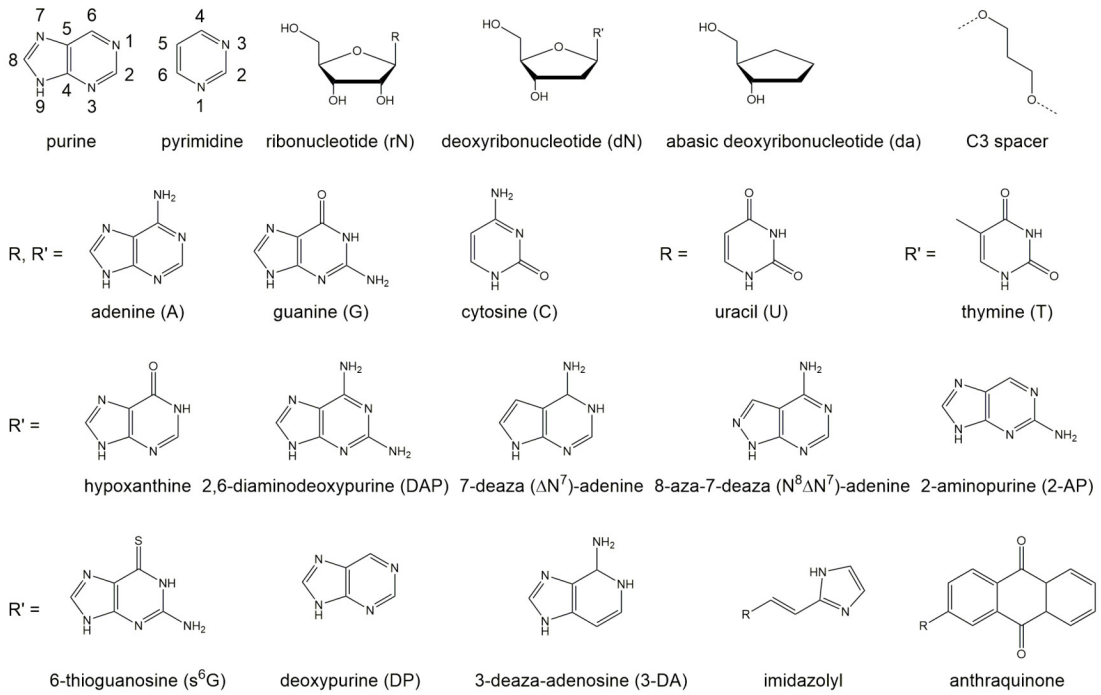


Figure 2.3: Structures of different nucleotides and C3 spacer used in mutagenesis studies for the 10-23 DNAzyme. The nucleobase hypoxanthine is present in the nucleotide inosine.

introduction of methylene between the intercalator and the amide bond significantly lowered the DNAzyme activity. The most enhancing effect on the DNAzyme was detected using an antraquinone intercalator. Furthermore, results obtained with additional intercalators and linkers revealed that planar molecules have an enhancing effect on DNAzyme activity. It was also reported that in the catalytic loop of the 10-23 DNAzyme, dA_{15} could be exchanged by a nucleobase lacking its 6-amino group to improve the catalytic rate. Furthermore, it was shown that the DNAzyme can be further optimized by adding an extra functional group to the 6-amino group via different C2- or a C3-linkers (Zhu *et al.*, 2016). Other modifications at dA_{15} appeared to show a negative effect on the cleavage activity (He *et al.*, 2011; Wang *et al.*, 2012).

Furthermore, studies with 2'-O-methyl modified nucleotides as well as locked nucleic acid (LNA) substitutions within the binding arms of the 10-23 DNAzyme have shown that modifications that promote an A-form helix of the binding arms enhance the activity of the DNAzyme (Schubert *et al.*, 2003; Vester *et al.*, 2002). While DNA double helices form B-type helices, the A conformation is common for RNA double helices and has also been reported for DNA-RNA helices (Milman *et al.*, 1967).

Taken together, a certain degree of flexibility in the base pairing between the substrate nucleotides next to the cleavage site and the corresponding deoxynucleotides in the DNAzyme binding arms is crucial for cleavage activity of the 10-23 DNAzyme. The required conformational

freedom might be a hint for a significant rearrangement of the RNA cleavage site towards functional deoxynucleotides in the catalytic loop sequence. This assumption is further supported by the observation that the transition from the catalytic loop to the substrate binding arm, i. e. the connection between dA_{15} and dN_{-1}, can be modified without major influence on the activity, indicating that no rigid conformation is required here.

2.5 Impact of deoxynucleotide substitutions and modifications within the catalytic loop

The structure of the DNAzyme in complex with its RNA substrate in a catalytically relevant conformation remains unknown. Therefore, much effort has been spend into determining how the deoxynucleotides within the catalytic loop affect DNAzyme function. Several studies have investigated the role of each individual nucleobase (Asanuma *et al.*, 2006; He *et al.*, 2011; Li *et al.*, 2014b; Nawrot *et al.*, 2007; Räz & Hollenstein, 2015; Santoro & Joyce, 1998; Schubert *et al.*, 2003; Smuga *et al.*, 2010; Wang *et al.*, 2010a, 2012; Zaborowska *et al.*, 2002; Zhu *et al.*, 2015), deoxyribose sugars (Robaldo *et al.*, 2012; Schubert *et al.*, 2003; Wang *et al.*, 2010a), and the phosphate groups (Breaker *et al.*, 2003b; Nawrot *et al.*, 2007). In addition deletion studies provide insights into the relevance of the nucleotides in different positions (Okumoto & Sugimoto, 2000; Sugimoto *et al.*, 1999; Zaborowska *et al.*, 2005). In this paragraph, we will discuss mutational studies with natural and non-natural deoxynucleotides. Moreover, we will outline the impact of substituting individual deoxynucleotides with abasic deoxynucleotides (da) and acyclic C3 spacers for each deoxynucleotide within the catalytic loop region. In contrast to mutations of deoxynucleotides that disturbed the given structure of the DNAzyme, the introduction of abasic deoxynucleotides or spacers allow for more flexibility of the structure. Figure 2.4 provides an overview of the relevance of each nucleotide in the catalytic loop on 10-23 DNAzyme activity at a glace.

The deoxynucleotides at position 1 to 6, 13, and 14 are crucial for activity

The sequence requirements in the catalytic loop of the 10-23 DNAzyme were analyzed by Zaborowska *et al.* using systematic mutagenesis studies with regard to DNAzyme activity in the presence of Mg^{2+}. The reaction of each DNAzyme variant was performed for 20 min at 37 °C in the presence of tenfold molar excess of the DNAzyme over a 19 nucleotide (nt) RNA substrate in the presence of 10 mM $MgCl_2$. They used the total yield of the reaction product after 20 min as an indicator of DNAzyme activity. This approach is less labor-intense, however, the results are less reliable as compared to kinetic measurements. Zaborowska *et al.* found that the deoxynucleotides dG_1, dG_2, dT_4, dG_6 and dG_{14} could not be exchanged independently by any other naturally occurring deoxynucleotide without a complete loss of cleavage activity. Surprisingly, Wang *et al.* found that the exchange of the deoxynucleotides dG_1, dG_2 and dT_4 by either an abasic deoxynucleotide (da) or a C3 spacer did not lead to a considerably loss of activity. It

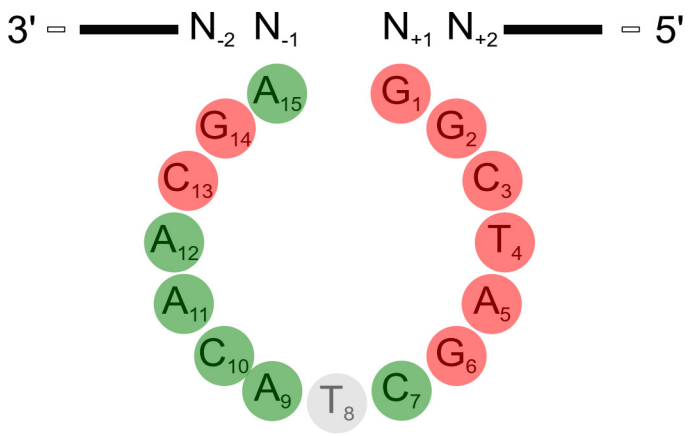


Figure 2.4: Schematic representation of the 10-23 DNAzyme with emphasis on the deoxynucleotides within the catalytic loop. Modification of the deoxynucleotides 1-6, 13, and 14 (red circles) greatly affect the cleavage rate, whereas exchanges of deoxynucleotides 7-12 and 15 (green circles) only slightly affect the DNAzyme activity. Deletion of the deoxynucleotide dT_8 leads to an active 10-23 DNAzyme (grey circle)

should be pointed out that the kinetic measurements with the abasic deoxynucleotides and C3 spacers by Wang *et al.* were performed in the presence of a 100-fold excess of DNAzyme over the substrate in the presence of 200 mM NaCl and 500 mM $MgCl_2$. The introduction of an abasic deoxynucleotide substantially reduces the observed rate constant k_{obs} values compared to the unmodified DNAzyme in the following order: $da_4 > da_2 > da_1$, with the substitution of dT_4 showing the largest effect. Mutation of these deoxynucleotides to a C3 spacer does not lower the k_{obs} values significantly further compared to the substitution by an abasic deoxynucleotide. In case of da_1, da_2 and da_4 , the overall yields after a given time (Y_t; see (2.4)) were only slightly reduced when compared to the unmodified DNAzyme, suggesting that the nucleobases may only be necessary for the correct positioning of other nucleotides that are directly involved in the reaction mechanism. Notably, the cleavage assays for the abasic and C3 spacer substitutions were conducted using chimeric DNA/RNA substrates with a single ribonucleotide linkage at the cleavage site (Wang *et al.*, 2010a). As mentioned previously, DNA:DNA duplexes form B-type helices, while RNA:DNA duplexes form A-type helices. The formation of a different helix type is caused by the different sugar conformation, i. e. different sugar puckers, and a different number of base pairs per turn. Additionally, the 2'-OH group on the RNA ribose is not compatible with the deep and narrow minor grove of B-type helices (Arnott *et al.*, 1986a,b; Milman *et al.*, 1967). In summary, the experiments performed by Wang *et al.* and Zaborowska *et al.* point towards a significant importance of the nucleotides dG_1, dG_2, dT_4, dG_6 and dG_{14} under single turnover conditions.

Exocyclic functional groups at position 6 and 14 are crucial for the activity of the 10-23 DNAzyme

The dramatically reduced DNAzyme activity after replacement of dG_{14} by dI indicates the importance of the 2-amino group. When exchanging the dG_{14} deoxynucleotide by a 2- aminopurine (2-AP) deoxynucleotide that lacks the 6-keto group of dG, the DNAzyme becomes almost completely inactive (Zaborowska *et al.*, 2002), emphasizing that both exocyclic functionalities of the guanine base are crucial for the RNA cleavage activity of the 10-23 DNAzyme, although dG_6 cannot be exchanged to dA without a complete loss of DNAzyme activity, the deoxynucleotide can be changed to dI with no detectable difference in the cleavage rate. This data suggests that the 6-keto group plays an important role in the cleavage mechanism, whereas the 2-amino group does not appear to be critical for function (Zaborowska *et al.*, 2002) (Figure 2.5). This was further confirmed by mutagenesis studies by Nawrot *et al.*, where modified DNAzymes with a 2-AP substitution as well as a ΔN^7G or a 6-thioguanosine (s^6G) replacement at position 6 were assayed for their activity in the presence of either Mg^{2+} or Mn^{2+}. The 6-keto group of guanine is lacking in the 2-AP variant, while it is replaced by a thiol group in the s^6G mutant. The ΔN^7G lacks the nitrogen atom at position 7 of the guanine base. According to the thio effect based on the Pearson acid-base concept, also termed Hard and Soft Acid and Base (HSAB) model, a non-polarizable hard Mg^{2+} has a lower affinity to the polarizable soft sulfur-containing DNAzyme (Pearson, 1968). As a result, the cleavage rate is lower, whereas a recovery of the cleavage rate, the so-called rescue effect, can be observed when soft Mn^{2+} are used as cofactor instead of Mg^{2+}. The different kinetic behavior of the 2-AP and s^6G DNAzyme mutants in the presence of either Mg^{2+} or Mn^{2+} (Nawrot *et al.*, 2007) confirmed the importance of the 6-keto group proposed by Zaborowska *et al.* in 2002. The substitution of dG_6 by ΔN^7G results in a more than 100-fold loss of activity in the presence of Mg^{2+}, which suggests that the N7 nitrogen is involved in the formation of intramolecular hydrogen bonds that are crucial for a functional conformation of the 10-23 DNAzyme (Nawrot *et al.*, 2007). Substitution of the deoxynucleotides dG_6 and dG_{14} by an abasic deoxynucleotide completely abolished the cleavage activity, in accordance with the proposed relevance of the exocyclic functional groups at the nucleobases at positions 6 and 14 (Wang *et al.*, 2010a).

The deoxynucleotides at position 7 to 12 could be replaced by other naturally occurring deoxynucleotides without severe effects. Only the replacement of dC_7 to a dA and of dA_9 to a dC reduced the cleavage activity by 80% to 90%. While mutation of the deoxynucleotide dC_7 to a dG or dI led to a decrease in cleavage activity by more than fourfold, the activity was only slightly affected by substitution to adenosine or thymidine (Zaborowska *et al.*, 2002). Substitution by an abasic deoxynucleotide or a C3 spacer significantly reduced the k_{obs} value of the cleavage reaction, but only slightly affected the Y_t (Wang *et al.*, 2010a). However, this observation is surprising, since the cleavage reaction was performed for 24 h under single-turnover conditions, in which excess DNAzyme (DNAzyme:substrate ratio 100:1) was used in the presence of 200 mM NaCl and 500 mM $MgCl_2$. Under these conditions complete cleavage of the RNA substrate would be expected. Thus, incomplete cleavage can be a hint for either insuf-

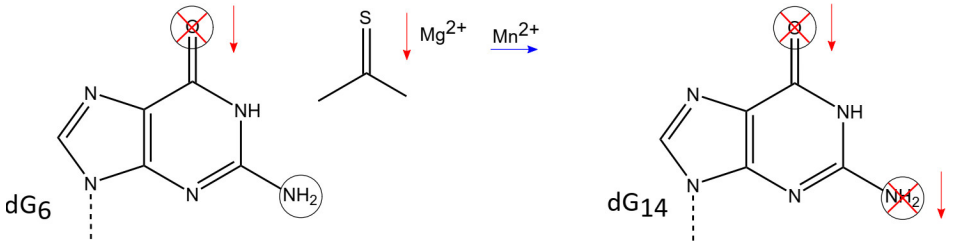


Figure 2.5: Proposed important exocyclic groups within the 10-23 DNAzyme. The 6-keto group and the 2-amino group of dG_{14} as well as the 6-keto group of dG_6 are of functional imporatance for the DNAzyme activity. Deletion of either the 6-keto group or the 2-amino group of dG_{14} as well as the deletion of the 6-keto group of dG_6 lead to an dramatic decrease in catalytic acitivity of the 10-23 DNAzyme as indicated by the red arrows. Substituting the oxygen atom of the 6-keto group in dG_6 by a sulfur atom leads to a DNAzyme with significantly reduced cleavage activity in the presence of Mg^{2+} (indicated by a red arrow), but the activity can be rescued in the presence of Mn^{2+} (indicated by a blue arrow). In addition, deletion of dG_{14} does significantly affect the acitivity of the DNAzyme.

ficient association of the uncleaved substrate with the DNAzyme or for complex formation in an inactive conformation. The same is true for dA_9, where only the substitution to a dC, an abasic deoxynucleotide or a C3 spacer leads to a significant decrease in the cleavage rates (Zaborowska *et al.*, 2002, 2005). Replacement of dA_{11} or dA_{12} by dI did not have a strong effect on the cleavage activity. In case of dC_7, dT_8 and dC_{10}, the exchange to dI or dG does not abolish DNAzyme activity. Deletion studies performed with a DNAzyme lacking dC_7 confirm that this variant still retains a relatively high activity (approx. 60%-80%, depending on the substrate sequence) towards a 19 nucleotide RNA substrate (Zaborowska *et al.*, 2005). At positions dA_9, dA_{11} and dA_{12} the exchange of the dA by dI led to a twofold decrease in the cleavage activity, which makes a significant importance of the 6-amino group for the cleavage mechanism at these positions unlikely (Zaborowska *et al.*, 2002). Substituting dC_{10}, dA_{11}, dA_{12} by abasic deoxynucleotides decreased the k_{obs} values by 90%. For the deoxynucleotides dA_{12} and dA_{11} no effect on Y_t was observed for substitutions with abasic nucleotides and a C3 spacer, while the Y_t was reduced by about 50% when changing dC_{10} for an abasic deoxynucleotide or a C3 spacer. Surprisingly, the mutants with an abasic deoxynucleotide or a C3 spacer at position 8 were found to exhibit a slightly increased k_{obs} and Y_t value compared to the unmodified DNAzyme (Wang *et al.*, 2010a). Introduction of specific functional groups at position dA_9 appear to have an enhancing effect on the DNAzyme activity (He *et al.*, 2011; Li *et al.*, 2014b; Zhu *et al.*, 2016).

The exocyclic amino groups at dC_3 and dC_{13} display additional functional importance

The exchange of the nucleotides at positions dC_3 and dC_{13} to dG, dT or dI drastically decreases DNAzyme activity, whereas a mutation to dA is well tolerated (Zaborowska *et al.*, 2002). Substitution of dC_{13} by an abasic deoxynucleotide completely abolishes activity, in accordance with the importance of the exocyclic 4-amino group. In comparison, total removal of the base at

dC_3 only reduces the k_{obs} by 50% and does not affect the Y_t (Wang *et al.*, 2010a), while removal of the exocyclic 4-amino group completely abolishes the DNAzyme activity (Zaborowska *et al.*, 2002). Substitution of dA_{15} by dI only had a minor effect on activity and substitution by a ΔN^7-deoxyadenine with an amino acid side-chain led to a fully active DNAzyme (Smuga *et al.*, 2010), whereas the substitution by dC reduces cleavage activity by 90% (Zaborowska *et al.*, 2002).

Non-bridging oxygen atoms of the phosphate group between dT_4 and dA_5 may be involved in metal ion coordination

At position dA_5 in the DNAzyme, the exchange of the adenosine by dI significantly reduced the cleavage activity, whereas the exchange to dC only led to a twofold decrease in the activity (Zaborowska *et al.*, 2002). Replacement of the nucleobase to a deoxypurine (DP) did not alter the activity compared to the unmodified DNAzyme (Zaborowska *et al.*, 2002). In contrast, an abasic substitution at this position completely abolishes the function of the DNAzyme (Wang *et al.*, 2010a). These findings strongly suggest that the exocyclic amino group at this position (position 6 of the purine ring system) is not of functional importance, but that probably the nitrogen in the ring system plays an important role in hydrogen bond formation (Zaborowska *et al.*, 2002). Studies by Räz & Hollenstein revealed that also the amine at position 3 of the purine ring does not play a critical role in 10-23 DNAzyme function. Nawrot *et al.* analyzed the relevance of particular phosphates within the catalytic loop by a systematic modification of each phosphate with a phosphorothioate (PS) analogue. Within the internucleotide phosphodiesters two identical unesterified oxygen atoms that share a negative charge are attached to the sp^3-hybridized phosphorous atom. The descriptors used for these oxygen atoms are *pro*-R and *pro*-S and they are used to distinguish between the two atoms. Substitution of one of these two non-bridging oxygen atoms by a sulfur atom leads to either the S_P or the R_P diastereoisomer (Figure 2.6). The activity of the modified DNAzymes can then be measured in the presence of Mg^{2+} and Mn^{2+}. Nawrot *et al.* performed such an experiment in tris(hydroxymethyl)aminomethane (Tris) buffer containing 100 mM NaCl and with a 100-fold molar excess of the DNAzyme over the RNA substrate. The results of this study confirm the involvement of both non-bridging oxygens between dT_4 and dA_5 (referred to as $d5^P$) in the hydrolysis reaction. With regard to the importance of the 6-keto group in the proximal deoxynucleotide dG_6, Nawrot *et al.* propose a model that involves both oxygen atoms at position $d5^P$ as well as the 6-keto group of dG_6 in the binding of one or more divalent metal ions. Experiments with thio-deoxyribozymes of stereodefined P-chirality also suggested that the R_P at position $d9^P$ is directly involved in metal ion coordination. Since significant thio and rescue effects were also detected for the oxygen atoms at positions $d2^P$, $d4^P$, $d10^P$, $d11^P$, $d12^P$, and $d13^P$, these ligands may also be involved in the coordination of metal ions (Nawrot *et al.*, 2007).

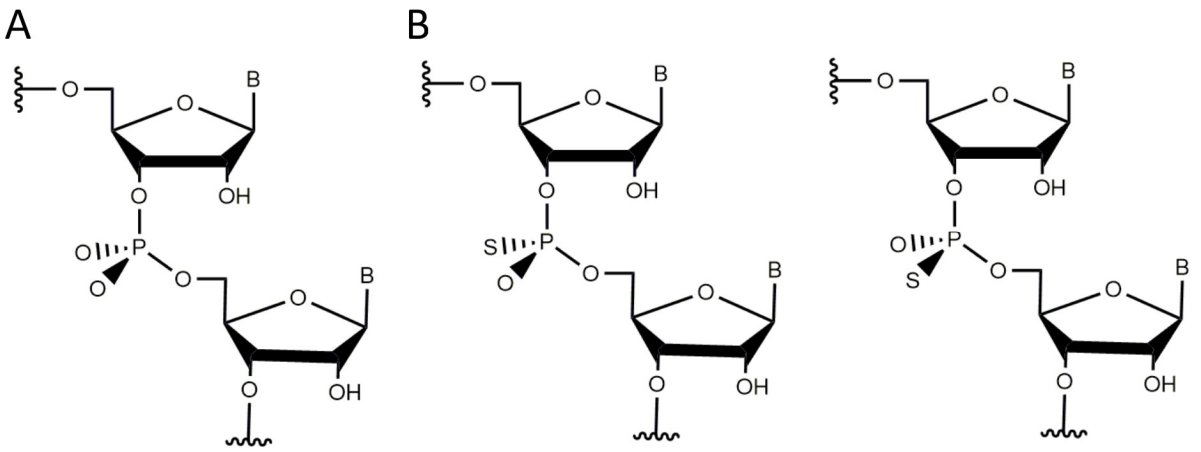


Figure 2.6: Conformations of the non-brinding oxygen atoms in a phosphodiester bond. (A) Prochiral phosphodiester bond between two nucleotides illustrating the absolute positions of the *pro*-R and *pro*-S non-bridging oxygens. (B) S_P or R_P phosphorothioate diastereomers after substitution of one of the identical non-bridging oxygen atoms by a sulfur atom.

Deletion of the deoxynucleotide dT_8 leads to an active 10-23 DNAzyme

Thymidine dT_8 was found to be the most tolerant nucleotide towards substitution by other nucleotides including dI (Zaborowska *et al.*, 2002), an abasic nucleotide, a C3 spacer (Wang *et al.*, 2010a) or a complete deletion (Zaborowska *et al.*, 2005). It has been reported that the participation of dT_8 in catalysis can be modulated by varying its ability to contribute to stacking interactions with 5-substituted azobenzene groups, where the specific stacking of the aromatic group with a different configuration could induce a positive or negative effect on the catalysis enhancement (Keiper & Vyle, 2006). Introduction of a rigidly conjugated imidazolyl group at dT_8 for expanded base stacking interaction and hydrogen-bonding network led to a slight decrease in the catalytic activity (Li *et al.*, 2017a). The weak base stacking interaction of the unmodified dT_8 has a positive effect on the reaction rate. This effect has also been confirmed by studies with modified sugars, in which (R)- and (S)-2'-C-methyl-purine as well as locked conformations had a negative effect on the catalytic activity of the DNAzyme (Robaldo *et al.*, 2012, 2010). However, deletion of dT_8 was even found to increase the cleavage activity of a DNAzyme targeting a 19-mer RNA substrate (Zaborowska *et al.*, 2005). Activity assays performed with different RNA substrates showed that in all cases dT_8 could be deleted without dramatic negative impact on the cleavage activity (Zaborowska *et al.*, 2005).

Adenine minor groove interactions play a role in 10-23 DNAzyme-mediated catalysis

Mutagenesis analyses by Räz & Hollenstein with substitutions of adenosine residues within the catalytic loop by 3-deaza-adenosine (3-DA) residues reveal the importance of hydrogen bonds that arise from the N3 atoms of adenosine (dA_5, dA_9, dA_{11}, dA_{12}, dA_{15}). N3 nitrogen atoms

in the purine rings of guanine and adenine are known to be involved in adenine minor groove interactions (A-minor interactions), which were first discovered in ribosomal RNA (Pley *et al.*, 1994a). Albeit 3-DA substitutions at each location decrease the activity of the 10-23 DNAzyme, dA_{12} appears to be the most affected deoxynucleotide with a decrease in k_{obs} of approximately 70% and of 25% in Y_t. This is remarkable, since dA_{12} was reported to be not of crucial importance for the DNAzyme activity. Substitution by other naturally occurring nucleobases and $N^8\Delta N^7$-deoxyadenine analogues did not affect the activity significantly and substitution by an abasic deoxynucleotide or a C3 spacer did not decrease Y_t suggesting that dA_{12} is not important for catalysis (He *et al.*, 2011; Zaborowska *et al.*, 2002, 2005).

A DNAzyme variant with an 11 nt-containing catalytic loop requires Ca^{2+} for its activity

Sugimoto *et al.* analyzed the effect of the DNAzyme sequence on its RNA cleavage activity by the removal of nucleotides from the catalytic loop region. Therefore, a reaction mixture consisting of a 200-fold molar excess of DNAzyme variant compared to RNA substrate in Tris buffer containing 25 mM divalent metal ions was used. After a reaction time of 90 min the mixture was analyzed using polyacrylamide gel electrophoresis. Their results show that a DNAzyme variant that has $d(A_5G_6C_7T_8)$ deleted from the original sequence retains a cleavage activity of 22.2% compared to the unmodified DNAzyme in the presence of Ca^{2+}. In the presence of either Mg^{2+} or Mn^{2+}, the activity was reduced to 10% compared to the unmodified sequence. A DNAzyme in which the nucleotides $d(A_5G_6C_7T_8)$ are replaced by abasic deoxynucleotides did not show any activity in the presence of either Mg^{2+}, Ca^{2+}, or Mn^{2+} (Sugimoto *et al.*, 1999). Since a DNAzyme with abasic substitutions at positions 5-8 were not functional and the $\Delta d(A_5G_6C_7T_8)$ variant only shows cleavage activity in the presence of Ca^{2+}, but not in the presence of Mg^{2+}. This cleavage mechanism of this new subclass was proposed to be different from the original 10-23 DNAzyme that requires the presence of Mg^{2+}. Ca^{2+} and Mg^{2+} do not only vary in the ionic radii and their charge density, they also prefer different coordination geometries: while Ca^{2+} forms complexes with eight ligands, Mg^{2+} prefers the formation of octahedral complexes with six ligands (Brown, 1988).

A summary of the effects of substitutions and modifications of single deoxynucleotides and functional groups on the activity of the 10-23 DNAzyme is given in Table 2.2. Mutagenesis studies by Zaborowska *et al.* and Wang *et al.* have identified functional groups that are important for DNAzyme catalysis: the exocyclic 4-amino groups of dC_3 and dC_{13}; the 6-carbonyl group of dG_6 and dG_{14}; the 2-amino group of dG_6; the nucleobases of dG_1, dG_2, and dT_4. Replacement of either position by any other naturally occurring nucleotide leads to complete loss of activity.

2.6 Influence of metal ions on 10-23 DNAzyme activity

The role of metal ions in ribozyme-mediated catalysis (Doherty & Doudna, 2001; Donghi & Schnabl, 2011; Fedor, 2002; Hanna & Doudna, 2000; Lilley, 2005; Schnabl & Sigel, 2010;

Table 2.2: Effect of selected substitutions of single deoxynucleotides on the activity (Y_t and k_{obs}) of the 10-23 DNAzyme towards short RNA substrates. Parameters were reported for reactions performed in the presence of tenfold excess (Zaborowska *et al.*, 2002, 2005) or in the presence of 100-fold excess of the DNAzyme over the RNA substrate (Nawrot *et al.*, 2007), and in the presence of 100- to 200-fold excess of the DNAzyme over a chimeric DNA/RNA substrate (Räz & Hollenstein, 2015; Wang *et al.*, 2010a). Unless specified differently, the values refer to Y_t. If values refer to k_{obs} this is explicitly mentioned. "–" indicates that this modification was not investigated on this position, "x" indicates that no yield Y_t could be measured for this modification. All values are given in percentages compared to the substrate cleavage performed by the unmodified DNAzyme; that is, $k_{obs} = 100 \cdot k_{obs,\,mutant}/k_{obs,\,wild\text{-}type}$ and $Y = 100 \cdot Y_{t,\,mutant}/Y_{t,\,wild\text{-}type}$.

dN	G[a]	A[a]	T[a]	C[a]	I[a]	abasic[b]	spacer[b]	deletion[c]	2-AP[a]	DP[a]	3-DA[d]	s^6G[e]	ΔN^7G[e]
G_1	–	<5[f]	<5[f]	<5[f]	40[f]	106[g] $k_{obs} = 65$	100[g] $k_{obs} = 43$	<10[f]	–	–	–	–	–
G_2	–	<5[f]	<5[f]	<5[f]	60[f]	97[g] $k_{obs} = 15$	88[h] $k_{obs} = 19$	<10[f]	–	–	–	–	–
C_3	10[f]	70[f]	20[f]	–	<5[f]	100[g] $k_{obs} = 54$	103[g] $k_{obs} = 40$	10[f]	–	–	–	–	–
T_4	<5[f]	<5[f]	–	5[f]	<5[f]	73[h] $k_{obs} = 7$	75[h] $k_{obs} = 6$	<5[f]	–	–	–	–	–
A_5	<5[f]	–	5[f]	55[f]	10[f]	6[h] $k_{obs} = 2$	–	<5[f]	–	105[f]	100[g] $k_{obs} = 50$	–	–
G_6	–	<5[f]	<5[f]	<5[f]	100[f]	x	x	<6[f]	–	–	–	$k_{obs} < 5$	$k_{obs} < 1$
C_7	20[f]	75[f]	60[a]	–	20[f]	79[h] $k_{obs} = 7$	61[h] $k_{obs} = 9$	78[f]	–	–	–	–	–
T_8	90[f]	90[f]	–	75[f]	110[f]	106[g] $k_{obs} = 124$	109[g] $k_{obs} = 129$	105[f]	–	–	–	–	–
A_9	40[f]	–	90[f]	10[f]	50[f]	71[h] $k_{obs} = 6$	69[h] $k_{obs} = 5$	30[f]	–	–	98[g] $k_{obs} = 49$	–	–
C_{10}	60[f]	75[f]	60[f]	–	60[f]	47[h] $k_{obs} = 4$	58[h] $k_{obs} = 4$	15[f]	–	–	–	–	–
A_{11}	55[f]	–	85[f]	30[f]	75[f]	94[h] $k_{obs} = 9$	94[h] $k_{obs} = 10$	30[f]	–	–	101[g] $k_{obs} = 31$	–	–
A_{12}	35[f]	–	85[f]	75[f]	95[f]	97[h] $k_{obs} = 12$	101[h] $k_{obs} = 11$	30[f]	–	–	74[g] $k_{obs} = 17$	–	–
C_{13}	5[f]	50[f]	10[f]	–	10[f]	x	x	<5[f]	–	–	–	–	–
G_{14}	–	<5[f]	<5[f]	<5[f]	10[f]	x	x	<5[f]	5[f]	–	–	–	–
A_{15}	70[f]	–	50[f]	10[f]	90[f]	84[h] $k_{obs} = 9$	86[h] $k_{obs} = 11$	<5[f]	–	–	103[g] $k_{obs} = 32$	–	–

[a] according to Zaborowska *et al.* (2002); [b] according to Wang *et al.* (2010a); [c] according to Zaborowska *et al.* (2005);
[d] according to Räz & Hollenstein (2015); [e] according to Nawrot *et al.* (2007); [f] given as $Y_{t=20\ min}$; [g] given as $Y_{t=120\ min}$; [h] given as $Y_{t=24\ h}$

Sigel & Pyle, 2007) has been reviewed in a large number of articles, which is not the case for DNAzymes. Metal ions can either directly be involved in catalysis by forming inner-sphere contacts with functional groups of the nucleic acid, or they can play a role as cofactors without direct interaction with the nucleic acid through outer-sphere contacts. In addition, they can stabilize transition states or active conformations of a DNAzyme. Moreover, metal ions are capable of generating an electrostatic environment that changes the pK_a value or the nucleophilic character of a water molecule. They are also able to alter the properties of functional groups by polarization. K^+, Na^+ and Mg^{2+} are the most abundant metal ions in living cells. They are associated with nucleic acids and function as metal ion cofactors for ribozymes *in vivo* (an excellent review article on the interaction of metal ions with nucleic acids can be found elsewhere (Pechlaner & Sigel, 2012)). Metal binding sites within DNA and RNA molecules are highly abundant as indicated in Figure 2.7. However, except the bridging and non-bridging oxygen phosphates the potential of most donor atoms to participate in metal coordination is very low due to either low or high basicity or steric hindrance, resulting in a very dynamic and unspecific binding of metal ions to nucleic acid strands (reviewed by Sigel & Pyle). Therefore, the identification of specific metal binding sites in catalytically active nucleic acids remains a challenging task. Despite some studies on metal ion dependency of the 10-23 DNAzyme, the coordination sites and the specific roles of metal ions are poorly understood. The 10-23 DNAzyme was obtained by *in vitro* selection in the presence of Mg^{2+} ions (Santoro & Joyce, 1997). Its activity strongly varies in the presence of different divalent metal cations, since the different ionic radii, the pK_a values of coordinating water molecules, and the affinity to functional groups tremendously affect the catalytic properties of the metal ion. Studies on the effect of metal ions on the DNAzyme activity have shown that divalent metal ions promote the reaction in the following order: Mn^{2+} (in 3-[4-(2-Hydroxyethyl)piperazin-1-yl]propane-1-sulfonic acid, EPPS) $>$ Pb^{2+}, Mg^{2+}, Ca^{2+} $>$ Cd^{2+} (in Tris buffer) $>$ Sr^{2+}, Ba^{2+}, Zn^{2+}, Co^{2+} (Santoro & Joyce, 1998). These findings correlate with the studies of Sugimoto *et al.*, where the metal ions enhance the cleavage reaction in the following order: $Mn^{2+} > Mg^{2+} > Ca^{2+} \gg Ba^{2+}$ in Tris buffer. The variation in activity may either be due to the occupation of different binding sites within the nucleic acids as reported for either Mg^{2+} (Scott *et al.*, 1996), Mn^{2+} (Scott *et al.*, 1996), or Cd^{2+} (Murray *et al.*, 2000) in the hammerhead ribozyme or the binding behavior with regard to inner- or outer-sphere contacts with the nucleic acid as reported for the same ribozyme with either Mg^{2+} (Chi *et al.*, 2008) or Mn^{2+} (Martick *et al.*, 2008). For all studies performed with Mn^{2+} at pH 7.5 one has to keep in mind that the actual Mn^{2+} concentration might be overestimated due to oxidation of the metal at pH 7.5 (Santoro & Joyce, 1998).

He *et al.* compared the observed activity of the 10-23 DNAzyme depending on the pK_a value of different metal ions and they found that the logarithm of the k_{obs} value was inversely proportional to the pK_a value. Folding of the DNAzyme in the presence of different Mg^{2+} concentrations has been analyzed using circular dichroism (CD) and Förster resonance energy transfer (FRET) in low ionic strength buffer, revealing three different folding events at Mg^{2+} concentrations of 0.5 mM (compact structure of the DNAzyme), 5 mM (proper positioning of binding arms to bind RNA) and 15 mM (completely organized catalytic domain) (Cieslak *et al.*,

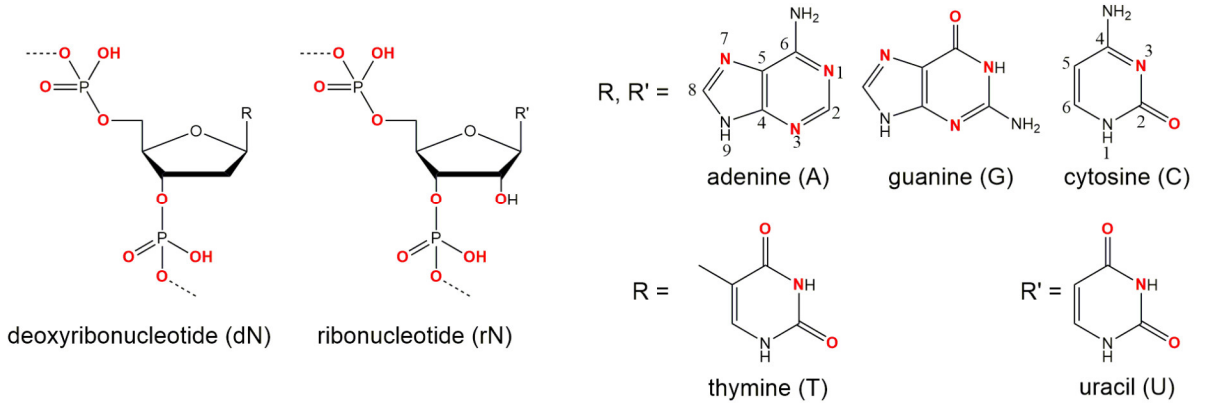


Figure 2.7: Potential metal interaction sites in DNA and RNA. The functional groups that are in principle capable of interacting with cations are highlighted in red. Besides the bridging and non-bridging oxygen atoms of the phosphate backbone of DNA and RNA and the 2' oxygen atom in RNA sequences, the nitrogen atoms N1, N3 and N7 of purine bases, the N3 nitrogen atom of cytosine and the carbonyl oxygen atoms at position C2 and C4 in pyrimidine bases and C6 in guanosine can function as donor atoms for the coordination of metal ions at physiological pH values as well as deprotonated $(N1)^-$ of guanosine and $(N3)^-$ of thymidine and uracil. The exocyclic amino groups also have the potential to participate in metal ion coordination as hydrogen-bond donors within the first coordination sphere (reviewed in Sigel & Pyle; and Lippert).

2003). They concluded that binding of Mg^{2+} to the 10-23 DNAzyme induces significant rearrangement of the catalytic loop leading to optimal folding of the catalyst. Besides what is known for small ribozymes (Curtis & Bartel, 2001; Murray *et al.*, 1998), in DNAzyme-mediated catalysis divalent metal ions cannot fully be replaced by monovalent metal ions. Moreover, our group recently found that the cleavage reaction of the 10-23 DNAzyme is drastically reduced at high ionic strength (Na^+ or K^+) in the presence of Mg^{2+}. Plashkevych & Chattopadhyaya investigated different 10-23 DNAzyme complexes with the RNA substrate in 2 ns MD simulations in the presence and absence of different metal ions. Controversially, in the presence of monovalent K^+ ions, the DNAzyme:RNA complex tends to form a typical B-type helix. This is surprising, since DNA:RNA duplexes are usually known to form A-type helices (Milman *et al.*, 1967). While the sugar pucker in the RNA strand of the duplex varies between 3'-endo (here termed *North*-type) and 2'-endo (here termed *South*-type) conformation, the DNA residues of the binding arms and the catalytic loop have the sugar moieties in the 2'-endo conformation, as in B-type DNA. In the presence of Mg^{2+} cations the complex is not only more stabilized, but also shows a significantly different structure compared to the one in the presence of only K^+ ions. Mg^{2+} ions lead to a significant change in the folding of the catalytic loop, that is no longer directed away from the DNA:RNA duplex, as found in the presence of K^+, but bent towards the DNA:RNA duplex, with a stretch of the catalytic domain being located close to the cleavage site. While in the presence of K^+, the sugar moieties of the flanking DNA strands predominantly show *South*-endo conformation, in the presence of Mg^{2+} *North*-endo conformation predominates. Comparing the structures of the native and mutated DNAzymes derived

from MD simulations, Plashkevych & Chattopadhyaya found that the presence of Mg^{2+} induces the formation of an electrostatic pocket by the catalytic loop and that the cleavage reaction may be assisted by three Mg^{2+} cations. Their findings are in agreement with a study by Victor *et al.*, which suggests a cooperative binding of three Mg^{2+}. Plashkevych & Chattopadhyaya reported that several cations are placed near the scissile phosphate and retained there during the simulation. However, the authors conclude that their model cannot explain the catalytic activity and proposed that a conformational change must happen prior to the cleavage reaction. A coarse-grained Brownian dynamics simulation by Kenward & Dorfman shows that the 10-23 DNAzyme bends its substrate away from the cleavage site, exposing the reactive site and buckling the DNAzyme catalytic loop. Breaker *et al.* have shown that—in general—metal ions may operate in the catalytic transesterification of the phosphodiester bonds in RNA substrates according to four main strategies that they termed: α-catalysis (facilitation of a proper geometric alignment for "in-line" nucleophilic attack), β-catalysis (charge neutralization on the non-bridging oxygens of the scissile phosphate bond), γ-catalysis (metal-assisted deprotonation of the reactive 2'-hydroxyl) and δ-catalysis (stabilization of the 5'-oxygen of the leaving group). The cleavage reaction of the 10-23 DNAzyme has also been studied with regard to the catalytic strategy (Breaker *et al.*, 2003b). A combinatorial analysis of Mg^{2+} and pH dependency reveals that the catalytic strategy of this DNAzyme exceeds the exclusive use of the combinatorial catalysis of α- (Soukup & Breaker, 1999) and γ-catalysis (Li & Breaker, 1999a), since the rate constant reached under suboptimal conditions exceeds the combined $\alpha\gamma$-speed limit (Breaker *et al.*, 2003b; Santoro & Joyce, 1998). However, since no thio effect was observed if the non-bridging phosphate oxygen at the cleavage site was replaced by a sulfur atom (Nawrot *et al.*, 2008), it is likely that the 10-23 DNAzyme does not use metal coordination to such an oxygen atom for rate enhancement, excluding β-catalysis (Emilsson *et al.*, 2003) as an additional strategy for the cleavage reactio (Breaker *et al.*, 2003b).

Critical DNAzyme positions and their potential involvement in DNAzyme catalysis

The exhaustive characterization of different 10-23 DNAzyme mutants and chemical modifications provide insights into the impact of various deoxynucleotides and functional groups on DNAzyme catalysis, even in the absence of high-resolution structural information. The mutagenesis studies by Zaborowska *et al.* and He *et al.* clearly demonstrate the relevance of the exocyclic 4-amino group in dC_3 and dC_{13} as well as the 6-keto group in dG_3. At position dG_{14}, the 2-amino group and the 6-keto group play an important role in DNAzyme catalysis. In addition, the free electron pairs of N1 or N7 at position 5 are crucial for the RNA cleavage activity. The effect of substitutions of deoxynucleotides by abasic deoxynucleotides or C3 spacers that add a high degree of flexibility to the DNAzyme loop also reveal that structural prearrangement of some of the 2'-deoxyribose moieties is of importance for the correct function of the 10-23 DNAzyme (Wang *et al.*, 2010a). The most striking observations have been made for

dT_8, which can be deleted without significantly decreasing the DNAzyme activity. Introduction of 2'-(R)-CH_3 and 2'-(S)-CH_3 modified purine nucleosides leads to a decrease in RNA cleavage activity. However, the reduction in activity was higher when the 2'-(R)-CH_3 derivative was introduced (Robaldo *et al.*, 2010). The activity was even more reduced when locked nucleic acid building blocks were introduced at position 8, suggesting that a certain degree of conformational flexibility at this position is crucial (Robaldo *et al.*, 2012). Position 8 is also of interest with regard to metal ion coordination. Based on results from cleavage assays with phosphorothioate substitutions at the non-bridging phosphate oxygen atoms analysing the thio and rescue effects in the presence of Mn^{2+} and Mg^{2+} it is reasonable to assume that the R_P oxygen at position $d9^P$ is directly involved in metal ion coordination (Nawrot *et al.*, 2007). Based on the Pearson acid-base concept (Pearson, 1968), a Mg^{2+} ion has a lower affinity to the polarizable sulfur-containing DNAzyme resulting in a lower cleavage rate. A recovery of the cleavage rate in the presence of Mn^{2+} instead of Mg^{2+} is called rescue effect. In addition to phosphorothioate substitution within the catalytic loop region, the effect of phosphorothioate substitution at the scissile phosphate in the RNA substrate has been investigated (He *et al.*, 2002; Nawrot *et al.*, 2008). Here, it was found that the cleavage of the S_P thio analog was only slightly reduced compared to the cleavage of the unmodified RNA substrate in the presence of Mg^{2+} ions. In contrast, the R_P phosphorothioate could not be cleaved, even in the presence of high Mg^{2+} concentrations (100 mM). However, this thio effect could not be rescued in the presence of Mn^{2+}, suggesting that there is no direct coordination of a metal ion to the non-bridging oxygen atoms at the scissile phosphodiester bond. This hypothesis is in agreement with earlier suggestions of He *et al.* and Breaker *et al.*. Summarizing the mechanistic results by Nawrot *et al.*, Breaker *et al.* as well as Santoro & Joyce, one can exclude β-catalysis as well as the exclusive use of α- and γ-catalysis. A catalytic mechanism, which is in agreement with the data of Breaker *et al.* and Nawrot *et al.*, proposes two metal ions that both bind to the scissile site of the RNA substrate: the first metal ion acts as a Lewis acid that directly coordinates to the 2'-oxygen atom and hereby promotes the deprotonation of the 2'-hydroxyl group (γ-catalysis). The proton then migrates to the R_P oxygen of the scissile phosphate. A second metal ion is coordinated to the 5'-oxygen of the leaving group to neutralize the negative charge during the cleavage of the P-O bond (δ-catalysis) (Figure 2.8).

2.7 Conflicting results

Although the influence of deoxynucleotide substitutions and modifications, deoxyribose and phosphate modifications within the catalytic loop and the binding arms of the 10-23 DNAzyme, as well as its tolerance towards different metal ions and pH conditions have been extensively studied, the mechanism by which the DNAzyme cleaves its specific RNA substrate as well as the role of different functional groups is only poorly understood. The main cause is arguably the absence of satisfactory structural data of the 10-23 DNAzyme. However, one major obstacle when it comes to the interpretation of different mutational studies is the low level of compa-

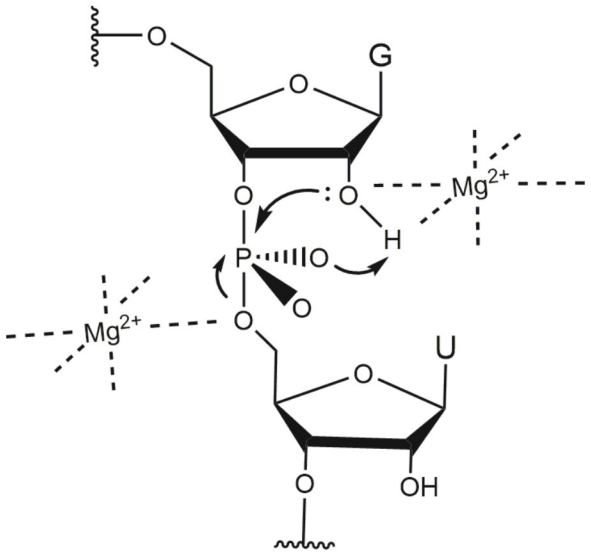


Figure 2.8: Proposed mechanistic model for the metal ion-mediated RNA cleavage catalysed by a 10-23 DNAzyme Two metal ions (here Mg^{2+}) are involved in a triester-like mechanism, that has originally been proposed for the RNA-cleavage reaction catalysed by a hammerhead ribozyme (Zhou *et al.*, 1998). One metal ion acts as a Lewis acid and coordinates to the 2'-OH group. Thereby, the OH bond is polarized and migration of the proton to the R_P oxygen is facilitated. The second metal ion coordinates to the 5'-oxygen leaving group and neutralizes the negative charge during the cleaveage of the P-O bond. Based on Nawrot *et al.*.

rability of the experimental setup. While some experiments were designed to determine only the yield after a given period of time (Y_t) of the cleaved substrate, others were designed to determine rate constants (k_{obs}). Furthermore, experiments vary in Mg^{2+} concentrations used to facilitate DNA-mediated RNA cleavage or the nature of the substrate: while some experiments were performed in the presence of *all*-RNA substrates other experiments were carried out using chimeric DNA/RNA substrates, with only the nucleotide at the cleavage site being a ribonucleotide. Also, DNAzyme and RNA substrate design can have a massive influence on the kinetic behavior of the DNAzyme, since the length and sequence dramatically influence the dissociation constants of the DNAzyme:RNA complex. In addition, temperature, monovalent cation levels as well as the type of reaction buffer can have an impact on DNAzyme activity. In Table 2.3, we summarize the experimental conditions used in the most important studies on DNAzyme activity.

2.8 Conclusion

Efforts to elucidate the reaction mechanism are essential for attempts to use DNAzymes for therapeutic applications. Structural information on DNAzymes is scarce, therefore, we rely on functional data for insights into DNAzyme catalysis.

The data summarized in this review provides interesting insights into the role of the different regions of the 10-23 DNAzyme. The binding arms of the DNAzyme are responsible for substrate recognition and form a double helical structure, while the catalytic loop forms an unusual three-

Table 2.3: Conditions used in selected important kinetic studies on the 10-23 DNAzyme with short substrates *in vitro*.

Reference	Substrate	Length of binding arms	Mg^{2+} [mM]	NaCl [mM]	RNA	DNAzyme	Duration and temperature	Output
Cairns *et al.*	RNA	7/7	10	150	0.2 µM	0.5–500 nM	various timepoints, 37°C	k_{obs}, single and multiple turnover
Cairns *et al.*	RNA	variable	10	–	0.04 µM	0.32 µM	various timepoints, 37°C	k_{obs}, single turnover
Cairns *et al.*	RNA	9/9	10	–	0.6 µM	5 µM	various timepoints, up to 1 h; 37°C	k_{obs}, single turnover
He *et al.*	RNA	7/7	25	–	0.06 - 0.1 µM	0.5-10 µM	various timepoints, 37°C	k_{cat}, multiple turnover
He *et al.*	DNA /w RNA	9/9	6 (2)	–	0.02 µM (0.2 µM)	2 µM (0.02 µM)	various timepoints, 37°C	k_{obs}, single turnover (multiple turnover)
Li *et al.*	DNA /w RNA	9/9	0.5-10	–	0.02 µM	2 µM	various timepoints, 37°C	k_{obs}, single turnover
Li *et al.*	DNA /w RNA	9/9	20	–	0.02 µM	2 µM	various timepoints, 37°C	k_{obs}, single turnover
Li *et al.*	DNA /w RNA	9/9	2.18	–	0.02 µM	2 µM	various timepoints, 37°C	k_{obs}, single turnover

*c*ontinued on next page

Table 2.3 – continued from previous page

Reference	Substrate	Length of binding arms	Mg^{2+} [mM]	NaCl [mM]	RNA	DNAzyme	Duration and temperature	Output
Nawrot *et al.*	DNA /w RNA	8/8	3	100	0.1 µM	10 µM	various timepoints, 37°C	k_{obs}, single turnover
Nawrot *et al.*	DNA /w RNA	8/8	0.02	100	0.1 µM	10 µM	various timepoints, 37°C	k_{obs}, single turnover
Okumoto & Sugimoto	RNA	6/8	25	–	0.01-1 µM	0.001 µM	various timepoints, 37°C	k_{cat}, multiple turnover
Robaldo *et al.*	RNA	8/8	0.5-10	–	0.5 µM	5 µM	various timepoints, 37°C	k_{obs}, single turnover
Robaldo *et al.*	RNA	8/8	0.5-10	–	0.5 µM	0.5 µM	various timepoints, 37°C	k_{obs}, single turnover
Santoro & Joyce	RNA	8/8	variable	variable	variable	variable	various timepoints, 37°C	k_{obs} and k_{cat}, single and multiple turnover
Santoro & Joyce	RNA	8/8	2	150	variable	variable	various timepoints, 37°C	k_{obs} and k_{cat}, single and multiple turnover
Schubert *et al.*	RNA	9/9	10	100	1 µM	0.01 µM	20 min, 37°C	initial velocity, multiple turnover

*c*ontinued on next page

Table 2.3 – continued from previous page

Reference	Substrate	Length of binding arms	Mg^{2+} [mM]	NaCl [mM]	RNA	DNAzyme	Duration and temperature	Output
Sugimoto *et al.*	RNA	8/6	25	–	5 nM (tenfold excess)	1 µM (n.a.)	various timepoints, 37 °C	Y_t (k_{cat}), single turnover (multiple turnover)
Victor *et al.*	RNA	9/9	0.1-1	100	0.1 µM	0.1 µM	various timepoints, 37 °C	k_{cat}, single turnover
Wang *et al.*	DNA /w RNA	9/9	500	200	0.005 µM	0.5 µM	various timepoints, 37 °C	Y_t and k_{obs}, single turnover
Wang *et al.*	DNA /w RNA	9/9	2	–	0.02 µM	2 µM	various timepoints, –	Y_t and k_{obs}, single turnover
Zaborowska *et al.*	RNA (RNA)	9/9	10	–	0.1 µM	1 µM	20 min, 37 °C	Y_t, single turnover
Zaborowska *et al.*	RNA (RNA)	9/9 (9/9)	10 (10)	– (–)	1 pmol (0.01- 1 µM)	10 pmol (0.001 µM)	20 min (various timepoints), 37 °C	Y_t (k_{cat}, multiple turnover)
Zhu *et al.*	DNA /w RNA	9/9	2	–	n.a.	100-fold excess	Various timepoints, 37 °C	k_{obs}, single turnover
Zhu *et al.*	DNA /w RNA	9/9	2	–	0.002 µM	2 µM	various timepoints, 37 °C	k_{obs}, single turnover

*e*nd of Table 2.3

dimensional conformation.

Taking a closer look at the cleavage site, we conclude that the neighboring nucleotides of the scissile A or G in the RNA and DNA strand play an important role in catalysis. All mutations and modifications at these positions indicate a flexibility or equilibrium of rN_{-1}–dA_{15} and rN_{+1}–dG_1. These nucleotides might act as hinges between the catalytic loop and the DNA-RNA arms. They allow the rigid catalytic loop to move against the rigid double helices and adopt a particular conformation that enable the cleavage reaction.

Studies focusing on the residues within the catalytic loop have shown that the exocyclic groups of the nucleotides dG_{14} and dG_6 are essential for DNAzyme function. Deletion of the 6-keto group in either dG_{14} or dG_6 leads to a dramatic decrease in catalytic activity, while the substitution of the oxygen atom of the keto-group by sulfur results in a decreased reaction rate in the presence of Mg^{2+}. The activity can be rescued in the presence of Mn^{2+}, suggesting the interaction of this atom with the divalent metal ion during hydrolysis. Deletion of the 2-amino group in dG_{14} also reduces the activity of the 10-23 DNAzyme, suggesting that these deoxynucleotides play an important role in the cleavage of RNA catalyzed by the 10-23 DNAzyme. Taken together, we have summarized molecular features of the 10-23 DNAzyme in this review in order to contribute to a better understanding of these fascinating biocatalysts. Our detailed analysis reveals important aspects to be considered in future attempts to elucidate the structure of DNAzymes in complex with the RNA substrate and unravel the mechanism of the reaction catalyzed by the 10-23 DNAzyme.

3 Influence of monovalent metal ions on metal binding and catalytic activity of the 10-23 DNAzyme

This chapter reflects content of the following publication.

3.1 Publication information

Hannah Rosenbach, Jan Borggräfe, Julian Victor, Christine Wübben, Olav Schiemann, Gerhard Steger, Manuel Etzkorn, and Ingrid Span

Submitted to: **Biological Chemistry**, 31.05.2020

3.2 Abstract

Deoxyribozymes (DNAzymes) are single-stranded DNA molecules that catalyze a broad range of chemical reactions. The 10-23 DNAzyme catalyzes the cleavage of RNA strands and can be designed to cleave essentially any target RNA, which makes it particularly interesting for therapeutic and biosensing applications. The activity of this DNAzyme *in vitro* is considerably higher than in cells, which was suggested to be a result of the low intracellular concentration of bioavailable divalent cations. While the interaction of the 10-23 DNAzyme with divalent metal ions was studied extensively, the influence of monovalent metal ions on its activity remains poorly understood. Here, we characterize the influence of monovalent and divalent cations on the 10-23 DNAzyme utilizing functional and biophysical techniques. Our results show that sodium and potassium ions affect the binding of divalent metal ions to the DNAzyme:RNA complex and considerably modulate the reaction rates of RNA cleavage. We observe an opposite effect of high levels of sodium and potassium concentrations on Mg^{2+}- and Mn^{2+}-induced reactions, revealing a different interplay of these metals in catalysis. Based on these findings, we propose a model for the interaction of metal ions with the DNAzyme:RNA complex.

3.3 Introduction

DNA molecules, like RNA molecules and proteins, are capable of folding into well-defined three-dimensional structures that can enzymatically catalyze chemical reactions. While protein enzymes and ribozymes occur naturally, deoxyribozymes (DNAzymes) have not been discovered in nature. The first DNAzyme was identified by Breaker & Joyce in 1994 using *in vitro* selection. They isolated a single-stranded DNA sequence from a random-sequence DNA pool that is capable of cleaving an RNA phosphodiester bond in a Pb^{2+}-dependent reaction. Since the selection of the first DNAzyme the number of DNA catalysts has dramatically increased and their catalytic repertoire has exceeded the cleavage of RNA substrates and now includes DNA phosphorylation (Li & Breaker, 1999b), peroxidation (Travascio *et al.*, 1998), thymine dimer photoreversion (Chinnapen & Sen, 2004) and DNA cleavage (Carmi *et al.*, 1998). RNA-cleaving DNAzymes are particularly attractive for therapeutic applications (Fokina *et al.*, 2015). Like RNA interference (RNAi)-based therapeutics (Bobbin & Rossi, 2016), the 10-23 DNAzyme has been considered for the treatment of diseases in which post-transcriptional gene silencing could be beneficial. They have been designed to target a broad variety of diseases, including allergy/asthma (Krug *et al.*, 2015), prion-protein related disorders (Victor *et al.*, 2018), cardiovascular diseases (Santiago *et al.*, 1999a), and cancer (Cho *et al.*, 2013; Wu *et al.*, 1999). In comparison to their RNA counterparts, DNA molecules show favorable characteristics, including higher stability and cost-effective synthesis. However, convincing evidence for DNAzyme activity inside the cell that exceeds antisense effects could not be demonstrated in these studies (Young *et al.*, 2010). A recent study suggests that the poor performance of the 10-23 DNAzyme*in vivo* is related to the low intracellular levels of accessible Mg^{2+}, which is crucial for DNAzyme catalysis (Victor *et al.*, 2018). Furthermore, the metal-dependence of DNAzymes makes them highly valuable for for the detection of metal ions as environmental sensors (McGhee *et al.*, 2017). The 10-23 DNAzyme is the most prominent RNA-cleaving DNAzyme. It cleaves its RNA substrate between a purine and a pyrimidine nucleotide in the presence of divalent metal ions, including Mg^{2+} or Mn^{2+} (Santoro & Joyce, 1997). The DNAzyme consists of a catalytic loop with a fixed sequence of 15 nucleotides, which is flanked by two substrate binding arms. These arms can be designed to bind virtually any RNA sequence through Watson-Crick base pairing. The first step in the reaction is the hybridization of the DNAzyme and the RNA substrate, followed by cleavage of the RNA strand and release of the two products (Figure S3.1A). Under optimum conditions the 10-23 DNAzyme exhibits a catalytic rate that is comparable to those of naturally occurring biocatalysts, such as the hammerhead ribozyme or the RNA-cleaving protein enzyme RNase A (Santoro & Joyce, 1998). Although the biochemical and biophysical properties of the 10-23 DNAzyme have been extensively studied for more than two decades, the mechanism of DNAzyme-mediated RNA cleavage and the role of the metal ions participating in the reaction are not fully understood. Studies on the pH dependence of substrate cleavage suggest a single deprotonation event, which most likely occurs at the 2'-OH group of the RNA nucleotide next to the scissile bond and is the rate limiting step in the RNA cleavage (Santoro & Joyce, 1998). Experimental evidence suggest that the mechanism of DNA-mediated catalysis is dif-

ferent from RNA-mediated catalysis. Particular evidence comes from the different behavior of the 10-23 DNAzyme and the hammerhead ribozyme towards lanthanide ions in the presence of Mg^{2+} (He *et al.*, 2002). Unlike ribozymes (Perrotta & Been, 2006; Schnabl & Sigel, 2010), RNA-cleaving DNAzymes are not active in the sole presence of monovalent cations. Previous studies focusing on the effect of different divalent metal ions on the 10-23 DNAzyme have shown that the different ions enhance the cleavage reaction in the following order: $Mn^{2+} > Mg^{2+} > Ca^{2+} \gg Ba^{2+}$ in tris(hydroxymethyl)aminomethane (Tris) buffer (Sugimoto *et al.*, 1999). This raises the question why these metal ions have such a different effect on the catalytic rate? He *et al.* proposed that the logarithm of the cleavage rate linearly decreased with the pK_a value of the divalent metal ions that promote the cleavage reaction. A similar correlation has been found for the hammerhead ribozyme (Dahm *et al.*, 1993; He *et al.*, 2002; Sawata *et al.*, 1993). The majority of previous studies have been focusing on the interaction of divalent metal ions with the 10-23 DNAzyme, however, monovalent cations, including sodium and potassium, have not been investigated carefully, despite their high levels inside the cell. In this study, we investigated the influence of sodium or potassium ion levels on the magnesium and manganese ion dependency of a 10-23 DNAzyme variant designed to cleave the mRNA of the human prion protein (PrP) (Victor *et al.*, 2018). To obtain deeper insights into metal binding and catalysis, we apply a set of complementary biophysical techniques including Förster resonance energy transfer (FRET)-based kinetic measurements, isothermal titration calorimetry (ITC), nuclear magnetic resonance (NMR), and electron paramagnetic resonance (EPR) spectroscopy. Our results show that monovalent cations such as sodium or potassium ions have an inhibitory effect on *in vitro* DNAzyme activity in the presence of magnesium ions. This inhibition presumably contributes to the poor performance of DNAzymes inside the cell, where monovalent cation levels are high compared to divalent metal ion levels. Moreover, we show that while DNAzymes are not active in the sole presence of monovalent cations, increasing monovalent metal ion levels have a beneficial effect on DNAzyme activity in the presence of manganese ions. The opposite behavior of DNAzyme activity induced by magnesium or manganese ions points to different modes of action of the respective metal ions.

3.4 Material and Methods

3.4.1 Oligonucleotide sequences

DNA and RNA oligonucleotides were acquired from biomers.net GmbH (Ulm, Germany) and BioSpring (Frankfurt a. Main Germany). In the stabilized RNA substrate T839-2'F the 2'OH group of the ribose sugar of the guanosine at the cleavage site was substituted by a fluorine atom. The not stabilized RNA substrate was labeled with a 6-Carboxyfluorescein (6-FAM) molecule at the 5' end and the Black Hole Quencher-1 (BHQ1) at the 3' end. The DNAzyme was designed to specifically bind the human prion protein mRNA (Victor *et al.*, 2018). The sequences of the DNAzyme and RNA substrates are listed in Table 3.1.

Table 3.1: Oligonucleotides used in this work.

Name	Sequence 5′ – 3′
T839-FRET	6-FAM-ACA UGC ACC GUU ACC CCA A-BHQ1
T839-FAM	6-FAM-AAA ACA UGC ACC GUU ACC CCA A
T839-2'F	ACA UGC ACC $G_{2'F}$UU ACC CCA A
DNA839	TTG GGG TAA CGG TGC ATG T
DNAzyme839	TTG GGG TAA GGC TAG CTA CAA CGA GGT GCA TGT

3.4.2 Activity assay with FRET-labeled RNA substrates

Cleavage reactions were carried out with 0.1 µM RNA and DNAzyme in 50 mM Tris-HCl pH 7.5 with different concentrations of Na^+, K^+, Mg^{2+}, Mn^{2+} at 37 °C. In the following the divalent metal ion is abbreviated as M^{2+}. 0.8 µM of RNA substrate and DNAzyme were denatured in buffer in the absence of Na^+ and/or K^+ and M^{2+} at 73 °C for 5 min, cooled down to room temperature for 15 min, and 20 µl of the solution were pipetted into the wells of a 384 well non-binding microplate (Greiner Bio-One, Kremsmünster, Austria). 10 µl of stock solutions with different NaCl concentrations in Tris-HCl pH 7.5 were added. The plate was sealed with tape (Polyolefine Acrylate, Thermo Scientific, Waltham, MA, USA), placed inside the plate reader (CLARIOStar, BMG LABTECH, Ortenberg, Germany) and equilibrated to 37 °C for 30 min. The reaction was started by the injection of $MgCl_2$ in buffer. Data points were obtained every 5, 3, or 2 s, depending on the $MgCl_2$ or $MnCl_2$ concentration and therefore on the speed of the reaction. Excitation and emission wavelength were 484 nm and 530 nm, respectively.

3.4.3 Activity assay with Fluorescein-labeled RNA substrates followed by denaturing polyacrylamide gel electrophoresis

Activity assays were performed with 0.4 µM fluorescein-labeled RNA substrate and 0.4 µM DNAyzme in 50 mM Tris-HCl pH 7.5 with 0.1 mM EDTA and 500 or 1000 mM NaCl in the presence and absence of 1 mM $MgCl_2$ for 3 h at 37 °C. RNA and DNAzymes were denatured in buffer in the absence of $MgCl_2$ at 73 °C for 5 min, followed by an incubation for 10 min at room temperature. Subsequently, the reaction was started by adding $MgCl_2$. Separation of the samples was carried out on 18% polyacrylamide gels with 7 M urea buffered with Tris-borate EDTA buffer (TBE) for 1 h at 20 W. Visualization of fluorescein-labeled RNA substrates was carried out by fluorescence detection. Images were acquired using the ChemiDoc MP System (Bio-Rad, Hercules, CA, USA).

3.4.4 ITC measurements

ITC was performed in 50 mM Tris buffer pH 7.5, with or without 100 mM NaCl, at 30°C on a Microcal iTC200 calorimeter (GE Healthcare, Chicaco, IL, USA). Affinities were determined with the pre-formed complex of 10-23 DNAzyme and RNA substrate as titrant in the cell at a concentration of 100 µM, and 10 mM $MgCl_2$ or 3 mM $MnCl_2$ as titrant in the syringe. Dissociation constants were obtained from a non-linear least-squares fit to either a 1:1 binding model or a *n*:1 binding model with *n* identical Mg^{2+}-binding sites per 10-23 DNAzyme:RNA complex using the Origin software (MicroCal) provided with the calorimeter.

3.4.5 Continuous wave (cw) EPR experiments

Continuous wave (cw) X-band EPR spectra were recorded on a Bruker EMX micro spectrometer (Bruker BioSpin, Rheinstetten, Germany) equipped with an ER 4103TM resonator (Bruker BioSpin, Rheinstetten, Germany). Titrations with Mn^{2+} were performed at room temperature with 40 µM of the DNAzyme and the stabilized RNA substrate in 50 mM Tris-HCl pH 7.5 in the presence and absence of NaCl (0 mM, 100 mM and 1000 mM) using a flat cell with a volume of 500 µl. For each data point, 2 µl $MnCl_2$ stock solution were added into the 500 µl of the initial sample. The spectra were measured at room temperature with a modulation frequency of 100 kHz, a modulation amplitude of 5.0 G, a microwave frequency of 9.79 GHz, and 1440 points in the field interval of 2890-4090 G. The resulting Mn^{2+} EPR signals were baseline corrected and doubly integrated (DI). The individual Q-values (∼2900-3500) and microwave powers (1.77 mW) were taken into account using (3.1).

$$DI = c \cdot Q \cdot \sqrt{P} \tag{3.1}$$

with DI = double integral, c = constant, Q = calculated Q-value from the spectrometer, and P = microwave power (Eaton *et al.*, 2010).

To determine the concentration of bound Mn^{2+} to the DNAzyme:RNA complex, the signal intensities of the increasing Mn^{2+} concentration in the presence of constant DNAzyme:RNA complex were compared to the signal intensities of standard samples containing 200 µM, 500 µM and 12.5 mM Mn^{2+} in the absence of nucleic acids. The reference samples of free Mn^{2+} in solution were measured using a freshly prepared stock solution ($MnCl_2$, 1 M, 10x1 mL, Sigma Aldrich, Darmstadt, Germany). The concentrations of these reference samples were checked via the spin counting tool of the Bruker EMXnano spectrometer (Bruker BioSpin, Rheinstetten, Germany). To construct Mn^{2+} binding isotherms, the concentrations of bound Mn^{2+} per DNAzyme:RNA complex were plotted against the concentration of free Mn^{2+} in the solution. These binding isotherms from titrations without NaCl (Figure 3.3A) were fit to (3.2) that assumes *n* independent, non-interacting Mn^{2+}-binding sites on a complex and each site has an identical association constant *K* (Poland, 1978):

$$\frac{[Mn^{2+}_{bound}]}{[Dz{:}RNA]} = \sum_{i=1}^{n} \frac{i \cdot [Mn^{2+}_{free}]}{\frac{1}{K} \cdot [Mn^{2+}_{free}]} \tag{3.2}$$

The binding isotherms in presence of 100 or 1000 mM NaCl (Figures 3.3B and C) were fit to the Hill equation (3.3) with Hill coefficient n, m binding sites on a complex, and apparent association constant K:

$$\frac{[Mn^{2+}_{bound}]}{[m \cdot Dz{:}RNA]} = \frac{K \cdot [Mn^{2+}_{total}]}{1 + K \cdot [Mn^{2+}_{total}]} \tag{3.3}$$

The experimental data was fitted using python v3.6.9[1] with lmfit v0.9.7 (Newville *et al.*, 2019); parameter space was scanned using the "brute" force grid-search method and obtained solutions refined with the leastsq method from lmfit.

3.4.6 Global fit of FRET data

Experimental data sets, for example FRET curves with c(Mg^{2+}) = [0.5, 1.0, 2.0, 3.0, 5.0] mM and c(Na^{+}) = [0, 20, 40, 60, 80, 100, 250, 500, 750, 1000] mM, were adjusted to common t = 0 and normalized to $0 \leqslant F \leqslant 1$ using gle[2]. Using python, the differential equations resulting from (3.4), (3.5) and (3.6), (3.7), respectively, were numerically integrated using odeint of scipy (Virtanen *et al.*, 2020) and parameters were fit to the experimental data using lmfit v0.9.7 (Newville *et al.*, 2019). Parameter space was scanned using the "brute" force grid-search method (for example see Figure S3.8) and obtained solutions refined with the leastsq method from lmfit. Optimized parameter values are summarized in Table 3.2; curves are shown in Figure 3.7.

[1] `https://www.python.org/`
[2] `http://glx.sourceforge.net/`

Table 3.2: Binding constants ($K_{M^{2+}} = k_{+2|+5}/k_{-2|-5}$, $K_{Na^+} = k_{+1|+4}/k_{-1|-4}$) were fixed to the given values. Pearson's correlation coefficient r of the experimental and fitted data was determined using numpy's[3] corrcoef.

Parameter	Unit	Optimal values	
		Mg^{2+}	Mn^{2+}
$K_{M^{2+}}$	1/mM	4.0	4.0
$K_{D,M^{2+}}$	mM	0.25	0.25
K_{Na^+}	1/mM	0.01	0.0001
K_{D,N^+}	mM	100	$1{\cdot}10^4$
$k_{1\|4}$	1/(mM·s)	$4.7{\cdot}10^{-3}$	$7.0{\cdot}10^4$
$k_{2\|5}$	1/(mM·s)	0.0195	10.3
$k_{3\|6}$	1/s	0.0272	101.0
f_{coeff}		9922.2	9987.0
r^2		0.910	0.927

3.4.7 NMR studies

NMR experiments were carried out on a Bruker Avance III HD^+ 600 MHz spectrometer with an inverse triple resonance cryo-probe. Two-dimensional [^{1}H,^{1}H]-TOCSY spectra were recorded at 37 °C in 3 mm sample tubes with 200 µl sample volume. The concentration of nucleic acid complex (DNAzyme839:T839-2'F, DNA839:T839-2'F) was 100 µM in 50 mM Tris-HCl, pH 7.5 with 100 or 500 mM NaCl and 10% (v/v) D_2O as deuterium lock. For $MgCl_2$ titration studies, 2 µl of $MgCl_2$ stock solution with different concentrations were sequentially added to the sample to final $MgCl_2$ concentrations of 0.25, 0.5, 1, 2, 5, 10, 20 and 50 mM. All data was processed with TopSpin 4.0.6 (Bruker), all spectra analyzed with CARA (ETH Zürich) and plotted with Sparky 3.114 (UCSF). CSP and peak intensity data were both fitted to a 1:1 binding model using OriginPro 9.

3.5 Results

3.5.1 Influence of monovalent metal ions on the binding of divalent metal ions to the DNAzyme:RNA complex

It is commonly known that monovalent metal ions such as Na^+ and K^+ are associated with nucleic acids in living cells. Usually they play an unspecific role as general charge screeners, rather than establishing a close packing of the negatively charged phosphate backbone (Pechlaner & Sigel, 2012). Furthermore, monovalent cations can have a strong impact on the interaction of divalent cations with nucleic acids (Record Jr. *et al.*, 1976). We investigated the effects

of the monovalent cations Na^+ and K^+ on the activity of a 10-23 DNAzyme variant designed to cleave the mRNA of the human prion protein (PrP) (Victor *et al.*, 2018). First, we characterized the influence of increasing Na^+ levels on the interaction of Mg^{2+} with the pre-catalytic DNAzyme:RNA complex using isothermal titration calorimetric (ITC) measurements. ITC is an approach for determining the heat energy associated with a molecular interaction (Ladbury & Chowdhry, 1996). To this end, we carried out a $MgCl_2$ titration to pre-formed DNAyzme:RNA complex pre-equilibrated in the absence or presence of 100 mM Na^+ (Figure 3.1). In order to prevent catalytic turnover during the measurements, we used a 2-fluoro-stabilized RNA substrate (T839-2'F), in which the exchange of a hydroxyl group at the cleavage site by a fluorine prevents the cleavage by the 10-23 DNAzyme (Figure S3.1B). Assuming a simple 1:1 binding model, our ITC data show that Mg^{2+}-binding to the DNAyzme:RNA complex occurs with high micromolar binding affinities ($K_D = (410 \pm 30)$ µM) and that binding is about twofold weaker in the presence of 100 mM Na^+ ($K_D = (870 \pm 30)$ µM) (Figure 3.1C and D). Under both conditions slight deviations of the fit to the experimental data can be compensated by using a model with multiple Mg^{2+}-binding sites in an *n*:1 binding mode (Figure 3.1E and F), suggesting the presence of multiple Mg^{2+}-binding sites. In addition, the results show that increasing levels of Na^+ have a negative effect on the binding affinity of Mg^{2+} to the DNAzyme:RNA complex. We performed the same experiments with Mn^{2+} in place of Mg^{2+}, however, the measurements led to inconclusive data pointing to different thermodynamic features of the Mn^{2+}-DNAzyme interaction (Figure S3.2).

To be able to detect and distinguish the effects of possible different Mg^{2+}-binding sites of the DNAzyme:RNA complex we utilized NMR spectroscopy. Homonuclear 2D [^{1}H,^{1}H] TOCSY NMR spectra obtained from the DNAzyme in complex with its stabilized RNA substrate show well resolved peaks that can directly be separated into nucleotides in the substrate binding arms and in the catalytic loop region via a simple comparison to a spectrum recorded on a complementary RNA:DNA helix without catalytic loop (Figure S3.3). Titration experiments with $MgCl_2$ concentrations ranging from 0 to 50 mM in the presence of 100 mM Na^+ clearly show Mg^{2+}-dependent chemical shift perturbations (CSP) for a number of nucleotides in the binding arms (Figure 3.2). Interestingly, while the CSPs observed for Mg^{2+} binding to the arms are indicative of fast exchanging interactions (NMR fast exchange regime), Mg^{2+} binding to the catalytic loop shows slower exchanging interactions (NMR intermediate exchange regime), which usually indicates a stronger binding and results in disappearing of effected peaks (Figure 3.2A). In general, the detected CSPs allow to calculate nucleotide-specific Mg^{2+}-binding affinities for each resolved peak. While the magnitude of the CSPs clearly varies between different nucleotides, the Mg^{2+}-binding affinity obtained from the CPSs is very similar for all resolved peaks that can be attributed to the binding arms and a dissociation constant $K_D = 2.2$ mM can be determined for the substrate binding arms of the stabilized DNAzyme:RNA complex (Figure 3.2B, exemplified for peak indicated in Figure 3.2A). Since all peaks show similar behavior and the CSPs cannot distinguish between direct interactions with the Mg^{2+} ion or structural changes relayed from a binding site distant to the observed peak, it is not possible to deduct reliable information about the number of different binding sites within the binding arms. The observed different be-

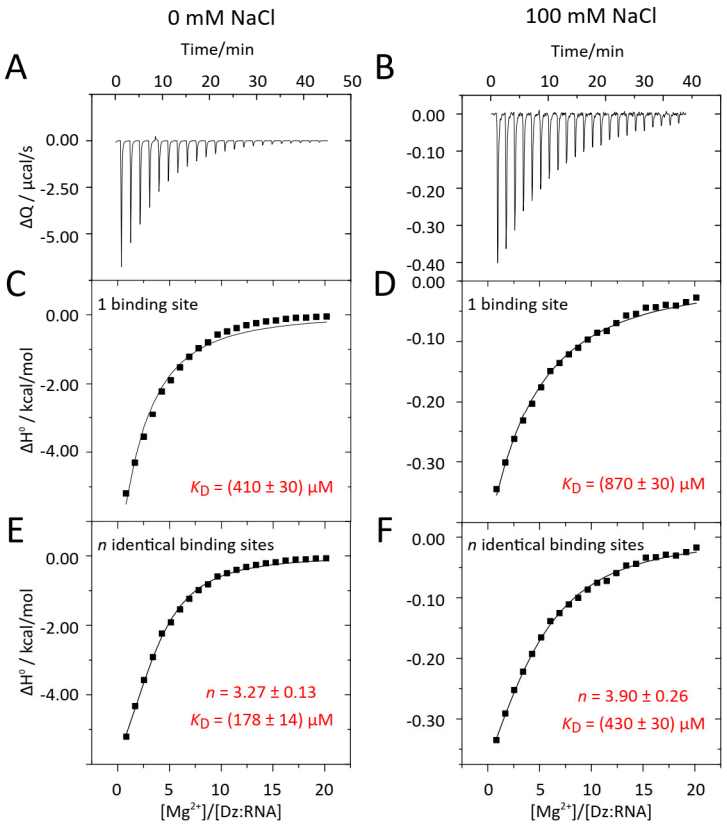


Figure 3.1: Interaction between Mg^{2+} and the DNAzyme:RNA complex at low (A, C, E) and high ionic strength (B, D, F) determined by ITC measurements. (A, B) Raw ITC data for the titration of Mg^{2+} (using a 10 mM stock solution) and DNAzyme:RNA complex (100 µM) at 30 °C after subtraction of the integration baseline. (C, D, E, F) Titration plots derived from the integrated raw data shown in (A) and (B). The solid lines represent the best fit to the data according to a single (C, D) and a multiple non-interacting (E, F) Mg^{2+} binding site model, respectively.

havior, i. e. fast vs. intermediate exchange, however, clearly indicates the presence of different Mg^{2+}-binding sites for the binding arms and catalytic loop of the DNAzyme. While the signal disappearance due to exchange contributions for the nucleotides in the catalytic loop prevents an accurate determination of the Mg^{2+}-binding affinity for this binding site, the data points to a considerably stronger Mg^{2+}-binding in the catalytic loop. A simple fit of the peak intensity decay suggests an up to tenfold stronger Mg^{2+}-binding with an exchange-biased lower limit for K_D in the range of 0.3 mM (Figure 3.2C).

Since ITC measurements with Mn^{2+} led to inconclusive results (Figure S3.2) and NMR spectroscopy of the Mn^{2+} binding is limited by its paramagnetic nature, we performed EPR spectroscopic measurements to also obtain insights into the effect of Na^{+} on the binding of Mn^{2+} to the DNAzyme:RNA complex. EPR spectroscopy is frequently used to directly extract the number of Mn^{2+} binding sites. Mn^{2+} is by far less abundant than Mg^{2+} in cells, however, it has similar physicochemical properties. Thus, it is often used as a Mg^{2+} mimick in combination with EPR spectroscopy (Horton *et al.*, 1998; Kisseleva *et al.*, 2005; Schiemann *et al.*, 2003). Since free paramagnetic $[Mn(H_2O)_6]^{2+}$ almost exclusively contribute to the EPR signal, the signal intensity can be converted to DNAzyme-bound and free Mn^{2+} concentrations using a calibration curve.

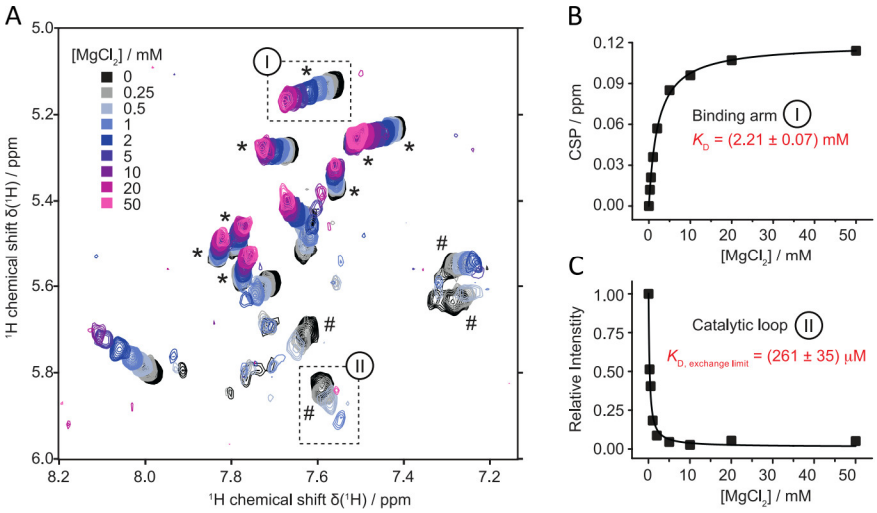


Figure 3.2: Site-resolved insights into the effects of Mg^{2+} binding to the stabilized DNAzyme:RNA complex. (A) Overlay of 2D [^{1}H,^{1}H] TOCSY NMR spectra of stabilized DNAzyme:RNA complex in presence of 100 mM Na^+ with indicated concentrations of Mg^{2+}. Different exchange rates are found for nucleotides within the binding arms (marked with $*$) and the catalytic loop (marked with #; see Supplementary Figure S3.3 for classification of arm and loop nucleotides). (B) Chemical shift perturbations (CSP) of indicated peak in (A) representing fast Mg^{2+}-induced exchange process of a nucleotide in the binding arms. The CSPs allow for accurate determination of the Mg^{2+}-binding affinity for this region ($K_D = 2.2$ mM). (C) Signals of the catalytic loop show strong Mg^{2+}-dependent signal loss due to intermediate exchange processes indicative of stronger interactions of this region with Mg^{2+}. Due to the NMR-unfavorable intermediate exchange only an exchange limited minimal K_D value in the range of 0.3 mM can be estimated.

We added Mn^{2+} incrementally to a pre-formed DNAzyme:RNA complex in the presence of 0, 100, and 1000 mM Na^+. The resulting Mn^{2+}-binding isotherms (Figure 3.3) show that the best fit for our data in the absence of Na^+ is provided by a simple model assuming a non-cooperative binding of six Mn^{2+} to the DNAzyme:RNA complex (Figure 3.3A). Binding of Mn^{2+} to the nucleic acid complex in the presence of either 100 mM or 1000 mM NaCl leads to a sigmoidal binding isotherm, suggesting a cooperative Mn^{2+} binding with a Hill coefficient of > 1 (Figure 3.3B, C; S3.4). This strongly indicates cooperative Mn^{2+} binding to the DNAzyme:RNA complex.

Taken together our data show that increasing levels of Na^+ lead to a decreasing binding affinity of Mg^{2+} to the DNAzyme:RNA complex. Further, we demonstrated that the binding of Mg^{2+} to the catalytic loop of the DNAzyme is considerably stronger than to the binding arms of the complex. For the binding of Mn^{2+} to the DNAzyme:RNA complex, we observed that the presence of Na^+ induces cooperative effects on Mn^{2+} binding.

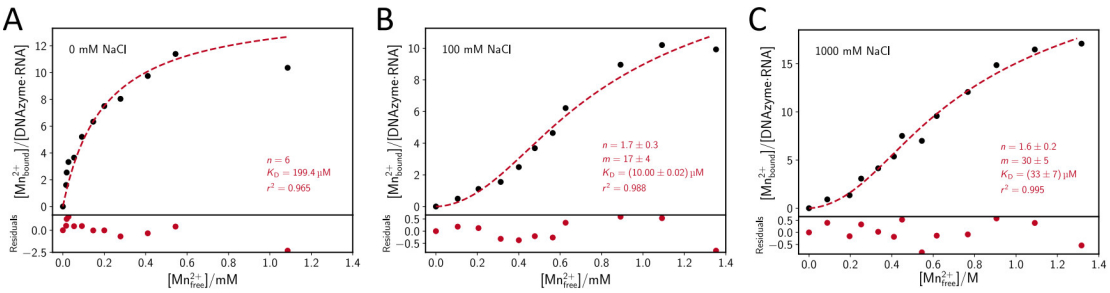


Figure 3.3: Binding isotherms for Mn^{2+} and the 10-23 DNAzyme:RNA complex determined by EPR spectroscopy. The binding isotherms were obtained obtained with no added Na^+ (A), 100 mM (B) or 1000 mM Na^+ (C) with 40 µM DNAzyme:RNA complex. The dashed curves are obtained by fitting the data to either a model that assumes n independent, non-interacting Mn^{2+}-binding sites, each with identical dissociation constant K_D (A; see (3.2)) or to the Hill equation (B, C; see (3.3)) with Hill coefficient n and m binding sites per DNAzyme:RNA complex; r^2 is the square of Pearson's correlation coefficient.

3.5.2 Structural implications of monovalent and divalent ion binding

We used 2D [^{1}H,^{1}H] TOCSY NMR spectroscopy to evaluate possible similarities and differences of the influence of mono- and divalent metal ions on the structural and dynamic features of the DNAzyme:RNA complex. We compared the effects of addition of moderate amounts of Mg^{2+} (1 mM, Figure 3.4A) to the effects of addition of high amounts of Na^+ (400 mM, Figure 3.4B) to the stabilized DNAzyme:RNA complex (pre-equilibrated with 100 mM Na^+). Interestingly, the addition of 400 mM monovalent Na^+ induces nearly identical changes in the spectrum as the addition of 1 mM divalent Mg^{2+}, including peak shifts in the binding arms and, in particular, also the disappearing of signals from nucleotides located in the catalytic loop. Our findings strongly indicate that both key features, i. e. structural modifications as well as the exchange dynamics, induced by Mg^{2+} on the DNAzyme:RNA complex can be mimicked by Na^+. This also implies that Na^+ can compete with Mg^{2+} for the same binding sites. In combination with our data on DNAyzme activity, this highlights a number of unforeseen features of the 10-23 DNAzyme (*vide infra*).

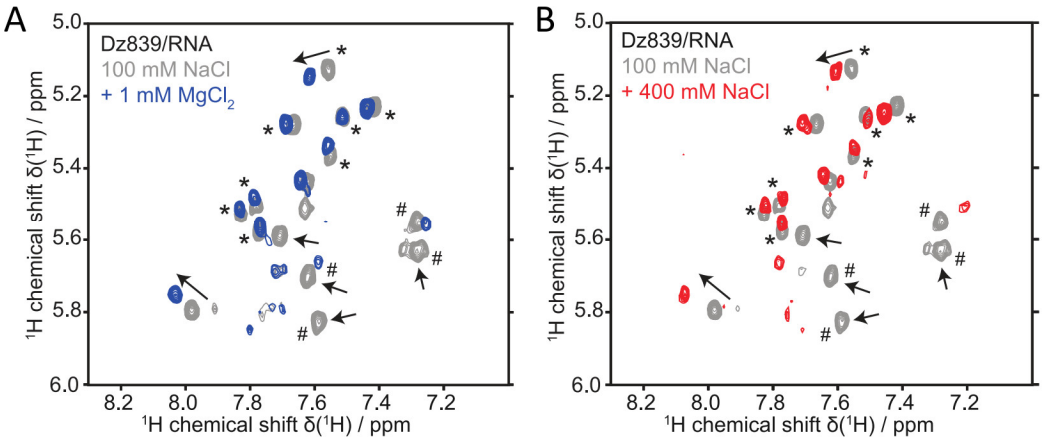


Figure 3.4: Structural changes of the DNAzyme:RNA complex induced by the presence of mono- and divalent cations. (A) Overview of CSPs and peak disappearance in 2D [^{1}H,^{1}H] TOCSY NMR spectra induced by addition of 1 mM Mg^{2+}. For simplicity the contour level was adjusted to the otherwise identical respective data shown in Figure 3.2A. (B) Increasing the Na^+ concentration from 100 mM (A and B, gray) to 500 mM (B, red) has a similar effect, both in respect to CSPs and peak disappearance, as the addition of 1 mM Mg^{2+} (A, blue).

3.5.3 Effects of monovalent ions on M^{2+}-induced 10-23 DNAzyme catalysis

To correlate the observations obtained by our calorimetric and spectroscopic measurements with the catalytic performance of the 10-23 DNAzyme, we determined the reaction rate in the presence of varying Na^+ (Figures 3.5 and 3.6) or K^+ concentrations (Figure S3.5) as well as at different Mg^{2+} or Mn^{2+} concentrations. We used a FRET setup with a non-stabilized RNA substrate (T839-FRET) labeled with a 6-Carboxyfluorescein (6-FAM) at the 5' end and a Black Hole Quencher-1 (BHQ1) at the 3' end to monitor cleavage in real-time. Cleavage and dissociation of the RNA substrate lead to an increase in the fluorescence signal due to the separation of fluorophore and quencher molecule that can easily be detected in a high-throughput real-time setup (Figure S3.1C). To be able to focus on the role of metal ions in the cleavage reaction, we used single turnover kinetics with a pre-formed DNAzyme:RNA complex. Figure 3.5A shows an example of such a FRET-based activity measurement in the presence of 3 mM Mg^{2+} and Na^+ concentrations ranging from 0 to 1000 mM. Fitting each individual curve to a simple exponential function allows for the determination of the observed rate constant k_{obs}.

In agreement with previous studies (Santoro & Joyce, 1997), our data shows that the k_{obs} values of the cleavage reaction increase with increasing Mg^{2+} concentrations at constant Na^+ levels. However, we observed a dramatic decrease in the k_{obs} values with increasing Na^+ levels for constant Mg^{2+} concentrations for all tested conditions (Figure 3.5B). Plotting the Na^+-dependence of the rate constants for each Mg^{2+} concentration to the respective maximal rate constant (individual normalization) shows that in each case the presence of Na^+ can reduce k_{obs} by approximately fivefold at very high Na^+ concentrations. At conditions representing a physiological relevant ionic strength, still an approximately 40% reduction of the rate constant as compared to the absence of monovalent ions is detected (Figure 3.5C; also see Table S3.1 for concentrations in mammalian cells of the different metal ions that are studied in this work). We

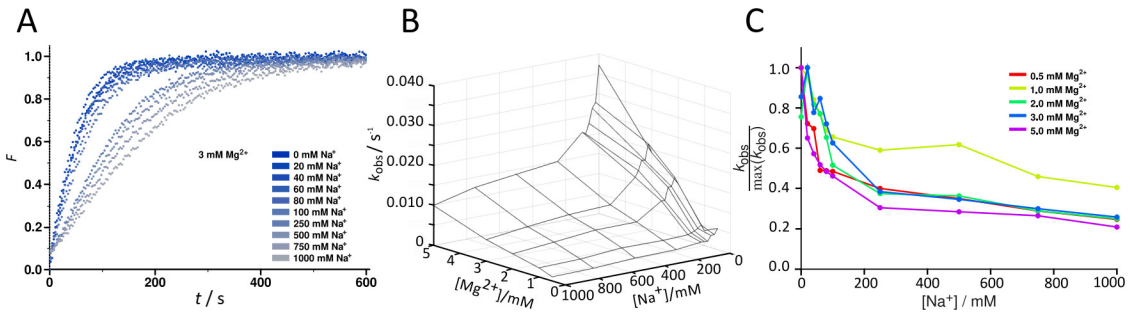


Figure 3.5: Influence of magnesium and sodium ions on DNAzyme catalysis determined by FRET-based activity assays. . (A) Normalized fluorescence intensities (F) at 3 mM Mg^{2+} concentration and different Na^+ concentrations. (B) Rate constants k_{obs} for the RNA cleavage as a function of Mg^{2+} and Na^+ concentrations obtained by curve fitting of the data as shown in A by $F=1\ \ e^{k_{obs} \cdot t}$. (C) Individual normalization of the rate constants indicates that the Na^+ effects on k_{obs} are similar for each Mg^{2+} concentration leading to an about fivefold reduction of the reaction rate at high Na^+ concentrations.

further carried out a comparable assay using Mn^{2+} instead of Mg^{2+}. Our data reveals that the cleavage reaction is significantly faster in the presence of Mn^{2+} than in the presence of Mg^{2+}, which is in accordance with previous studies (He *et al.*, 2002; Santoro & Joyce, 1998; Sugimoto *et al.*, 1999). Due to the high rate constants of the Mn^{2+}-induced reaction only a limited range of monovalent and divalent ion concentrations can be resolved with our experimental setup. Our data clearly show that within this range the Na^+ promote Mn^{2+} induced cleavage exposing a completely opposite effect on the Mn^{2+}-induced reaction than on the Mg^{2+}-induced reaction (Figure 3.6B, see Figure S3.6 for individually normalized k_{obs} values). To evaluate whether these effects can also be observed with K^+ or if they are specific for Na^+, we repeated the experiments in the presence of KCl. Our results show that K^+ influences DNAzyme catalysis similar to Na^+ (Figure S3.5).

In summary, our FRET-based activity measurements reveal that increasing levels of Na^+ or K^+ reduce the rates of Mg^{2+}-induced reactions of the 10-23 DNAzyme, while they increase the rates of Mn^{2+}-induced reactions. Thus, the functional data correlates well with the effects of monovalent ions on binding of divalent cations to the DNAzyme:RNA complex, suggesting different mechanisms of the reaction in presence of Mg^{2+} or Mn^{2+}.

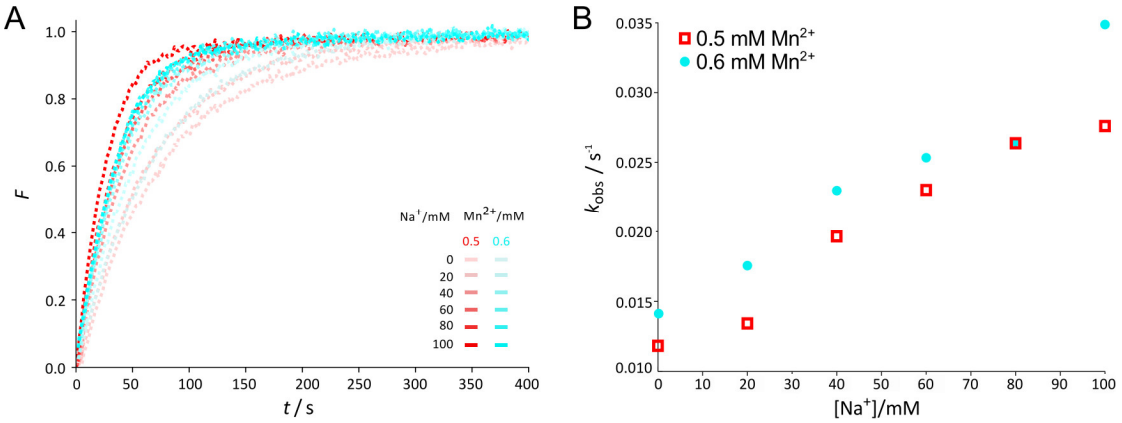


Figure 3.6: Effect of Na^+ on Mn^{2+}-induced DNAzyme reaction rates. (A) Normalized fluorescence intensities (F) resulting from the cleavage assays at 0.5 and 0.6 mM Mn^{2+} concentration and different Na^+ concentrations. After addition of the divalent metal ion cofactor at time zero the measured fluorescence intensity increases due to cleavage of the RNA substrate. (B) Observed rate constants (k_{obs}) for the DNAzyme catalyzed RNA cleavage in the presence of different concentrations of Mn^{2+}.

3.5.4 Model for the interaction of metal ions with the DNAzyme:RNA complex

In an attempt to link our binding and activation data, we propose two reaction schemes for the role of either Mg^{2+} or Mn^{2+} in 10-23 DNAzyme-mediated catalysisbased on our observation that monovalent cations have opposite effects on the cleavage reaction induced by Mg^{2+} or Mn^{2+}. In case of Mg^{2+}, the influence of Na^+ on the activity of the 10-23 DNAzyme resembles enzyme inhibition, whereas in the case of Mn^{2+} it indicates enzyme activation. Thus, simplistic but still computable models might be the following: For the Mg^{2+}-catalyzed reaction

$$\text{Dz:RNA} + \text{Na}^+ \underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}} \text{Dz:RNA:Na} \tag{3.4}$$

$$\text{Dz:RNA} + \text{Mg}^{2+} \underset{k_{-2}}{\overset{k_{+2}}{\rightleftharpoons}} \text{Dz:RNA:Mg} \xrightarrow{k_3} \text{Dz} + \text{RNA}_{\text{cleaved}} + \text{Mg}^{2+} \tag{3.5}$$

and for Mn^{2+}-catalyzed reaction

$$\text{Dz:RNA} + \text{Na}^+ \underset{k_{-4}}{\overset{k_{+4}}{\rightleftharpoons}} \text{Dz:RNA:Na} \tag{3.6}$$

$$\text{Dz:RNA:Na} + \text{Mn}^{2+} \underset{k_{-5}}{\overset{k_{+5}}{\rightleftharpoons}} \text{Dz:RNA:Na:Mn} \xrightarrow{k_6} \text{Dz} + \text{RNA}_{\text{cleaved}} + \text{Mn}^{2+} + \text{Na}^+ \tag{3.7}$$

with the abbreviations Dz for the 10-23 DNAzyme and $RNA_{cleaved}$ for the hydrolyzed RNA. For the Mg^{2+} system ((3.4) and (3.5)) Na^+ binding to the Dz:RNA complex competes with Mg^{2+} binding (i. e. $k_{+/-1}$ vs. $k_{+/-2}$). This discretion reflects our finding that increasing Na^+ levels create unfavorable conditions for the coordination of Mg^{2+} to the DNAzme:RNA complex, which leads to decreasing reaction rates (Figure 3.5) as well as our NMR observation that Mg^{2+} and Na^+

likely compete for binding sites on the DNAzyme:RNA complex (Figure 3.4). For DNAzyme catalysis in the presence of Mn^{2+} ((3.6) and (3.7)), we propose a model where increasing Na^{+} levels support the DNAzyme:RNA complex in adopting a conformation (i. e. Dz:RNA:Na in (3.6)) favorable for Mn^{2+} interactions that result in substrate cleavage. Taking this into account, reactions (3.4) and (3.6) have not to describe identical Na^{+} or Mg^{2+}/Mn^{2+} binding modes or positions, which is reflected by different rate constants in 3.4 and (3.6). Notably, binding of only a single M^{2+} is considered here. Binding of additional M^{2+}, as revealed by ITC (Figure 3.1) and NMR (Figure 3.2) measurements for Mg^{2+} and by EPR for Mn^{2+} (Figure 3.3), are assumed not to be involved in the cleavage process. We applied (3.4), (3.5) and (3.6), (3.7) in a global fit of FRET data recorded with Mg^{2+} and Mn^{2+}, respectively (Figure 3.7). While not all details are adequately captured, the approach is capable to reproduce the central features of the NaCl dependency of the 10-23 DNAzyme activity. To draw more definite conclusions, the range of optimal parameter combinations is, however, too large (Figure S3.8 and Table 3.2). Nevertheless, it provides an initial description of the interplay of the observed monovalent and divalent ions and its consequence for the 10-23 DNAzyme 's capability to cleave its substrate.

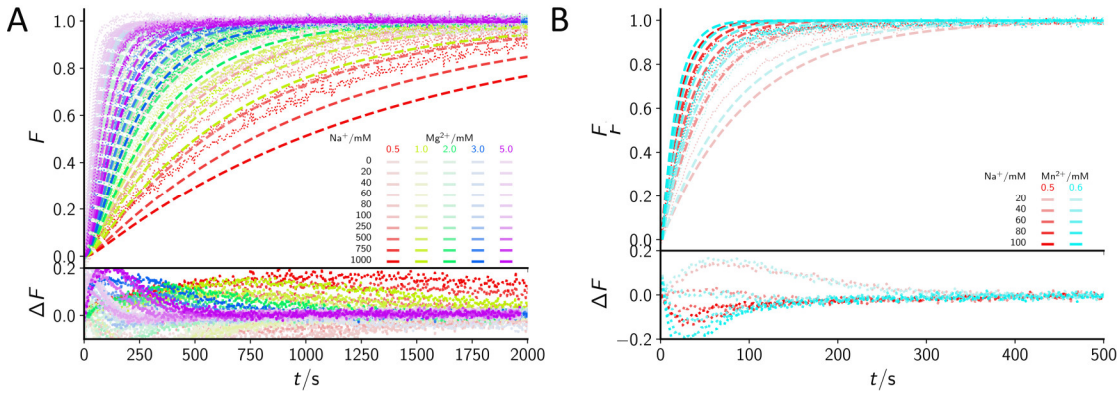


Figure 3.7: Global fit of the Na^{+}-dependency of the fluorescence-based cleavage assays at different Mg^{2+} (A) and Mn^{2+} (B) concentrations. (A) The observed rate of Mg^{2+}-induced cleavage drops with increasing Na^{+} concentration and was fitted using (3.4) and (3.5). (B) the rate of Mn^{2+}-induced cleavage increases with increasing Na^{+} concentration and was fitted using (3.6) and (3.7). Since we assume a model in which Mn^{2+} can only bind to a DNAzyme:RNA:Na^{+} complex, data obtained in the absence of Na^{+} were not taken into account for the calculation (see Methods for more details on data normalization, resulting fit parameters and fitting procedure according).

3.6 Discussion

Metal ions are essential for the function of most DNAzymes. Cations are required for secondary and tertiary structure formation, structural stability, and they may play a direct role in catalysis. Previous studies have predominantly focused on the role of divalent metal ions on DNAzyme activity (Okumoto & Sugimoto, 2000; Santoro & Joyce, 1998). The aim of this work is to pinpoint the relationship between mono- and divalent cations in 10-23 DNAzyme catalysis. We have shown that monovalent metal ions can indeed have a dramatic impact on DNAzyme activity. In this respect we focused on the monovalent ions' implications for (i) divalent cation binding to the DNAzyme:RNA complex, (ii) metal-ion induced structure and dynamic effects of the DNAzyme:RNA complex, (iii) divalent-ion induced DNAzyme activity, and (iv) a more generalized formal description of the 10-23 DNAzyme-mediated catalysis. Our study reveals that Na^+ as well as K^+ levels impact Mg^{2+}- or Mn^{2+}-induced catalysis in different ways, suggesting different modes of action for the divalent cations. This links a more complex interplay between mono- and divalent metal ions with DNAzyme function. A previous study using the same 10-23 DNAzyme variant suggested that three Mg^{2+} cations bind to the DNAzyme:RNA complex in the absence of Na^+, while a single Mg^{2+} cation binds to the complex in the presence of 100 mM Na^+ (Victor *et al.*, 2018). In this study a simple two-state model was assumed in which the DNAzyme:RNA complex is either completely free of Mg^{2+} or fully bound to all three Mg^{2+} (Victor *et al.*, 2018). Our calorimetric data shows that in in the absence as well as presence of NaCl a model assuming multiple Mg^{2+}-binding sites in an *n*:1 binding mode provides the best fit (Figure 3.1E and F) . The observation that more than one divalent metal ion is interacting with the DNAzyme:RNA complex in presence of all tested Na^+ conditions isf further supported by our EPR spectroscopic measurements. In the absence of Na^+ the best fit for our EPR data is achieved by a simple model assuming a non-cooperative binding of six Mn^{2+} to the DNAzyme:RNA complex. With increasing Na^+ concentration a cooperative binding behavior with a Hill coefficient n of > 1 is detected (Figures 3.3B and C, S3.4). A Hill coefficient $n > 1$ indicates positive cooperativity, i. e. once one Mn^{2+} is bound to the enzyme, its affinity for another Mn^{2+} increases. In comparison to the binding behavior in the absence of Na^+ (Figure 3.3A) our data indicate that the presence of Na^+ renders binding of the first Mn^{2+} ions less favorable. Possible explanations for the role of Na^+ in this behavior include competition for the same binding site, shielding of electrostatic interactions and/or stabilization of different complex conformations. Our NMR spectroscopic data could shed more light on the structural changes induced by mono- and divalent metal ions. For the interaction of Mg^{2+} with the binding arms of the DNAzyme, we observed fast exchange interactions, and the analysis of Mg^{2+}-induced CSP of selected peaks allowed for the accurate determination of a K_D value of 2.2 mM. For the interaction of Mg^{2+} with the catalytic loop, intermediate exchange interactions have been detected that lead to the disappearance of the affected peaks. Therefore, we were not able to determine an accurate K_D for this region. However, our data suggest that the affinity of nucleotides in the catalytic loop could be about tenfold higher as compared to residues in the binding arms (Figure 3.2C). Additionally, we compared the influence that addition of 1 mM Mg^{2+} or addition

of 400 mM Na^+ have on complex conformation and exchange dynamics. Surprisingly both conditions induced comparable peak shifts in the arms as well as disappearance of signals from nucleotides located in the loop (Figure 3.4). This observation indicates that Na^+ can compete with Mg^{2+} for the same binding sites in the catalytic loop and the substrate binding arms. In addition, it also suggests that Na^+ and Mg^{2+} have in many aspects comparable structural effects on the DNAzyme:RNA complex.

Albeit these similarities, we observed that the presence of high amounts of Na^+ is not sufficient to catalyze DNAzyme-mediated RNA cleavage in the absence of Mg^{2+} (Figure S3.7). This shows that the mechanism for the 10-23 DNAzyme-mediated reactions has to be different from the hammerhead ribozyme which is capable of accelerating the catalytic reaction in the presence of only high Na^+ levels (Perrotta & Been, 2006; Schnabl & Sigel, 2010). Based on our data we propose that high levels of Na^+ can induce a conformation that largely resembles the RNA cleavage competent state but remains inactive. In this picture the presence of Mg^{2+} or Mn^{2+} is than essential for the catalytic reaction itself.

Our calorimetric and spectroscopic studies were performed with a stabilized RNA substrate to prevent M^{2+}-induced cleavage and allow for a detailed analysis of the binding properties of metal ions towards the DNAzyme:RNA complex. In order to link our binding data with its implications for DNAzyme function we used FRET-labeled RNA substrates lacking the stabilizing fluoride substituent. We observed that, as expected, the k_{obs} of the cleavage reaction increases with increasing Mg^{2+} concentrations at constant Na^+ concentrations. Additionally, we observed a decrease in the k_{obs} values at constant Mg^{2+} concentrations with increasing Na^+ concentrations. For each tested Mg^{2+} concentration ranging between 0.5 and 5 mM, we observed a similar relative decrease in the k_{obs} value of approximately 40% at ion concentrations reflecting physiological relevant ionic strength as compared to the k_{obs} value in the absence of monovalent cations. Furthermore, the k_{obs} value drops further to about fivefold lower rates at elevated ionic strength (1000 mM NaCl) (Figure 3.5C). We performed comparable assays with Mn^{2+} instead of Mg^{2+}. In addition to the fact that Mn^{2+}-induced reaction is much faster than Mg^{2+}-induced cleavage, we observed a completely different effect of monovalent cations on the Mn^{2+}-induced reaction: the higher the concentration of Na^+ at constant Mn^{2+} concentration, the faster is the DNAzyme reaction (Figure 3.6B). Understanding the influence of Na^+ and K^+ on 10-23 DNAzyme catalysis is crucial for the development of biosensors or therapeutic gene-silencing applications. In addition to a limited availability of Mg^{2+} inside the cell, levels of Na^+ and K^+ are considerably high in the cytosol (Table S3.1). Therefore, our data suggest that the poor catalytic performance of RNA-cleaving DNAzymes in cells is not only due to low Mg^{2+} levels, but could also result from inhibition from monovalent cations, such as Na^+ and K^+. As for the applications of RNA-cleaving DNAzymes as biosensors, the accelerating or inhibiting effect of monovalent cations such as Na^+ or K^+ may lead to a distorted read-out.

While we can only speculate on the underlying mechanism, the observed effects are likely related to the metal ions' different intrinsic chemical and physical properties as listed in Table S3.2. Although Mg^{2+} and Mn^{2+} ions exhibit similar electrostatic properties, the two ions significantly differ in their affinity to their nucleic acid ligands (reviewed in Freisinger & Sigel

(2007)). On the one hand, Mg^{2+} and Mn^{2+} have many similar physicochemical properties, including their similar ion radius, leading to a similar charge-to-radius ratio that should provide similar electrostatic properties. On the other hand, Mg^{2+} is classified as a hard ion according to the Pearson model, while Mn^{2+} is classified as a borderline ion. Mg^{2+} has a [Ne] electron configuration with filled s- and p-orbitals and no contribution of d-orbitals to bonding. It preferentially forms ionic interactions with ligands such as oxygen that are also classified as hard. In contrast, Mn^{2+} has an electron configuration of [Ar] $3d^5$ with a half-filled d-orbital shell that contributes to bonding. It is classified as a borderline ion that strongly interacts with both nitrogen and oxygen ligands. Mn^{2+} binds stronger to the nitrogen atoms of nucleobases than Mg^{2+}, and its affinity for oxygen atoms of the phosphodiester is slightly higher than in case of Mg^{2+}. The affinity of Mg^{2+} to water is higher than for Mn^{2+}, indicating that Mg^{2+} is less capable of losing its hydration shell when binding to nucleic acids. Consequently, Mn^{2+} is more likely capable of directly binding to a nitrogen atom of a nucleobase by an inner-sphere contact than Mg^{2+}, which probably binds indirectly through hydrogen-bonded outer-sphere contacts (Hunsicker-Wang *et al.*, 2009). The different binding modes lead to an apparently higher affinity of Mn^{2+} for nucleic acid binding sites than for Mg^{2+} (Hunsicker & DeRose, 2000). Combining our findings with the intrinsic properties of the different metal ions it is tempting to speculate that the Na+ can rather adequately mimic the hydration shell contacts of divalent ions which are responsible to induce a cleavage-competent conformation of the pre-catalytic DNAzyme:RNA complex. This state, however, still needs at least one additional divalent cation for the reaction itself. In this speculative model the monovalent cations may be even better suited to satisfy the hydration shell binding sites of the DNAzyme:RNA complex as compared to Mn^{2+}. As a result, the presence of Na^+ could allow the Mn^{2+} to focus on their role in the cleavage reaction, increasing the catalytic rates. The favorable hydration shell forming properties of Mg^{2+} could then be responsible for the observed opposite influence of monovalent ions on Mg^{2+}-induced cleavage. Here an increased preference of Mg^{2+} to the hydration shell binding sites could occupy a fraction of the Mg^{2+} in a competition with Na^+ for the same binding sites in the DNAzmye:RNA complex. In this picture, the presence of increasing Na^+ concentrations could explain the observed decreased Mg^{2+}-induced cleavage rates. Our generalized formalism ((3.4), (3.5) and (3.6), (3.7)) captures the essential parts of our results but will need additional adjustments to further untangle the different roles of metal ions in DNAzyme catalysis. In conclusion, the complex interplay of mono- and divalent metal ions has considerable implications on DNAzmye DNAzyme function and should be considered in future investigations and applications of the 10-23 DNAzyme.

3.7 Supporting data

Supplementary Tables

Table S3.1: Concentrations of Mg^{2+}, Mn^{2+}, K^+ and Na^+ in mammalian cells, as reviewed by Pechlaner & Sigel (2012).

M^{n+}	Mammalian cell [mM]
Mg^{2+}	30[a]
Mn^{2+}	–
Na^+	10
K^+	140

[a]The free concentration of Mg^{2+} is about 1 mM.

Table S3.2: Physical and chemical properties of Mg^{2+}, Mn^{2+}, K^+, and Na^+. Ionic radius, preferred coordination numbers and ligands, pK_a values of the coordinated water molecules, hardness on the Pearson scale, water exchange rate from the first hydration shell at 298 K (k_{ex}) and the free enthalpy for hydration (ΔG_{hyd}) as reviewed by Pechlaner & Sigel (2012). Values in parentheses belong to the coordination number also given in parentheses.

M^{n+}	Ionic radius [Å]	Coordination numbers	Preferred ligands	pK_a	Pearson hardness	k_{ex} [s^{-1}]	ΔG_{hyd} [kJ/mol]
Mg^{2+}	0.72	6	O	11.44	32.6	$5.3 \cdot 10^5$ to $6.6 \cdot 10^5$	-1858
Mn^{2+}	0.83	6	O/N	10.59	9.0	$2.1 \cdot 10^7$	-1821
K^+	1.38 (1.51)	6 (8)		14.46	13.6	10^9	-322 (-308)
Na^+	1.02 (1.18)	6 (8)		14.18	21.1	10^9	-412 (-383)

Supplementary Figures

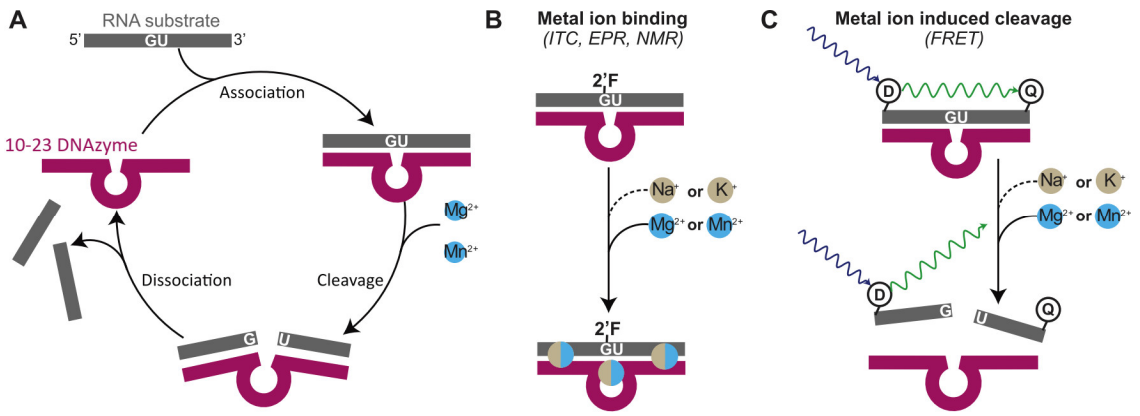


Figure S3.1: Experimental design. (A) Reaction scheme of the 10-23 DNAzyme. The RNA substrate (indicated in grey) hybridizes with the DNAzyme (indicated in purple) via Watson-Crick base pairing. Cleavage of the RNA is induced by Mg^{2+}. The ternary complex dissociates and the DNAzyme is released to bind another target molecule. (B) Schematic overview of the experimental setup for metal ion binding experiments using a stabilized RNA substrate. To prevent the RNA substrate from DNAzyme-mediated cleavage, the 2'-OH group at the nucleotide at the cleavage site is substituted by a fluorine atom. (C) Schematic overview of the experimental setup to study metal ion induced RNA cleavage using FRET-labeled RNA substrates. When the RNA substrate is intact, excitation of the 6-FAM fluorophore D leads to a (partial) transfer of energy to the BHQ-1 Q moiety based on the principle of Förster resonance energy transfer where the energy dissipates as heat. DNAzyme-mediated cleavage of the RNA substrate separates fluorophore from quencher and leads to increased fluorescence after excitation of the 6-FAM molecule.

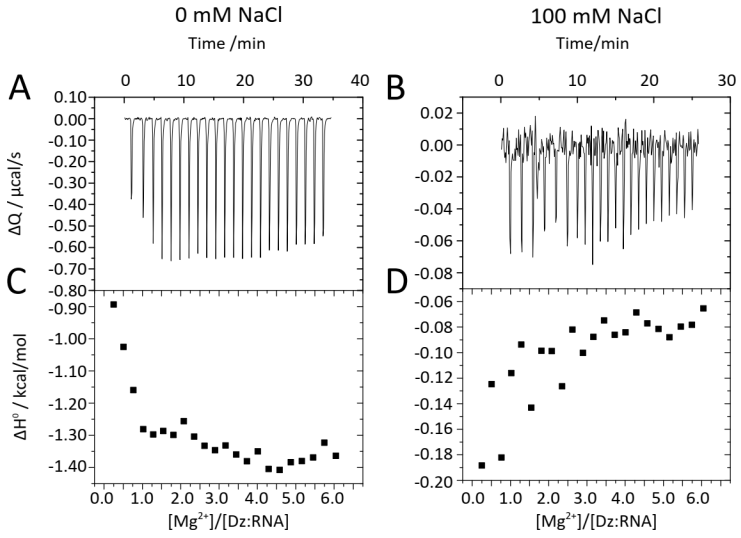


Figure S3.2: Titration of Mn^{2+} into pre-formed DNAzyme:RNA complex in the absence (A, C) or presence (B, D) of 100 mM Na^{+}, monitored at 30 °C. The upper panels show the baseline-corrected instrumental response. The lower panels show the integrated data (filled squares).

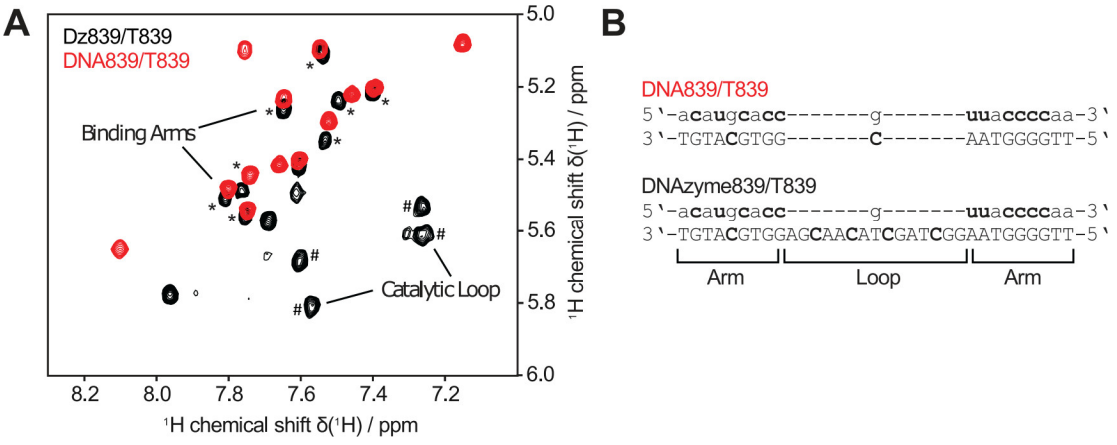


Figure S3.3: Assignment of catalytic loop and binding arms. Overlayed [^{1}H, ^{1}H]-TOCSY spectra of DNAzyme:RNA complex (black) and complementary RNA:DNA helix without catalytic loop (red). Signals correspond to the pyrimidines U and C. Signals with low chemical shift difference were assigned to the binding arms (*), while newly appearing signals were assigned to the catalytic loop (#). (B) Sequences of DNAzyme:RNA and DNA:RNA complexes. Pyrimidines, which appear in the spectral region of (A), are written in bold, binding arms and catalytic loop are highlighted.

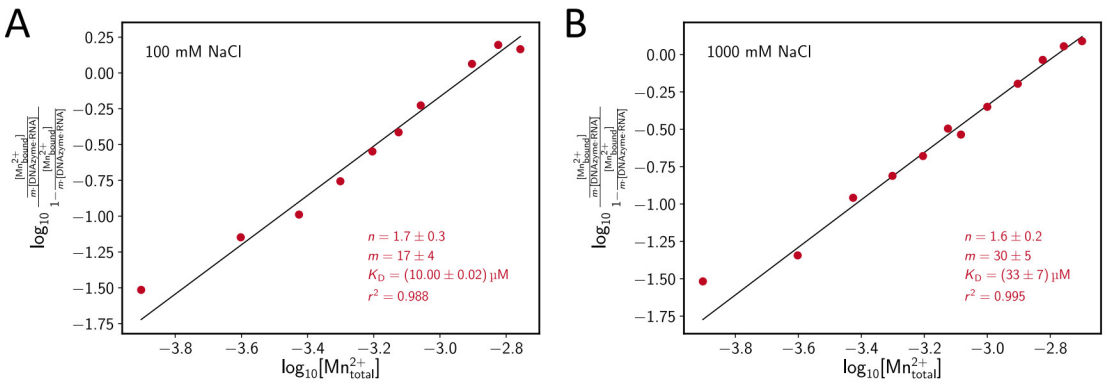


Figure S3.4: Hill plots with the experimental data as in Figure 3.3 obtained by EPR experiments.

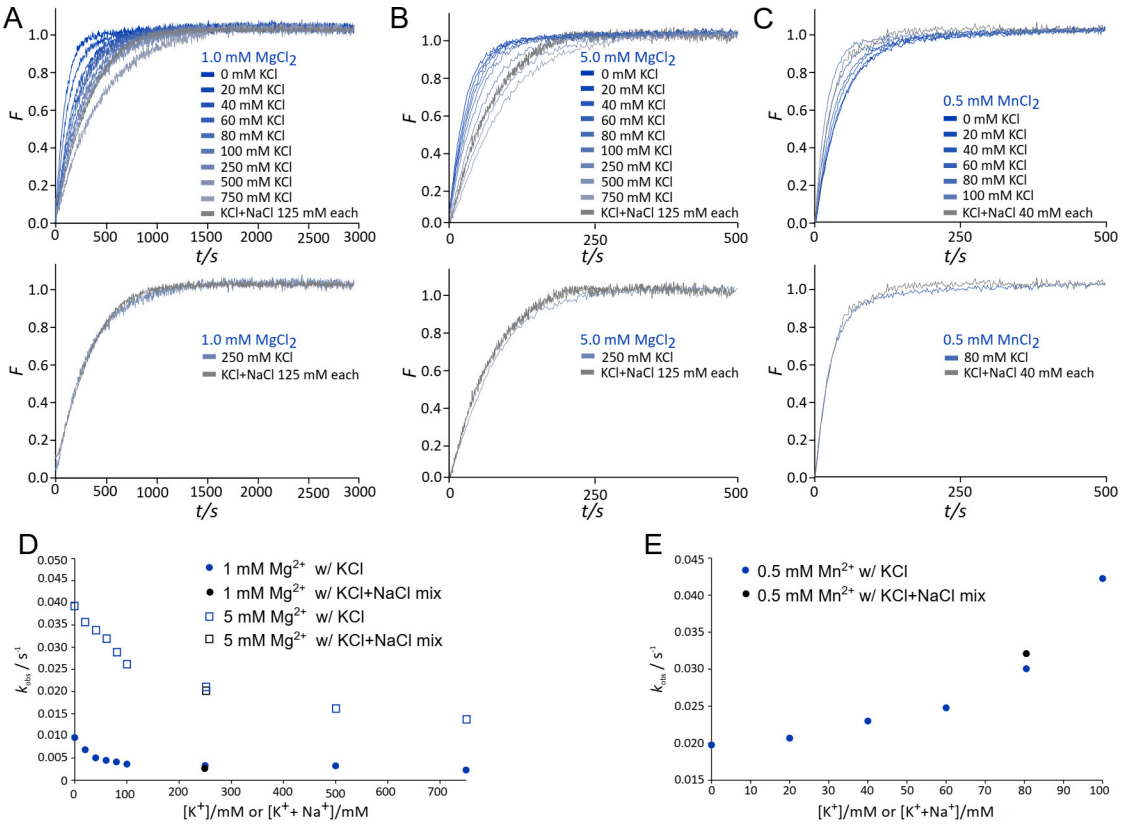


Figure S3.5: Influence of K^+ levels on DNAzyme catalysis. Normalized fluorescence intensities resulting from the FRET-based activity assays at (A) 1 mM Mg^{2+}, (B) 5 mM Mg^{2+}, and (C) 0.5 mM Mn^{2+} and different K^+ concentrations (top) or a mixture of K^+ and Na^+ (125 mM each) and 250 mM K^+ (bottom). (D) Observed rate constants (k_{obs}) for the DNAzyme catalyzed RNA cleavage in the presence of different concentrations of Mg^{2+}. (E) Observed rate constants (k_{obs}) for the DNAzyme catalyzed RNA cleavage in the presence of different concentrations of Mn^{2+}.

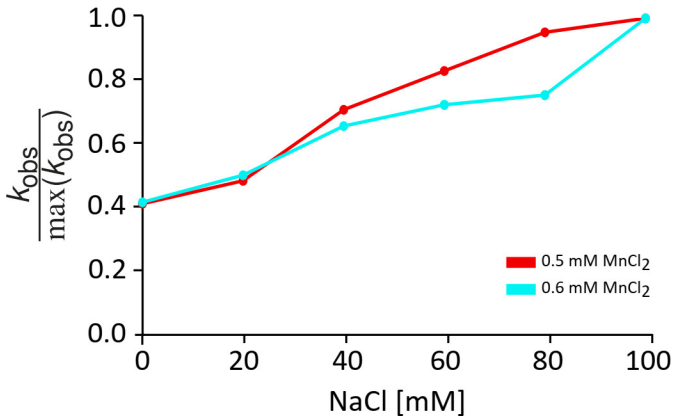


Figure S3.6: Individual normalization of the rate constants of the 10-23 DNAzyme-mediated RNA cleavage reaction for each Mn^{2+} concentration.

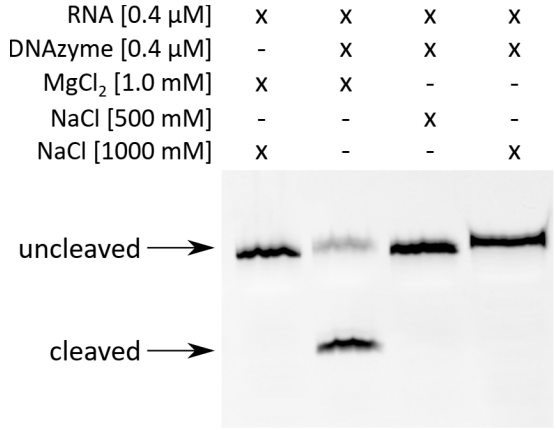


Figure S3.7: Cleavage assay of the 10-23 DNAzyme with the fluorescein-labeled RNA substrate (T839-FAM) in the presence and absence of 1 mM $MgCl_2$ and 500 or 1000 mM NaCl. PAGE analysis of the reaction mixtures shows that the DNAzyme is not capable of cleaving its RNA substrate in the sole presence of Na^+ ions.

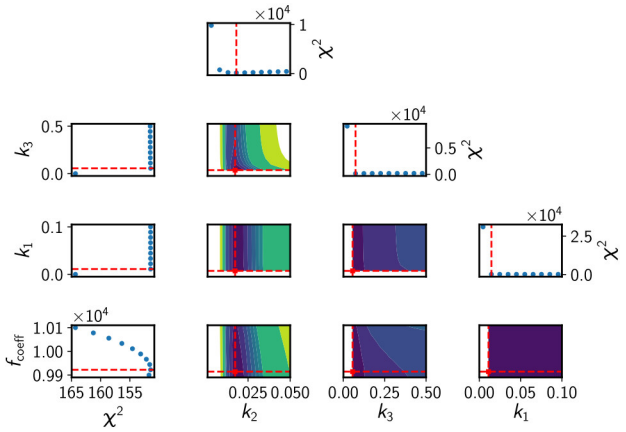


Figure S3.8: Example output of scan for parameter solutions with “brute" force grid-search method. The graphs display the chi-square [χ^2] value per parameter and contour plots for all combinations of two parameters with colors from blue to yellow for χ^2 values normalized from 0 to 1. The dotted red lines depict the optimal parameter value.

4 Expanding crystallization tools for nucleic acid complexes using U1A protein variants

This chapter reflects content of the following publication.

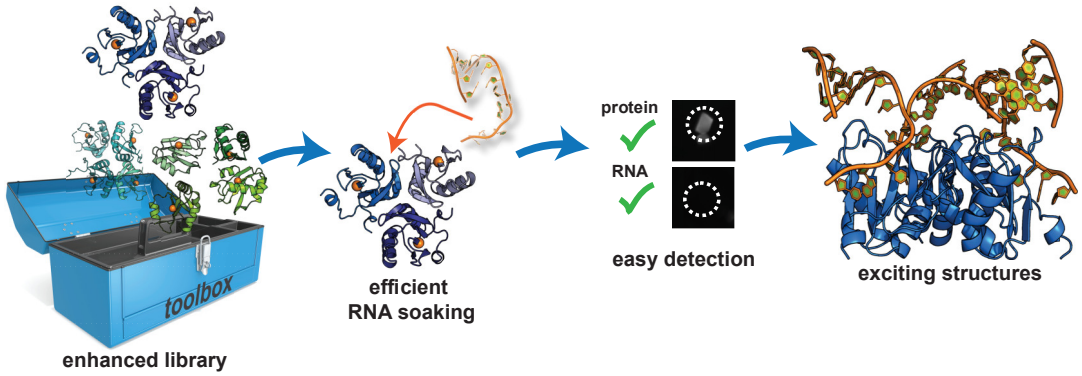
4.1 Publication information

Hannah Rosenbach, Julian Victor, Jan Borggräfe, Ralf Biehl, Gerhard Steger, Manuel Etzkorn, and Ingrid Span

Journal of Structural Biology, 2020, Vol. 210(2), 107480

4.2 Abstract

The major bottlenecks in structure elucidation of nucleic acids are crystallization and phasing. Co-crystallization with proteins is a straight forward approach to overcome these challenges. The human RNA-binding protein U1A has previously been established as crystallization module, however, the absence of UV-active residues and the predetermined architecture in the asymmetric unit constitute clear limitations of the U1A system. Here, we report three crystal structures of tryptophan-containing U1A variants, which expand the crystallization toolbox for nucleic acids. Analysis of the structures complemented by SAXS, NMR- and optical spectroscopy allow for insights into the potential of the U1A variants to serve as crystallization modules for nucleic acids. In addition, we report a fast and efficient protocol for crystallization of RNA by soaking and present a fluorescence-based approach for detecting RNA-binding *in crystallo*. Our results provide a new tool set for the crystallization of RNA and RNA:DNA complexes.



Graphical Abstract: New tool set for the crystallization of RNA and RNA:DNA complexes.

4.3 Introduction

Nucleic acids are capable of adopting 3D structures that are considerably more complex than double stranded helices. The structure of ribozymes or DNAzymes enables them to accelerate chemical transformations, which expands the biological role of nucleic acids from carrier of genetic information to effective biocatalysts. Solving the structure of these nucleic acids in a catalytically relevant conformation is the key to understanding the reaction mechanism. X-ray crystallography is a widely used method for structure solution, however, the formation of highly ordered nucleic acid crystals is often prohibited by their surface properties, which are dominated by negatively charged and regularly ordered phosphate groups (Ferré-D'Amaré & Doudna, 2000b). Thus, numerous attempts to crystallize nucleic acids often resulted in the formation of crystals with poor long-range order. One of the approaches to tackle this problem involves the co-crystallization of RNA bound to the RNA-binding domain of U1A (Ferré-D'Amaré & Doudna, 2000a; Ferré-D'Amaré *et al.*, 1998). Proteins consist of a larger number of building blocks than nucleic acids, which increases the number of possible interactions between the molecules in the asymmetric unit and the probability of obtaining highly ordered crystals. Thus, using an nucleic acid-binding protein as crystallization helper has proven to be a successful strategy to solve the structure of ribozymes or DNAzyme:substrate complexes.

The RNA-binding fragment of the U1A protein consists of 98 amino acids and binds a specific RNA hairpin loop with high affinity (Oubridge *et al.*, 1994). U1A-RBD adopts a globular shape with a compact fold in presence or absence of RNA (Nagai *et al.*, 1990; Oubridge *et al.*, 1994). This feature makes U1A-RBD particularly suitable for the formation of highly ordered crystals (Ferré-D'Amaré, 2010). The crystal structures of U1A-RBD:RNA complexes reveal that contacts between RNA and protein comprise polar and non-polar interactions, the latter of which are independent of ionic strength. Therefore, complex formation is not strongly affected by extreme crystallization conditions, such as solutions with very low or high salt concentrations (Ferré-D'Amaré, 2010). Several RNA structures have been solved using U1A-RBD as crystallization module, among them the self-cleaving hepatitis delta virus ribozyme (Ferré-D'Amaré *et al.*, 1998), the hairpin ribozyme (Rupert & Ferré-D'Amaré, 2001) the GlmS riboswitch (Cochrane *et al.*, 2007), a self-splicing group I intron (Adams *et al.*, 2004), and the catalytic core of an RNA-polymerase ribozyme (Shechner *et al.*, 2009). Although U1A-RBD has successfully been used as a crystallization module (Ferré-D'Amaré, 2010), one major drawback is the lack of tryptophan residues in the endogenous amino acid sequence. Tryptophan residues allow for the determination of the protein concentration by measuring the absorbance at 280 nm and enable the distinction between protein and salt crystals by fluorescence imaging (Meyer *et al.*, 2015).

Thus, we designed five U1A-RBD variants containing one tryptophan in their protein sequence, evaluated their crystallization properties, solved the structures of three variants and accessed their suitability for soaking experiments. We also analyzed the obtained protein crystals using fluorescence imaging and monitored the changes in fluorescence after soaking the RNA into preformed crystals. In addition, we investigated the application of an U1A variant

for structural characterization of an RNA-cleaving DNAzyme in complex with its corresponding substrate using native PAGE, solution NMR spectroscopy, and SAXS measurements. Further, we present an improved protocol for obtaining U1A:RNA complex crystals by soaking. This method provides a time and cost effective alternative to generate complex crystals with respect to co-crystallization and even more importantly prevents degradation of fragile nucleic acid molecules rendering it in particular valuable for RNA structure determination.

4.4 Material and Methods

4.4.1 Cloning and site-directed mutagenesis

The *U1A* gene, encoding for the amino acids 1 to 98 of the human U1A-RBD (UniProtKB accession number P09012) was obtained as DNA string by GeneArt (Thermo Fisher Scientific, Waltham, MA, USA). An additional sequence to generate an overlap with the vector was introduced by amplification of the DNA string using the primers 5'-AGG AGG TCT AGA ATG GCA GTT CCC GAG ACC-3' and 5'-CCT GAA AAT ACA AAT TTT CAT CTT GGC AAT GAT ATC TGA GTC G-3'. The plasmid pET16b-TEV was linearized using the primers 5'-TAA CTA GCA TAA CCC CTT GGG GC-3' and 5'-CAT ATG TCC CTG AAA ATA CAG GTT TTC ATG GC-3'. The plasmid pET16b-TEV-U1A was engineered using the In-Fusion HD cloning kit (Takara Bio, Kisatsu, Japan) and verified by sequencing (GATC Biotech, Konstanz, Germany). To introduce point mutations within the U1A gene we used primers with the desired codon exchange (Table S4.1). All polymerase chain reactions were performed with the Phusion High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, USA). The construct design will lead to proteins encoding for amino acid residues 1 to 98 of the U1A gene with two additional residues, glycine and histidine, before the sequence as a result of using the TEV recognition site.

4.4.2 Protein preparation

Cultures of *E. coli* BL21(DE3) containing the respective plasmid were cultivated in Terrific Broth medium supplemented with Ampicillin (100 µg/ml). Cells were cultivated at 37 °C and 160 rpm until the optical density measured at 600 nm reached a value of 2 followed by induction of gene expression by adding 0.5 mM IPTG to the cell culture. Cells were grown over night at 25 °C. Cells were harvested by centrifugation at 6,000 x g for 10 min, and 4 °C. For lysis cells were resuspended in lysis buffer [50 mM 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES) pH 7.5, 100 mM KCl] using a volume of 5 ml/g cell pellet. 2,2',2",2"'-(Ethane-1,2-diyldinitrilo)tetraacetic acid (EDTA)-free cOmplete Protease Inhibitor Cocktail Tablets (Roche, Basel, Switzerland) were added according to the manufacturer's instructions. The suspension was sonicated (Bandelin electronic, Berlin, Germany) for 20 min with an amplitude of 60% and a pulse of 1 s every 3 s using a VS70/T sonotrode. The cell debris was separated from the lysate by centrifugation for 45 min at 40,000 x g. The supernatant was fractionated with 0.5% polyethyleneimine, pH 7.5 followed by adding 75% (v/v) of a saturated ammonium sulfate solu-

tion as described previously (Ferré-D'Amaré, 2016). The protein in the pellet was then resuspended in wash buffer (50 mM HEPES pH 7.5, 100 mM KCl) and isolated using a HisTrap excel column (GE Healthcare, Chicago, IL, USA) with a bed volume of 5 ml equilibrated with wash buffer using an ÄKTAprime plus system (GE Healthcare, Chicago, IL, USA). The column was then washed with 9 column volumes (CV) 25 mM imidazole in wash buffer and subsequently eluted with 250 mM imidazole in wash buffer. Fractions containing the protein of interest were pooled and dialyzed against 2 l of wash buffer for 2 h at 4 °C. To remove the polyhistidine tag the protein solution was treated with the TEV protease in a molar ratio of 1:100 in the presence of 0.5 mM EDTA and 1 mM dithiothreitol (DTT) for 16 h. The TEV protease was removed using a HisTrap excel column, which was equilibrated with wash buffer containing 37.5 mM imidazole. The protein was then applied to the column and the flow-through was collected. The protein was concentrated by ultrafiltration using an Amicon Ultra-15 centrifugal filter (Merck Millipore, Burlington, MA, USA) with a 3 kDa molecular weight cut-off. For the next steps the chromatography systems and columns were washed with 0.5 M NaOH before use and then flushed with diethyl pyrocarbonate-treated H_2O. All solutions used in the following procedure were prepared with RNAse-free water and buffers containing RNase-free grade chemicals. U1A was further purified by size-exclusion chromatography (SEC) using a Superdex-75 PG column (GE Healthcare, Chicago, IL, USA) in wash buffer. For the final purification step a CHT-1 hydroxyapatite column (BioRad, Hercules, CA, USA) was used as described previously (Ferré-D'Amaré, 2016) and the protein solution was stored at -80 °C.

4.4.3 Design and synthesis of oligonucleotides

Starting points for the design of the DNA:RNA complex were the variant A5C of the DNAzyme Dz839, its target RNA T839 (Victor *et al.*, 2018) and the U1A-binding hairpin (Oubridge *et al.*, 1994). We used the software RNAcofold (Bernhart *et al.*, 2006b) and RNAfold from the Vienna package (Lorenz *et al.*, 2011) to design the DNA (Dz839_A5C) and RNA (T839_hairpin) constructs with RNA default parameters at 37 °C (Table S4.2). The design was performed according to these criteria: (i) low tendency for intramolecular folding; (ii) low tendency for intermolecular interactions; (iii) unequivocal formation of the target complex with minimum free energy; (iv) the hairpin loop should be easily accessible for the U1A protein; (v) the hairpin stem should be stable; and (vi) the backbone discontinuity at the adjacent 3' RNA and 5' DNA ends should be stacked to stiffen the overall conformation. All the DNA and RNA oligonucleotides (Table S4.2)were acquired from biomers.net GmbH (Ulm, Germany).

4.4.4 RNA and DNA binding assays using native PAGE

To analyze the capability of binding nucleic acids the U1A variants in wash buffer were mixed with 1 mM $MgCl_2$, RNA, and/or DNA (in H_2O) at equimolar ratios and incubated at room temperature for 20 min. The samples were mixed with DNA Loading Dye (6x) (Thermo Fisher Scientific, Schwerte, Germany) and 10 µM of the complex were then analyzed on a 10% na-

tive Polyacrylamide gel for 3 h at 180 V. For visualization the gel was incubated in Tris-borate EDTA (TBE) buffer containing a 1:10,000 dilution of the fluorescent nucleic acid dye GelRed (Merck, Darmstadt, Germany) for 1 h. Images were acquired using the ChemiDoc MP System (Bio-Rad, Hercules, CA, USA).

4.4.5 Analytical size-exclusion chromatography

To analyze the dispersity of the DNA:RNA:protein complex analytical SEC was performed using a Superdex 75 PC 3.2/30 column (GE Healthcare, Chicago, IL, USA). The column was run and equilibrated with 50 mM HEPES pH 7.5, 100 mM KCl and a sample volume of 10 µl at a concentration of 50 µM was injected per run.

4.4.6 Denaturing PAGE

To prove the presence of RNA in preformed protein crystals after the soaking experiments, the crystals were dissolved in 2x RNA loading buffer. The samples were analyzed via denaturing PAGE. Separation of the samples was carried out on 18% polyacrylamide gels with 7 M urea buffered with TBE for 1 h at 20 W. Images were acquired using a ChemiDoc MP System (BioRad, Hercules, CA, USA).

4.4.7 Small angle X-ray scattering

The scattering patterns were recorded with the laboratory based SAXS system “Ganesha-Air” from (SAXSLAB/XENOCS) at Forschungszentrum Jülich. The X-ray source is a D2-MetalJet (Excillum) with a liquid metal anode operating at 70 kV and 3.57 mA with Ga-Kα radiation (wavelength λ = 0.1314 nm) providing a very brilliant and a very small beam (<200 µm). A PILATUS 300K (Dectris) was used to record the 2D scattering patterns. All samples were sealed in glass capillaries of 2 mm inner diameter. Data were circular averaged, normalized to incident intensity and transmission corrected. The covered q range was 0.15-5.8 nm^{-1} measured at a detector distance of 54 cm and 114 cm. To extract the form factor concentrations of 5, 10 and 15 mg/ml were measured and the corresponding buffer was subtracted as background. The intensity was extrapolated to zero concentration to avoid any influence of a structure factor. To reconstruct a low-resolution model the form factor was analyzed by a sequence of programs from the ATSAS (Franke *et al.*, 2017) suite. After transformation of the form factor to a distance distribution, the program DAMMIF (Franke & Svergun, 2009) was used to generate 30 low resolution bead models by an ab initio bead modelling algorithm. A simulated annealing algorithm is used to construct a compact bead model that fits the experimental data. As the reconstruction did not result in unique shape models DAMAVER (Volkov & Svergun, 2003) is used to align a given set of models and to calculate a most probable model. The crystal structure of the U1A:RNA complex (PDB ID 6SQT) and a generic model of a DNAzyme:RNA complex were manually aligned to the experimental SAXS model to illustrate that size and shape of the SAXS data are overall consistent with the expected ternary complex.

4.4.8 NMR spectroscopy

NMR acquisition was performed on Bruker Avance III HD+ spectrometer operating at 600 MHz and equipped with a quadruple resonance QXI/QCI (^{1}H, ^{13}C, ^{15}N, ^{31}P) cryoprobe. All spectra were recorded as [^{1}H, ^{1}H]-2D-TOCSY using 2048x512 points and 10 ppm spectral width for both dimensions. 200 µM of U1A, DNA:RNA and DNA:RNA:U1A were measured at 37°C in 50 mM Tris-d3/HCl pH 7.5, 100 mM NaCl, 1 mM $MgCl_2$ and 10% (v/v) D2O. NMR spectra were processed with TOPSPIN 4.0.6 (Bruker) and plotted with Sparky 3.114 (T. D. Goddard, D. G. Kneller, UCSF).

4.4.9 Cleavage assay

Cleavage assays were performed with 10 µM RNA substrate and 10 µM DNAzyme in 50 mM Tris-HCl pH 7.5 in the presence of 100 mM Mg^{2+} for 3 h at 37°C. Prior to cleavage the RNA and DNAzymes were denatured in buffer in the absence of Mg^{2+} at 73°C for 5 min, followed by an incubation for 10 min at room temperature. Subsequently, the reaction was started by adding Mg^{2+}. For cleavage assays in the presence of the U1A protein, 10 µM of the U1A variant F56W were incubated for 30 min at room temperature to allow for formation of the ternary complex before starting the reaction with Mg^{2+}. The reaction was terminated by adding 95% formamide and 25 mM EDTA. Samples were heated to 96°C for 10 min and then cooled down on ice. Separation of the samples was carried out on 18% polyacrylamide gels with 7 M urea buffered with Tris-borate EDTA buffer (TBE) for 1.5 h at 20 W. For visualization the gel was incubated in TBE buffer containing a 1:10,000 dilution of the fluorescent nucleic acid dye GelRed (Merck, Darmstadt, Germany) for 1 h. Images were acquired using the ChemiDoc MP System (Bio-Rad, Hercules, CA, USA).

4.4.10 Crystallization and soaking experiments

Crystals were grown using the sitting-drop vapor diffusion method at 12°C for soaking experiments and at room temperature for co-crystallization experiments. The reservoir solutions consisted of 2.2 M ammonium sulfate and 0.2 mM tri-potassium citrate for F56W, 1.8 M ammonium sulfate and 0.1 M sodium acetate pH 4.6 for R70W, as well as 1.0 M ammonium sulfate and 0.1 M Tris/HCl pH 8.5 for K98W. For co-crystallization a sample volume of 1 µl with a concentration of 0.5 mM RNA:protein complex in 12.5 mM Tris pH 7.0, 50 mM KCl, 2 mM spermine, and 5 mM $MgCl_2$ was mixed with an equal volume of a reservoir solution containing 1.2 M ammonium sulfate and 0.1 mM tri-potassium citrate. Crystals were obtained within one month. After soaking for 1 min in cryoprotecting solution (50% glycerol) the crystals were mounted on loops, immediately flash-frozen, and stored in liquid nitrogen. For soaking experiments a volume of 0.7 µl hairpin RNA with a concentration of 1.5 mM in wash buffer was mixed with an equal volume of reservoir solution. Then preformed protein crystals were transferred into the solution and incubated for 4 d before they were flash-frozen and stored in liquid nitrogen. Tryptophan fluorescence of protein crystals was detected using the camera of a Rock Imager

RI1000 (Formulatrix, Bedford, MA, USA). For the soaking experiments a fluorescent crystal was transferred into a drop with a volume of 2 µl containing 1.5 mM of hairpin RNA and mixed with the reservoir solution at a 1:1 ratio. The crystal was incubated for 1 h and then transferred back into its mother liquor prior to imaging.

4.4.11 Data collection and structure determination

The X-ray diffraction data was collected at beamline PETRAIII of the Deutsches Elektronensynchrotron (German Electron Synchrotron, DESY) in Hamburg (Burkhardt *et al.*, 2016; Meents *et al.*, 2013), Germany at 100 K at an energy of 12 keV. Data processing was carried out using the xia2/xds program (Winter *et al.*, 2013). All structures were solved by molecular replacement using the Phaser program (McCoy, 2006) embedded in the CCP4i2 suite. The structure with the PDB IDs 1OIA (Nagai *et al.*, 1990) and 1URN (Oubridge *et al.*, 1994) were used as search models. The electron density maps revealed the protein or the protein and RNA, which were manually built using COOT (Emsley & Cowtan, 2004). The model was then refined against the diffraction data using REFMAC5 program (Murshudov *et al.*, 2011). During refinement, 5% randomly selected data were set aside for free R-factor calculations (cross validation). The 2*Fo-Fc* and *Fo-Fc* electron density maps were regularly calculated and used as a guide for model building and adjustment. The Ramachandran Plot analysis of the crystal structures reflect the following values for conformational space available: F56W (6SQT) 98.47% favored and 0.00% outlier; R70W (6SQV) 85.81% favored and 0.60% outlier; K98W (6SR7) 99.12% favored and 0.00% outlier; F56W:RNA co-crystallized (6SQN) 98.22% favored and 0.00% outlier and F56W:RNA soaked (6SQQ) 97.79% favored and 0.00% outlier. The structural refinement statistics are summarized in Tables S4.3 and S4.4.The atomic coordinates for the U1A variants have been deposited at the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) (Berman, 2000). Atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data bank under accession numbers 6SQT for the variant F56W, 6SQV for the variant R70W, 6SR7 for the variant K98W, 6SQQ for the F56W:RNA complex obtained by soaking, and 6SQN for the co-crystallized complex.

4.5 Results

4.5.1 Design, biosynthesis, and characterization of U1A triple mutants

To identify amino acid positions that likely do not affect structure and RNA-binding properties of U1A we used BLAST (Altschul *et al.*, 1990) to search for U1A-RBD homologs that contain tryptophan residues. Whereas most U1A-RBDs from *Animalia* lack tryptophan residues, most homologs from *Viridiplantae* contain a tryptophan instead of a phenylalanine at position 56 (Figure S4.1). This position is critical for RNA binding as the phenylalanine stacks to a certain nucleotide in the complex formed by U1A and RNA (Shiels *et al.*, 2002). Notably, the U1 RNA hairpin sequence is conserved in *Viridiplantae* and *Animalia*. This indicates that the mutation

F56W should not prohibit RNA-binding to U1A-RBD (Law, 2005; Shiels *et al.*, 2002). Previously, it has been shown that the exchange of amino acid residues in the protein U1A affects crystallization behavior and the properties of the protein in solution (Price *et al.*, 1995). Since an U1A variant containing the two mutations Y31H and Q36R has been reported to have superior crystallization properties compared to the wildtype sequence (Oubridge *et al.*, 1994) we used this double mutant as parent sequence for the introduction of all tryptophan mutations. In addition to position 56, we selected positions A2, H10, R70, and K98 in the U1A-RBD for site-directed mutagenesis, since they are located distant to the RNA-binding pocket and on the protein surface (Figure 4.1). Starting from the *U1A* gene hosted within the expression plasmid pET16b-TEV encoding the parent amino acid sequence, we exchanged the codon for amino acids A2, H10, F56, R70, and K98 by W using site-directed mutagenesis. The genes were then expressed in *Escherichia coli* BL21(DE3) cells resulting in high protein yields. All proteins were successfully isolated and purified to high purity using a combination of several precipitation and chromatography steps (Figure S4.2).

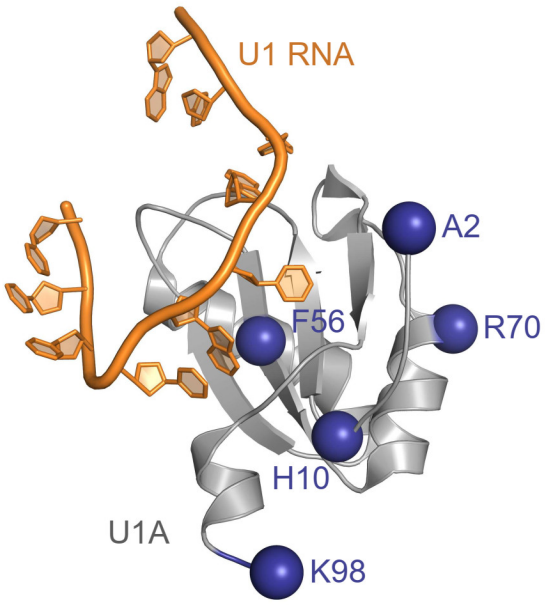


Figure 4.1: Crystal structure of U1A protein (gray) in complex with the U1 RNA hairpin loop (orange) (Oubridge *et al.*, 1994) with the positions that were selected for amino acid exchange highlighted as spheres (blue).

Next, we tested the capability of the rationally designed U1A variants to bind RNA. Therefore, we incubated each protein with the U1 RNA hairpin sequence at equimolar concentrations and, subsequently, analyzed the samples via native PAGE (Figure S4.3). The shift in electrophoretic mobility showed that all U1A variants retain their ability to bind RNA in solution.

4.5.2 Expanding the scope of U1A to DNA:RNA complexes

It has previously been shown that U1A can be used as a crystallization module for RNA sequences, including RNA hairpins and ribozymes (Ferré-D'Amaré *et al.*, 1998). In this study, we have evaluated the potential of the new U1A variants to serve as tools for structural studies of DNA:RNA complexes. Therefore, we used the 10-23 DNAzyme, a prominent member of the DNAzyme family, which is capable of binding and, subsequently, cleaving RNA substrates (Santoro & Joyce, 1997). DNAzymes are single-stranded DNA sequences that exhibit catalytic activity, such as the specific cleavage of RNA sequences. Previously reported attempts to solve the crystal structure of this 10-23 DNAzyme in its catalytically active form failed due to the formation of a catalytically irrelevant duplex conformation, which is formed due to a self-complementary sequence within the catalytic core (Nowakowski *et al.*, 1999b, 2000). To avoid the formation of such an artefact, we used a catalytically active mutant in which the self-complementary sequence is disrupted by the introduction of a point mutation Dz839_A5C (Zaborowska *et al.*, 2002).

In order to bind to the DNAzyme and the U1A protein simultaneously, a suitable RNA sequence had to be designed. The RNA substrate (T839) has therefore been elongated by the addition of the U1A specific hairpin loop (T839_hairpin). Since flexible regions in macromolecules often hinder crystallization of the sample, we designed the RNA sequence in the most rigid conformation where the nucleotide at the very 5'-end of the DNAzyme stacks onto the last nucleotide at the 3'-end of the elongated RNA substrate. Next, we verified our strategy by analyzing the formation of the ternary complex consisting of wildtype U1A, RNA substrate and DNAzyme by native PAGE in combination with a nucleic acid dye (Figure 4.2A). As expected, the free protein is not visible with the nucleic acid stain, whereas free DNA or RNA show one single band. As a negative control, we used a sample consisting of protein and DNA and we observe no shift in electrophoretic mobility, indicating that, as anticipated, complex formation between protein and DNA is not possible. DNA and RNA as well as RNA and U1A are capable of forming a binary complex, leading to a significant shift in electrophoretic mobility. The size of the ternary complex is not significantly higher than the binary complex, 33 kDa and 22 kDa, respectively. The small shift in electrophoretic mobility between binary and ternary complexes is therefore in line with the molecular weight ratio. Our results provide evidence that the designed RNA is capable of binding U1A and DNA simultaneously. Analysis of the ternary complex using SEC also supports the quantitative formation of the complex resulting in a monodisperse sample with high purity (Figure S4.4). Our results show that the molecules form a well-defined, monodisperse ternary complex, which is a prerequisite for structural studies.

As a next step, we investigated the structural features of the DNA:RNA:U1A complex in solution using SAXS measurements as well as NMR spectroscopy. Since the exchange of phenylalanine by tryptophan is the only naturally occurring mutation, we used the U1A F56W variant for this study. To analyze the global shape of the ternary complex and to obtain information on the homogeneity of our sample we performed SAXS measurements with the ternary DNA:RNA:protein complex at different concentrations and extracted the form factor of the com-

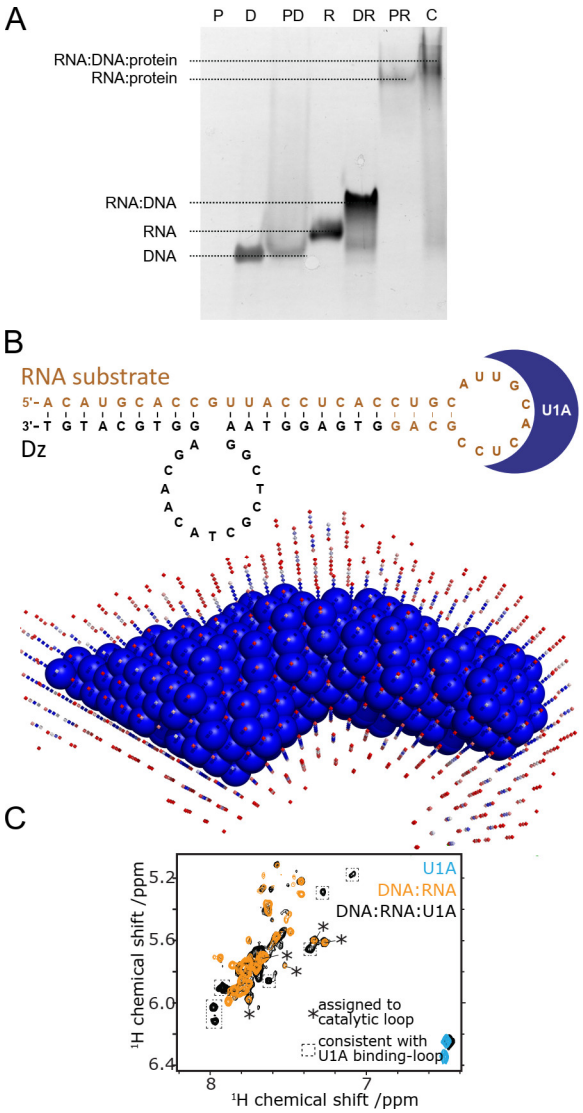


Figure 4.2: Formation and analysis of a ternary DNAzyme:RNA:U1A complex. (A) Native PAGE visualizes the formation of a ternary DNA:RNA:protein complex. P: U1A protein, D: DNAzyme, R:RNA substrate, C: ternary complex (additional information available in the supplementary data file). (B) Top: Scheme of the ternary complex consisting of the RNA-binding protein U1A in blue, the DNA in black and the RNA substrate in red. Bottom: *Ab initio* reconstruction of the low-resolution shape of the ternary complex from SAXS data analysis. (C) Section of homonuclear TOCSY spectra of ternary complex (black), DNA:RNA complex without protein (orange), and U1A protein alone (light blue). Signals corresponding to cytosines within the DNAzyme catalytic core (asterisk) are not affected by U1A binding, while shifted signals (boxes) indicate uracil/cytosine signals in the RNA hairpin loop (see Figure S4.6 for larger extract of spectra).

plex at infinite dilution (Figure S4.5). SAXS measures the contrast in scattering between the macromolecule and buffer, thus, it is particularly well suited for nucleic acids, which have high contrast due to the electron-rich phosphate backbone (Burke & Butcher, 2012). In line with our previous results from size-exclusion chromatography, the SAXS analysis shows that no aggregates are formed during the assembly of the ternary complex. The reconstruction of a low-resolution model (Franke & Svergun, 2009) (Figure 4.2B) shows an elongated structural

model of about 11 nm length, which is consistent with roughly 23 base pairs (0.34 nm/bp) and the protein with a diameter of 3 nm. The less bulky part on the right side of the molecule with a diameter of 3 nm is in agreement with the size of the U1A protein, therefore, we assume that the bulky part of the model can be assigned to the catalytic core of the DNAzyme including the complementary part of the RNA substrate. To test this hypothesis we have fitted a structural model of the DNA:RNA:protein complex into the low-resolution data obtained by SAXS measurements (Figure S4.7). The results support the idea that the catalytic core of the DNAzyme can be accommodated by the bulky part of the bead model. We then performed solution NMR spectroscopy to check for a potential interference of the U1A protein with the structure of the DNA:RNA complex, in particular with the region relevant for catalysis. A comparison of the homonuclear 2D TOCSY spectra of the U1A-RBD F56W protein and the DNA:RNA complex with the ternary DNA:RNA:U1A complex (Figure 4.2C and S4.6) shows that the few observed spectral changes are localized in the RNA hairpin extension. In contrast, all peaks that can be linked to the catalytic loop do not shift between the binary DNA:RNA or the ternary DNA:RNA:U1A complex. These results confirm that the presence of the U1A protein does not affect the structure of the catalytic core of the DNA:RNA complex. We then investigated the influence of binding to U1A on the catalytic performance of the DNAzyme using a cleavage assay. Our results show that binding of the U1A protein to the hairpin extension of the RNA substrate does not influence the activity of the DNAzyme (Figure S4.8) leading to the conclusion that structure and function of the DNAzyme are not affected by the formation of the ternary complex.

4.5.3 Structures of the U1A triple mutants in absence of RNA

To examine the structure in detail and to evaluate the crystallization properties of all U1A variants, we aimed at crystallizing all triple mutants. We obtained protein crystals for the variants A2W, F56W, R70W and K98W in ammonium sulfate conditions, the A2W protein did not lead to high quality data and we were not able to solve the structure. The crystal structures of the three U1A variants reveal the same overall architecture as the U1A wildtype (Figure 4.3A) with RMSD values of 0.32 Å for F56W, 0.29 Å for R70W, and 0.21 Å for K98W. The positions of most secondary structural elements are virtually identical, whereas the terminal regions vary slightly. Interestingly, symmetry and content of the AU are different for each variant, due to the different amino acids involved in formation of the crystal contacts. Variant F56W (Figure 4.3B) crystallized in the trigonal space group $P3_121$ with three molecules per AU, whereas variant R70W (Figure 4.3C) crystallized in the orthorhombic space group $P2_12_12_1$ and with four molecules per AU. Even an amino acid exchange at the very last position in the sequence led to altered crystal contacts compared to the U1A wildtype. While the wildtype U1A-RBD crystallized in the body centered cubic space group ($I4_132$), mutation K98W (Figure 4.3D) resulted in the primitive tetragonal space group $P4_32_12$ with four molecules per AU. The crystallographic and refinement parameters are summarized in Tables S4.3 and S4.4.

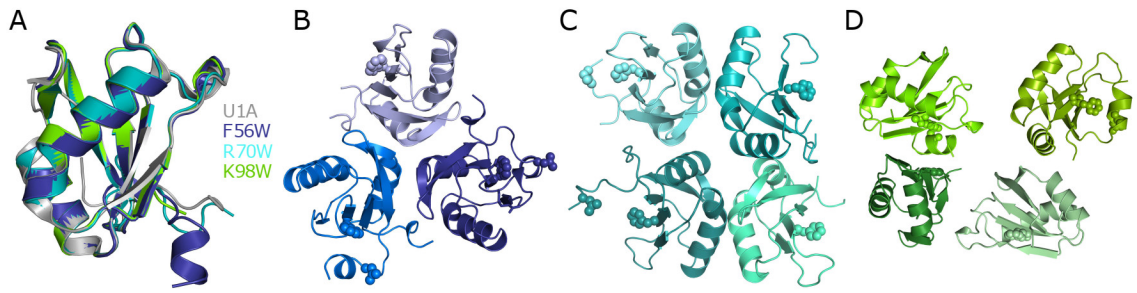


Figure 4.3: Crystal structures of U1A triple mutants in absence of RNA. (A) Superposition of U1A-RBD wildtype (Nagai *et al.*, 1990) in grey with the U1A variants F56W in blue; R70W in cyan, and K98W in green. Biological assembly in the asymmetric unit of the protein crystals formed by the U1A variants F56W (B), R70W (C), and K98W (D). Residues 56 and 92 are shown as spheres.

Our structural data show that while the exchange of the amino acids at positions 56, 70, and 98 by tryptophan does not influence the structure or RNA-binding properties in solution significantly, it does have a dramatic impact on the assembly in the unit cell. The orientation of the protein molecules in the AU is different for each mutant, which impacts the accessibility of the RNA-binding site. Based on the crystal structures we have estimated the accessibility of the binding pocket by measuring the distance between residue 56 located deep inside the RNA-binding cavity and the closest corresponding residue of a neighboring molecule within the same AU or symmetry-related molecules (Figure S4.9). The distance increases from K98W with 12 Å to R70W with 16 Å and F56W with 19 Å, suggesting that the RNA-binding site is most accessible for F56W in the crystalline form. When analyzing the crystal packing, we also observed that the RNA-binding site is oriented towards the bulk solvent in the structures of F56W and R70W, whereas it is facing towards the inside of the biological assembly in the structure of K98W. Although the termini of the amino acid sequences are not well-defined, no difference electron density was observed that would suggest an obstructed RNA-binding site in the F56W structure. In the structure of R70W, some positive difference density suggests that the C-terminus may hinder the access to the RNA binding pocket. Unfortunately, this density could not be modelled satisfactory. Interestingly, previous studies of the RNA-binding domain that is slightly longer than the construct that we have been working on and consists of residues 2-102 have shown that the C-terminus is critical for the formation of U1A dimers (Varani *et al.*, 2000). Residues 93-102 form helix c which forms the protein-protein interface in the dimer. The excess electron density close to the C-terminal region in R70W could indicate a partial formation of helix C. Furthermore, the different arrangement in U1A variant K98W may be explained by the influence of the C-terminal residues on protein-protein interaction.

4.5.4 Binding of U1A variants to the U1 RNA hairpin motif in the crystalline form

To ameliorate the process of generating U1A:RNA complex crystals, we established a protocol for soaking U1A crystals with the U1 RNA hairpin motif in combination with an efficient

method of estimating the success of RNA binding exploiting the fluorescence of the introduced tryptophan residue. We produced crystals of the U1A variants F56W, R70W, and K98W, which displayed tryptophan fluorescence when irradiated with UV light (Figure 4.4, left column). Subsequently, we treated these crystals with a solution containing the specific U1 RNA hairpin motif. The RNA sequence consists of 21 nucleotides and should fold into a characteristic hairpin motif, which binds into the RNA-binding pocket of the U1A protein (Table S4.2). Inspection of the crystals using bright light shows that the crystals were not severely damaged by the treatment (Figure 4.4, middle column). Exposure of the preformed protein crystals with the U1A specific RNA motif lead to a drastically reduced fluorescence in case of the F56W variant (Figure 4.4, right column). We speculate that the loss in fluorescence intensity is a result of fluorescence quenching due to the interaction of RNA and the tryptophan residue at position 56. Notably, the lack of fluorescence could also be a consequence of absorption of the UV light by the RNA.

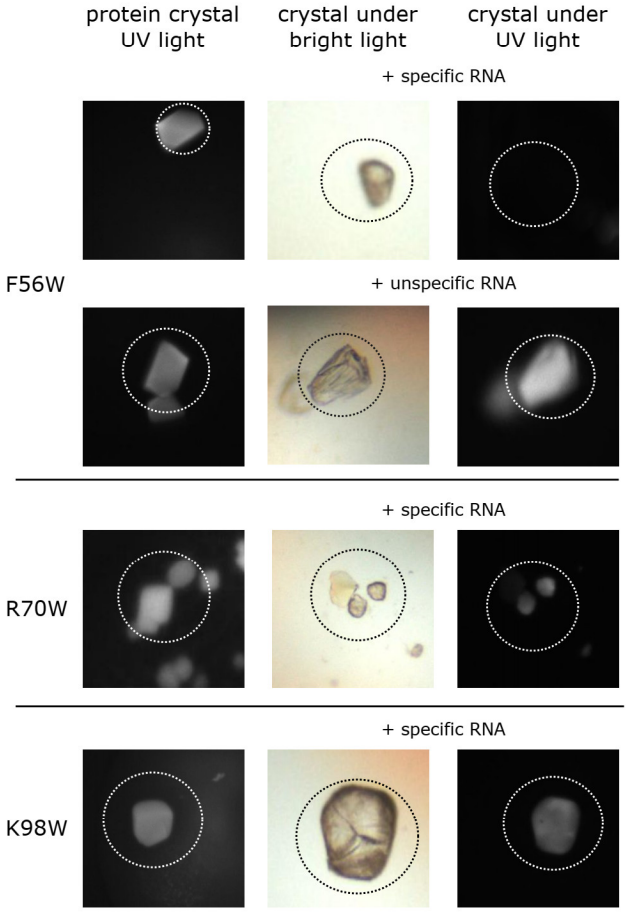


Figure 4.4: Binding of RNA to U1A variants *in crystallo*. Protein crystals are UV active in the absence of RNA (left column). Soaking of RNA did not lead to dissolution of the same crystals (middle column). Addition of the specific U1 RNA hairpin motif to the solution leads to quenching of the fluorescence in case of F56W, but not for R70W and K98W (right column), whereas addition of an unspecific RNA retains the fluorescence (only shown for F56W).

To validate that the decrease in fluorescence is an indication for RNA binding to U1A, we performed this experiment using an unspecific RNA sequence (T839, see Table S4.2) that is not capable of binding to the protein as a control. In line with our hypothesis, the crystals retain their fluorescence after treatment with an unspecific RNA. In contrast, soaking of the specific RNA hairpin motif into preformed R70W and K98W crystals does not lead to a significant reduction in fluorescence intensity. To determine whether fluorescence quenching correlates with binding of the RNA to U1A, we analyzed the crystals after soaking with specific and unspecific RNA. The crystals were dissolved in RNA loading buffer and analyzed by denaturing PAGE using a fluorescent nucleic acid dye (Figure S4.10). We detect RNA in all samples that were soaked with the specific RNA hairpin motif, whereas no RNA was present in the sample with unspecific RNA. Our results show that we detect an increased amount of nucleic acids in F56W crystals compared to R70W and K98W crystals, suggesting that a higher amount of RNA binds to the F56W protein *in crystallo*. These findings support the hypothesis that fluorescence quenching is a result of RNA binding to the F56W variant. In line with these observations, all structures that were determined of the R70W and K98W variants soaked with specific RNA lack additional electron density that could be assigned to RNA. Taken together, our investigation of the RNA binding properties of the U1A variants F56W, R70W, and K98W suggest that while they are all capable of binding the RNA hairpin motif in solution, only the F56W variant is capable to incorporate the RNA sequence *in crystallo*.

4.5.5 Structure of the F56W variant bound to RNA

To investigate the molecular arrangements and 3D structure of protein and RNA after soaking and co-crystallization, respectively, we solved the structures of the RNA:F56W complex using crystals obtained by each method (F56W:RNA_{soaked} or F56W:$RNA_{co\text{-}crystal}$, respectively). The best results for the soaking were obtained by adding RNA to a final concentration of 1.5 mM to a preformed crystal followed by an incubation time of 4 d at room temperature. We were able to collect data and refine the structural model of the RNA:F56W complex obtained by soaking to a resolution of 2.37 Å with an R_{work}/R_{free} of 0.248/0.294. Interestingly, the space group of the soaked crystal changed from $P3_221$ for the protein in the absence of RNA into $P6_522$ in the presence of RNA. Therefore, indexing the crystals is sufficient to determine whether RNA is bound to the F56W crystals or not. In addition to the fluorescence quenching, a change in the crystal symmetry is another indicator of the binding of RNA to the F56W variant in the crystalline form. As a consequence, binding of RNA can be assessed quickly by indexing the crystal and determining the space group, rather than collecting data, processing, phasing, and model building.

The structure of F56W:RNA_{soaked} is virtually identical to the previously reported double mutant U1A:RNA complex structure obtained via co-crystallization (Figure 4.5A) (Oubridge *et al.*, 1994). All nucleic acids are well-defined except the residues 13–15 (UCC), which are located in the loop region of the hairpin motif (Figure S4.11). The same nucleotides are also disordered in the double mutant U1A:RNA complex structure (Oubridge *et al.*, 1994). Interestingly,

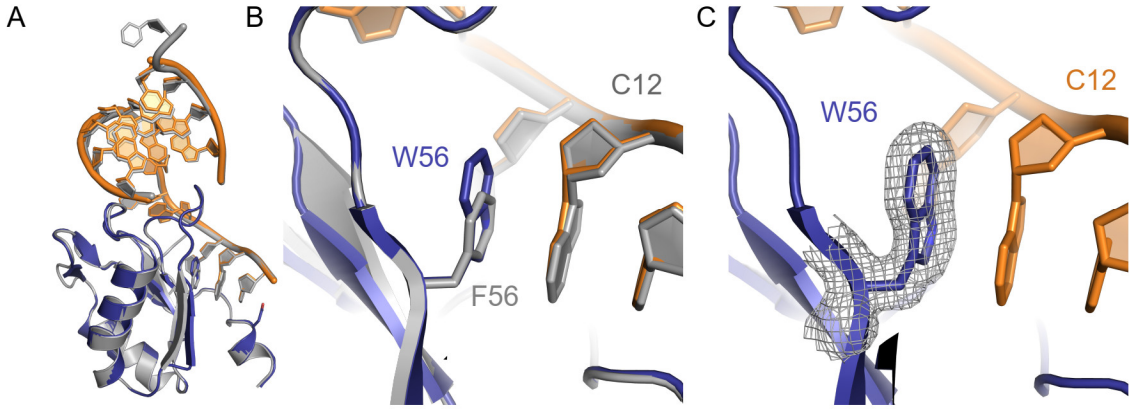


Figure 4.5: Structures of U1A double mutant and variant F56W bound to the RNA hairpin. (A) Superposition of the F56W:RNA$_{co\text{-}crystal}$ complex (protein in blue and RNA in orange) and the previously reported U1A double mutant:RNA complex (Oubridge *et al.*, 1994) (grey). (B) Close-up of the superposition as depicted in A showing the stacking interactions between RNA and the residue at position 56. (C) Close-up of the F56W:RNA$_{co\text{-}crystal}$ complex showing the 2*Fo-Fc* electron density of the introduced amino acid W56 that is involved in RNA binding.

in crystal structures of U1A in complex with large RNA sequences these residues are ordered due to interactions with symmetry related nucleic acids (Ferré-D'Amaré *et al.*, 1998; Rupert, 2002). Residues A11 and C12 in the RNA sequence are crucial for the interaction with the U1A residues 56 (Figure 4.5B) and 92, resulting in a four-element-stack. Besides the amino acid residue at position 56, also Y13, N16, K22, R52, Q54 and N92 are involved in RNA binding. The electron density around the introduced tryptophan at position 56 is well-defined (Figure 4.5C) and shows that the side chain adopts the same orientation as the phenylalanine side chain in the wildtype retaining the four-element-stacking.

Furthermore, we screened for novel crystallization conditions for the RNA:F56W complex to be able to compare the structures obtained by soaking or co-crystallization. We were able to refine the structural model to a resolution of 2.05 Å with an R_{work}/R_{free} of 0.230/0.249. The RNA:F56W complex crystallizes in the hexagonal space group $P6_522$ with three molecules in the AU. This space group is also observed for F56W:RNA$_{soaked}$ and is consistent with the previously reported U1A:RNA complex structure (Oubridge *et al.*, 1994). Superposition of structures from both crystallization methods yield an RMSD value of 0.17 Å (Figure 4.6) demonstrating that both crystallization methods result in virtually identical structures. Our results reveal that binding of the RNA hairpin loop to the protein is not prohibited in the crystalline form. Moreover, the change of the space group upon RNA binding allows for detecting complex formation by indexing. Taken together, the presented soaking approach may be a generally applicable strategy for efficient structure solution for protein:nucleic acid and protein complexes in general.

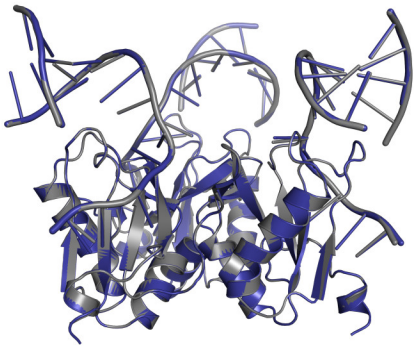


Figure 4.6: Superposition of the RNA:F56W crystal structures obtained by co-crystallization (blue) and soaking experiments (grey).

4.6 Discussion

While the surfaces of proteins in general possess hydrophobic patches that favor intermolecular interactions and counteract the repulsive forces arising from charged functional groups contributing significantly to the formation of protein crystals, the surface of nucleic acids is dominated by a regular array of negatively charged phosphate groups leading to the formation of crystals with poor long-range order. In addition to the phasing problem, this has hampered the structural characterization of nucleic acids using X-ray crystallography for decades. Several strategies have been developed to overcome these challenges, such as the addition of the polyamine spermine (Kim *et al.*, 1971) to the crystallization solution or the application of nucleic acid-binding proteins as crystallization module. This has led to significant progress in the structure elucidation of nucleic acids, which improves our understanding of structural principles of nucleic acids in general. The structure and function of nucleic acids that are capable of adopting unusual three-dimensional conformations and to perform catalysis, such as ribozymes and DNAzymes, is particularly intriguing.

Our presented crystallization strategy relies on using the U1A RNA-binding domain to aid the crystallization process and allows for solving the phasing problem using molecular replacement. Initially, we generated a series of variants of the RNA-binding domain of U1A that contain a tryptophan residue. The U1A-RBD has already been recognized as invaluable tool for the crystallization of RNA (Ferré-D'Amaré & Doudna, 2000a) based on (i) its globular and compact fold (Nagai *et al.*, 1990); (ii) the tight binding of the U1A-RBD to its cognate site (K_d of approximatley 10^{-11} M) (van Gelder *et al.*, 1993) and (iii) the protein-RNA interface comprises polar (salt bridges) and non-polar (stacking of aromatic amino acid side chains between nucleobases) interaction. A key feature of U1A-RBD is that mutations in the solvent-exposed residues affect the solution and crystallization properties of the protein variants (Oubridge *et al.*, 1995).

We have exploited this feature to generate four new U1A variants introducing single tryptophan residues. In addition, we have also performed a bioinformatics analysis to screen for naturally occurring tryptophan residues within U1A sequences discovering that homologs from

Viridiplantae possess a tryptophan residue that is involved in the stacking interaction with nucleobases of the U1 RNA hairpin motif. This directed approach has led to a fifth, and in many aspects most promising, U1A variant. Noteworthy, for large RNAs it has been shown that interactions between nucleic acids from adjacent AU also contribute to the crystal packing (Rupert, 2002), thus, other U1A variants may be more suitable for other RNA sequences.

The presence of a tryptophan residue has several advantages, which make the U1A variants presented in this study a valuable asset for structural studies in solution or in the crystalline form. First, tryptophan fluorescence can be used to determine the protein concentration using absorbance at a wavelength of 280 nm. Second, in comparison to alternative approaches (Raghunathan *et al.*, 2010; Sumner & Dounce, 1937; Wilkosz *et al.*, 1995), tryptophan fluorescence is a fast, reliable, and non-invasive tool to distinguish between protein and salt crystals. Third, the change in fluorescence intensity can be used to detect the binding of the RNA hairpin loop *in crystallo*. We have shown that the fluorescence intensity of F56W crystals decreases upon RNA binding (Figure 4.4), which most likely is a result of the stacking interaction of tryptophan with the nucleobases. To our knowledge, the detection of RNA binding to a protein in the crystalline form using tryptophan fluorescence has not been reported so far.

In previous studies, the structure determination of U1A in complex with RNA was achieved using co-crystallization. This is probably the most common method for obtaining crystals of a protein:ligand complex, where the ligand is added to the protein to form a complex that is subsequently used in crystallization trials (Hassell *et al.*, 2007). Ligands are usually well tolerated in crystal structures, thus, crystals of protein–ligand complexes can often also be obtained by soaking crystals with ligands. Owing to the ease of the method this is frequently the first approach to obtain the structure of a protein bound to small molecules or ions. The potential ligand can access the binding sites by diffusing through solvent channels within the crystal lattice, as long as the sites are not involved in crystal packing or the site is not blocked by parts of the protein. Although soaking of nucleic acids into protein crystals has sporadically been reported it is by far not a common strategy in structural biology (Horn *et al.*, 2004; Valegård *et al.*, 1994; Wojtas & Abrescia, 2012).

A prerequisite for soaking is the existence of a soakable crystal form, which can be assessed by inspection of the crystal structure of the protein in absence of RNA. Therefore, we have inspected the atomic model as well as the electron density map, to ensure that flexible regions are not blocking the binding site. The F56W structure shows that no close crystal-packing contacts are present in the vicinity of the RNA-binding site, whereas we observe some electron density near this site in the R70W structure and close contacts in the K98W structure. Thus, the F56W variant has more favorable characteristics for soaking experiments. Another important feature of U1A is that we expect no large conformational rearrangements upon RNA binding, which is crucial for soaking macromolecules into crystals.

The benefit of soaking is that once the crystallization conditions are optimized it is easy to produce a large amount of crystals (Hassell *et al.*, 2007). Although the conditions for soaking ligands require some optimization in terms of ligand concentration and soaking time, the process is significantly faster and low amounts of ligands are required. When working with fragile

ligands, such as RNA strands that are easily degraded by RNases or under the harsh conditions of crystallization solutions, soaking also improves the chances of preserving the ligand. The amounts of RNA required for soaking are also significantly lower, which makes this approach very cost effective. In general, soaking ligands into pre-existing crystals bears the risk of disrupting the crystal lattices and destroying the crystals. As mentioned previously, the binding of RNA involves no major structural rearrangements, which minimizes the risk of disrupting the crystal packing. Another issue is the crystal-to-crystal variation, even within the same crystallization drop. We have demonstrated that the new F56W variant results in UV active crystals and RNA-binding leads to a diminished fluorescence (Figure 4.4), thus, binding of the ligand can quickly be assessed using a crystallization imager with UV fluorescence detection. The lack of fluorescence of U1A variant crystals after soaking is therefore a good indication of ligand occupancy and can be used to overcome the issue of crystal-to-crystal variation. To ensure that the full range of conformational changes does not vary between the structures obtained from soaked or co-crystals we validated the soaking system with co-crystallization experiments. Our structural analysis confirms that the structures obtained by soaking and co-crystallization are virtually identical. In summary, we have developed the first soaking system for an U1A variant, which provides an efficient method to distinguish between protein:RNA complex and free protein crystals and to obtain protein:RNA complex structures.

4.7 Conclusion

In this study we have generated new variants of the RNA-binding protein U1A for the application in crystallization and phasing of RNA or DNA:RNA complex structures. We have determined three crystal structures of U1A variants F56W, R70W, and K98W, which revealed insights into the architecture of the protein within the AU as well as the accessibility of the RNA-binding site. Using a combination of native PAGE, NMR spectroscopy and SAXS measurements we were able to show that the F56W protein binds to RNA and DNA:RNA complexes, expanding the scope of applications for U1A. We introduced an RNA hairpin motif into preformed protein crystals by soaking and used fluorescence-based imaging to detect RNA binding to the protein in crystalline form. Finally, we report crystal structures from soaked crystals and co-crystals of the F56W:RNA complex, revealing that the soaking protocol results in a virtually identical structure. Taken together, our results expand the crystallographic toolbox for the structure determination of nucleic acids and to help unveiling structure of new RNA and DNA:RNA complexes. We expect that an increased number of structures will ultimately lead to a better understanding of the function of catalytically active nucleic acids.

4.8 Supporting data

Supplementary Information for Figure 4.2A

As expected, the free protein is not visible with the nucleic acid stain, whereas free DNA or RNA show one single band. As a negative control, we used a sample consisting of protein and DNA and we observe no shift in electrophoretic mobility, indicating that, as anticipated, complex formation between protein and DNA is not possible. DNA and RNA as well as RNA and U1A are capable of forming a binary complex, leading to a significant shift in electrophoretic mobility. The size of the ternary complex is not significantly higher than the binary complex, 33 kDa and 22 kDa, respectively. The small shift in electrophoretic mobility between binary and ternary complexes is therefore in line with the molecular weight ratio.

Supplementary Tables

Table S4.1: Primer sequences used in this study.

Name	Sequence 5′ – 3′
U1AA2W_for	CTG TAT TTT CAG GGA CAT ATG TGG GTT CCG GAA ACC CGT CC
U1AA2W_rev	GGA CGG GTT TCC GGA ACC CAC ATA TGT CCC TGA AAA TAC AG
U1AH10W_for	GTT GTT GAT ATA AAT GGT CCA ATT CGG ACG GGT TTC CGG AAC TGC CAT A
U1AH10W_rev	TAT GGC AGT TCC GGA AAC CCG TCC GAA TTG GAC CAT TTA TAT CAA CAA C
U1AF56W_for	CCT GAA AAT GCG TGG TCA GGC ATG GGT TAT TTT CAA AGA AG
U1AF56W_rev	CTT CTT TGA AAA TAA CCC ATG CCT GAC CAC GCA TTT TCA GG
U1AR70W_for	TTA TCA TAA AAC GGA AAA CCC CAC ATG CTA CGC AGT GCA TTG G
U1AR70W_rev	CCA ATG CAC TGC GTA GCA TGT GGG GTT TTC CGT TTT ATG ATA A
U1AK98W_for	GCG ATA TTA TCG CCA AAA TGT GGT AAC TAG CAT AAC C
U1AK98W_rev	GGT TAT GCT AGT TAC CAC ATT TTG GCG ATA ATA TCG C

Table S4.2: Oligonucleotide sequences used in this study.

Name	Sequence 5′ – 3′
T839_hairpin	ACA UGC ACC GUU ACC UCA CCU GCA UUG CAC UCC GCA G
Hairpin	AAU CCA UUG CAC UCC GGA UUU
T839	ACA UGC ACC GUU ACC CCA A
Dz839_A5C	GTG AGG TAA GGC TCG CTA CAA CGA GGT GCA TGT

Table S4.3: Data collection and refinement statistics for the crystal structures F56W (PDB ID 6SQT), R70W (PDB ID 6SQV), K98W (PDB ID 6SR7).

	F56W (6SQT)	R70W (6SQV)	K98W (6SR7)
Data collection			
Space group	$P3_221$	$P2_12_12_1$	$P4_32_12$
Cell dimensions			
a, *b*, *c* [Å]	89.0, 89.0, 85.6	76.1, 76.1, 136.9	76.3, 76.3, 151.5
α, β, γ [°]	90, 90, 120	90, 90, 90	90, 90, 90
Resolution [Å]	38.89-1.844	51.02-2.45	68.11-1.86
	(1.87-1.84)	(2.49-2.45)	(1.89-1.86)
R_{merge} [%]	5.3 (89.8)	4.9 (172.7)	13.9 (88.6)
I / σ(I)	25.4 (0.7)	25.8 (1.4)	14.8 (2.4)
Completeness [%]	82.1 (13.4)	100.0 (100.0)	100 (99.6)
Multiplicity	13.9 (1.2)	12.8 (12.8)	24.1 (15.2)
CC(1/2)	1.0 (0.27)	1.0 (0.66)	1.0 (0.54)
Refinement			
Resolution [Å]	38.92-1.84	51.08-2.45	68.31-1.86
No. reflections	27593	28682	36417
R_{work}/R_{free}	21.34/24.46	24.05/27.34	22.14/25.60
No. atoms			
Protein	2198	2806	2863
Ligand (RNA) + ion	0	0 + 15	0
Water	73	5	78
B-factors			
Protein	43.94	82.45	29.75
Ligand (RNA) + ion	–	116.82	–
Water	37.26	69.28	28.68
R.m.s. deviations			
Bond lengths [Å]	0.0145	0.0143	0.0142
Bond angles [°]	1.91	1.98	1.89

Table S4.4: Data collection and refinement statistics for the crystal structures RNA:F56W obtained by co-crystallization (PDB ID 6SQN) and RNA:F56W obtained by soaking (PDB ID 6SQQ).

	RNA:F56W$_{\text{co-crystallized}}$ (6SQN)	RNA:F56W$_{\text{soaked}}$ (6SQQ)
Data collection		
Space group	$P6_522$	$P6_522$
Cell dimensions		
a, *b*, *c* [Å]	96.6, 96.6, 257.6	96.2, 96.2, 258.2
α, β, γ [°]	90, 90, 120	90, 90, 120
Resolution [Å]	51.74-2.05	51.60-2.37
	(2.09-2.05)	(2.41-2.37)
R_{merge} [%]	19.7 (235.7)	28.7 (257.5)
I / σ(I)	18.2 (1.9)	12.4 (1.4)
Completeness [%]	100.0 (99.9)	100.0 (99.9)
Multiplicity	38.0 (39.6)	38.2 (40.2)
CC(1/2)	0.999 (0.683)	0.998 (0.753)
Refinement		
Resolution [Å]	48.52-2.05	51.03-2.37
No. reflections	46908	28072
R_{work}/R_{free}	22.97/24.88	24.8/29.4
No. atoms		
Protein	2335	2277
Ligand (RNA) + ion	993 + 0	1008 + 1
Water	95	26
B-factors		
Protein	37.10	44.33
Ligand (RNA) + ion	46.81	53.55 + 49.26
Water	33.94	36.11
R.m.s. deviations		
Bond lengths [Å]	0.127	0.0126
Bond angles [°]	1.81	1.82

Supplementary Figures

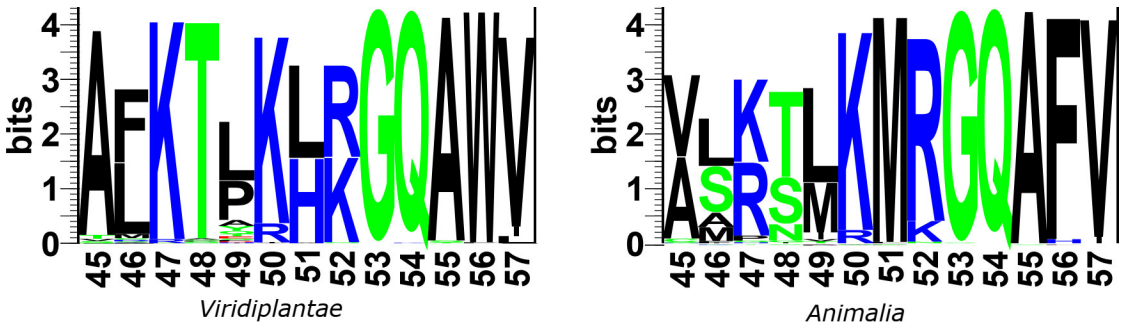


Figure S4.1: Sequence logos (Schneider & Stephens, 1990) of the central part of the RNA recognition motif (Pfam entry RRM_1, PF00076) for about 500 U1A homologs from *Viridiplantae* (left) and from *Animalia* (right). In Animalia phenylalanine is the predominant amino acid at position 56, whereas tryptophan is highly conserved in the group of Viridiplantae.

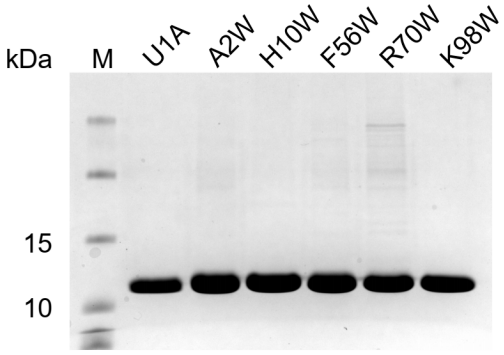


Figure S4.2: SDS polyacrylamide gel (15%) to analyze the purity of the different U1A variants after CHT1-column.

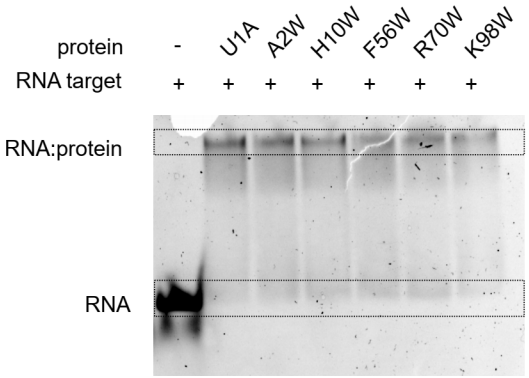


Figure S4.3: Native PAGE showing that the novel U1A variants retain their ability to bind the RNA hairpin. Lane 1 shows the RNA target in absence of protein, lane 2 the U1A Y31H\Q36R mutant, and lanes 3-7 the new U1A variants containing a triple mutation. Protein and RNA were incubated at a 1:1 ratio and a concentration of 10 µM each. The intensity of the bands deriving from protein:RNA complexes is weaker than the bands deriving from free RNA, since the protein hinders the interaction of the nucleic acid dye with RNA.

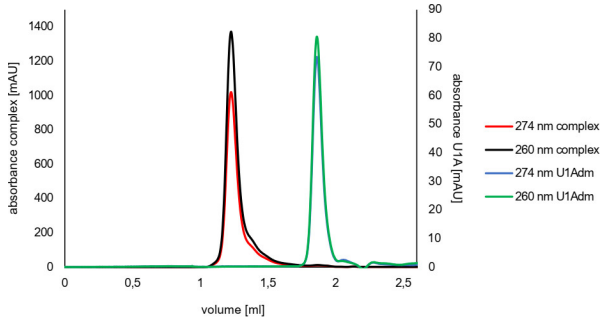


Figure S4.4: Formation of the ternary DNA:RNA:U1A complex. Analytical SEC run demonstrates a monodisperse ternary complex.

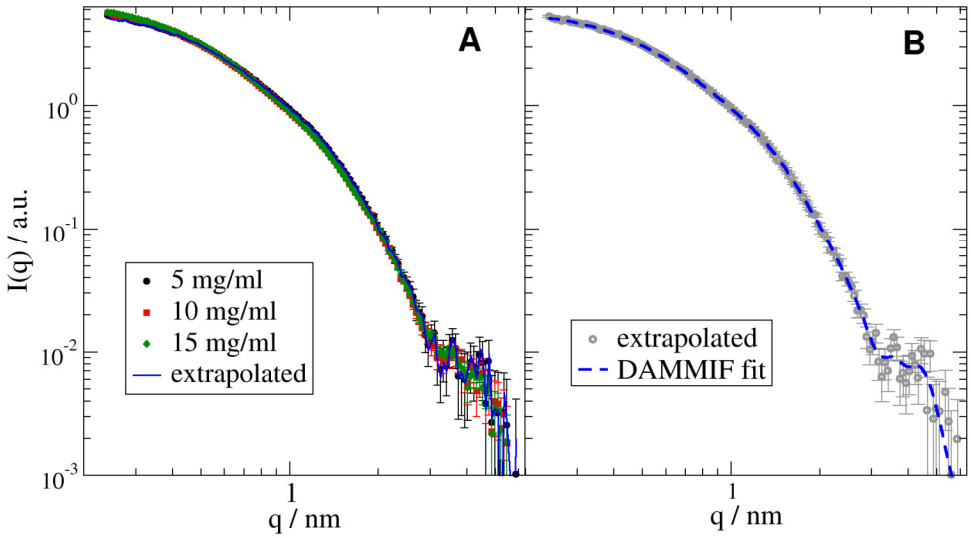


Figure S4.5: Analysis of the SAXS measurements. (A) Background corrected and concentration scaled SAXS data with the ternary complex at concentrations of 5, 10, 15 mg/ml. The data are extrapolated for each scattering vector individually to zero concentration to extract the form factor without influence of a structure factor due to interparticle interactions. (B) Exemplary comparison of a single DAMMIF result to the original form factor. We observe excellent description of all details of the measured form factor. The quality of other DAMMIF results is of the same quality.

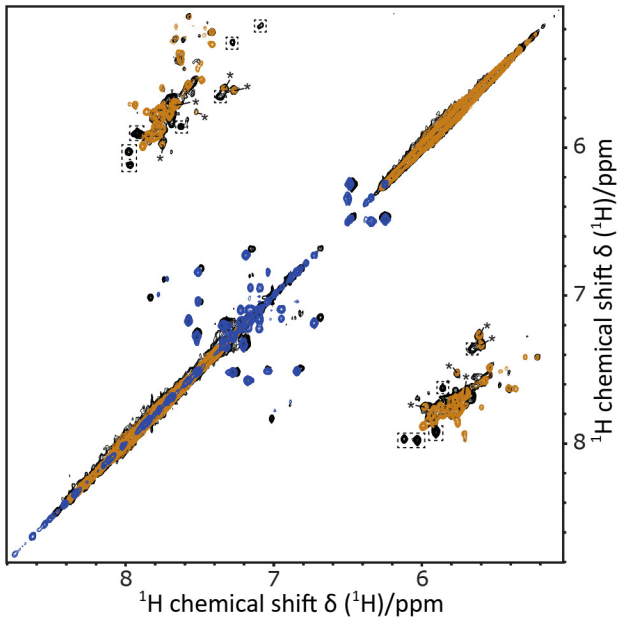


Figure S4.6: Homonuclear TOCSY spectra of ternary complex (black), DNA:RNA complex without protein (orange) and U1A protein alone (blue). Signals corresponding to cytosines within the catalytic loop (asterisk) are not affected by U1A binding, while shifted signals (boxes) indicating uracil/cytosine signals in the binding loop.

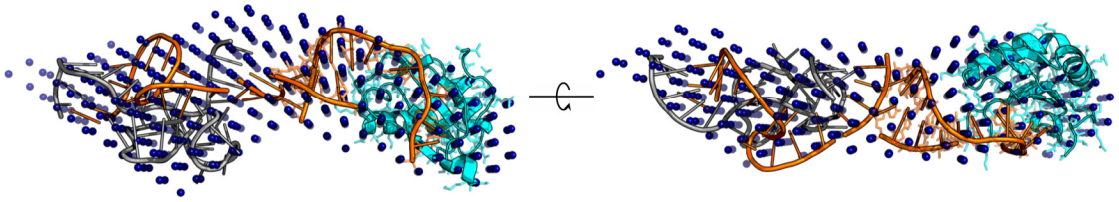


Figure S4.7: DNA:RNA:protein complex model fitted into the low-resolution shape obtained by SAXS with the DNAzyme shown in gray, RNA shown in orange, and U1A protein in cyan. The fitted model is shown from two different perspectives.

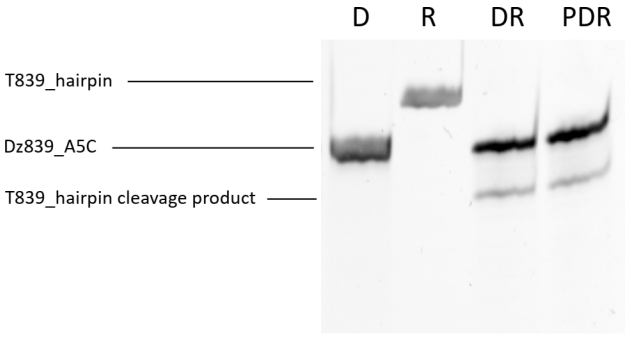


Figure S4.8: Cleavage assay of the DNAzyme in absence and presence of the U1A protein. PAGE analysis of the reaction mixtures containing DNAzyme (D), RNA substrate T839_hairpin (R), DNAzyme:RNA complex (DR), and protein:DNA:RNA coplex (PDR) shows that the presence of the protein does not influence the cleavage reaction.

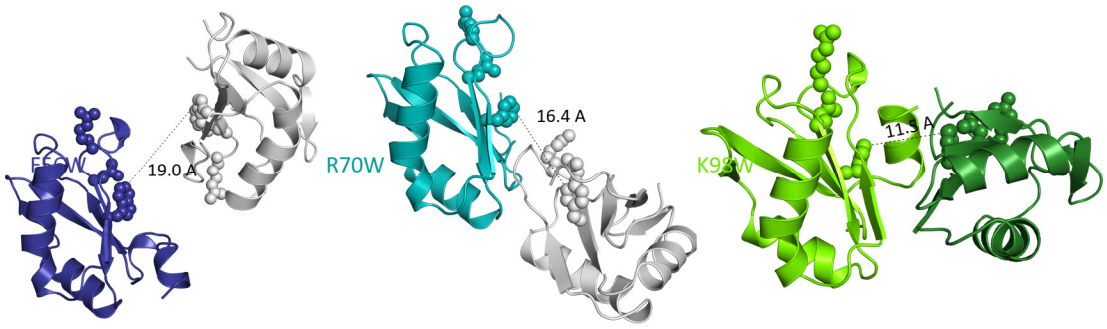


Figure S4.9: Orientation of one molecule of the asymmetric unit towards the closest symmetry-related molecule in grey (in case of the U1A variants F56W and R70W) or another molecule within the same AU, respectively (in case of the U1A variant K98W). The residues at position 52, 54 and56 are shown as balls representing the RNA binding site. The distances are measured between the amino acid residues at position 56 that are involved into RNA binding and can therefore be used as a reference for the accessibility of the RNA-binding site.

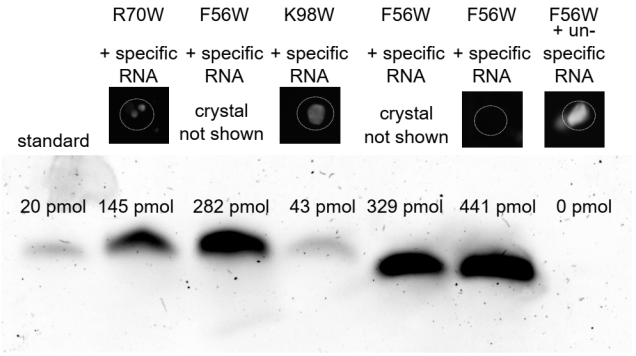


Figure S4.10: Analysis of bound RNA after soaking experiments with preformed protein crystals of the U1A variants F56W, R70W and K98W by denaturing PAGE (18%). The amount of RNA was estimated using the volume tool of the ChemiDoc MP System (Bio-Rad, Hercules, CA, USA). As a standard we used an amount of 20 pmol hairpin RNA. The samples were visualized using the nucleic acid stain GelRed.

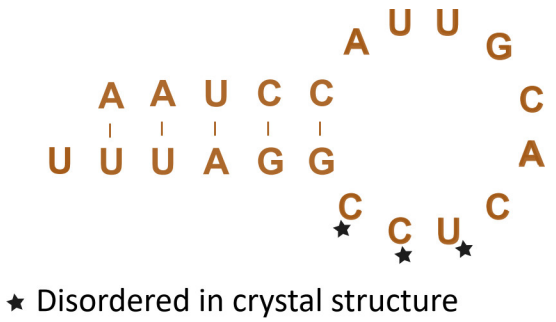


Figure S4.11: U1 hairpin loop sequence with the disordered residues 13—15 (UCC), which are located in the loop region of the hairpin motif.

5 Approaches to crystallize the 10-23 DNAzyme in a biologically relevant conformation

This chapter reflects content of the following publication.

5.1 Publication information

Hannah Rosenbach, Julian Victor, Thomas Pauly, Luitgard Nagel-Steger, Manuel Etzkorn, Gerhard Steger, and Ingrid Span

To be submitted to: **Crystals**

5.2 Abstract

X-ray crystallography is currently the most favoured technique for structure determination of nucleic acids. However, structure elucidation by X-ray diffraction requires well-ordered single crystals of high quality. For this purpose, large amounts of DNA or RNA in high purity and structural homogeneity are of great importance. Structure elucidation of the 10-23 DNAzyme in a biologically relevant conformation faces three mayor challenges: (i) The stabilization of the substrate to capture the DNA:RNA complex in the active conformation, (ii) preventing the formation of an artificial duplex conformation due to a self-complementary sequence in the catalytic core of the DNAzyme, and (iii) the difficulty to crystallize nucleic acids due to their surface properties. We address general challenges that arise with the crystallization of nucleic acids and discuss a variety of different strategies to crystallize the 10-23 DNAzyme in complex with its RNA substrate in a native conformation. These strategies include co-crystallization with the RNA-binding protein U1A; combinatorial screens for suitable crystallization conditions for the DNAzyme:RNA complex with varying length of the RNA substrate and the binding arms of DNAzyme resulting in duplexes with and without overhangs; DNAzymes with reduced flexibility within the catalytic core due to nucleotide deletion; and tetrahedral nanostructures consisting of three DNAzyme molecules and their RNA substrates.

5.3 Introduction

X-ray crystallography is an excellent method to obtain insights into the structure of nucleic acids on an molecular level as well as into structural relationships between nucleic acids and metal ions. However, structure elucidation of nucleic acids using X-ray crystallography is often impeded by their regularly ordered and negatively charged phosphate backbone that can lead to crystals with a poor long-range order. Approaches to tackle this problem involve co-crystallization of the nucleic acid with a crystallization helper such as the human U1 small nuclear ribonucleoprotein A (U1A) that was previously used for the crystallization of several large RNAs including ribozymes and riboswitches (Cochrane *et al.*, 2007; Ferré-D'Amaré *et al.*, 1998). The first crystal structure of an RNA-cleaving DNAzyme, namely the 8-17 DNAzyme (Santoro & Joyce, 1997), was obtained in the presence of the African swine fever virus DNA polymerase X (*Asfv*PolX) (Liu *et al.*, 2017). Proteins that serve as crystallization helpers can bind to a specific nucleic acid motif as it is the case for the U1A protein that specifically binds an RNA hairpin loop (Oubridge *et al.*, 1994), or to double-stranded DNA sequences as it is the case for *Asfv*PolX (Chen *et al.*, 2017c). However, what proteins serving as crystallization helpers for the crystallization of nucleic acids have in common is that they support the formation of well-ordered crystals due to their diverse surface properties and compact fold.

Furthermore, approaches to successfully crystallize nucleic acids involve a careful consideration of their sequence and length. This is true for both approaches involving the use of a protein as a crystallization helper or solely crystallizing the nucleic acids of interest, since the success of crystallization primarily depends on the nature of the nucleic acid molecules, rather than the composition of the crystallization buffers (Ferré-D'Amaré & Doudna, 2000b). In most cases, the nucleic acid that is aimed at being crystallized comprises a conserved motif or sequence of interest, while other regions may be varied, or even deleted or added. To obtain highly ordered crystals, it is beneficial to reduce the flexibility by minimizing the size of the nucleic acid. Additionaly, the crystal packing has to be considered (Holbrook & Holbrook, 2001). One of the most successfull methods for improving the quality of nucleic acid crytals is to vary the length and sequence of double helices present in a molecule (Jordan *et al.*, 1985). This approach is called helix-engineering approach. For DNA helices a length of 10 or 12 base pairs and for RNA helices 6, 8 or 12 base pairs are favorable since this results in an integral or simple fraction of the helical repeat length (Holbrook & Holbrook, 2001). Double helices are usually involved in crystal contacts resulting from stacking interactions: The strength of the stacking interaction is mediated by the sequence of the terminal nucleotides, while the distance to the neighbouring molecules within the crystal lattice is defined by the length of the helix. The generation of overhangs within the double helices can result in intermolecular base pairing, i. e. new crystal contacts between neighbouring molecules. Modification of the helices by either shortening or extending one of the strands led to several crystal structures of protein:DNA complexes, ribozymes, and small oligonucleotides (Anderson *et al.*, 1996; Jordan *et al.*, 1985; Scott *et al.*, 1995).

The 10-23 DNAzyme comprises a catalytic core that consists of 15 nucleotides and shows high flexibility. This flexibility might be an obstancle for the crystallization of the DNAzyme:RNA complex. In 2000, Okumoto & Sugimoto reported on a reduced variant of the 10-23 DNAzyme that lacks the nucleotides at position 5 to 8 within the catalytic core. This DNAzyme variant is still active in the presence of Ca^{2+} ions. Five years later, Zaborowska *et al.* reported that the nucleotide at position eight of the catalytic core region can be deleted without severe effects on the catalytic activity. We use both reduced variants of the 10-23 DNAzyme for crystallization experiments in order to obtain sequences with reduced flexibility, since this might be beneficial for the crystallization properties of the nucleic acid complex.

Although the success of crystallization of nucleic acids highly depends on the properties of the biological sample, purity of the nucleic acids, buffers, additives such as divalent metal ions and polyamines, and precipitants should not be disregarded. Nucleic acids can be synthesized either chemically or enzymatically. Chemical synthesis of RNA is more difficult than the chemical synthesis of DNA, since RNA provides an additional 2'-hydroxyl group that needs to be protected during and deprotected after the synthesis. However, incorporation of non-natural nucleotides is possible during the sythesis process, making chemical synthesis of nucleic acids for crystallographic purposes superior over enzymatic approaches. Enzymatic approaches include the production of DNA by polymerase chain reaction (PCR) and plasmid amplification or the production of RNA from a DNA template harboring a promotor for the T7 RNA polymerase (Holbrook & Holbrook, 2001). Purification of nucleic acids can be performed by either polyacrylamide gel electrophoresis (PAGE) or liquid chromatography such as High-performance liquid chromatography (HPLC) (François *et al.*, 2005). Buffers used for nucleic acid screens usually range between pH 5 and pH 8, to prevend hydrolysis of the molecules. Preferred buffers are acetate pH 4.5, citrate pH 5.5, cacodylate 6.5, or 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.5 (Holbrook & Holbrook, 2001). Potassium cacodylate pH 6.0–6.5 is a commonly used buffer, since the low pH prevents autocatalytic cleavage of RNA sequences (Golden & Kundrot, 2003). Furthermore, the addition of mono- and divalent metal ions were reported to have a great effect on the nucleic acid crystals (Correll *et al.*, 1999; Golden & Kundrot, 2003; Pley *et al.*, 1994b), and the addition of polyamines such as spermine or spermidine had favourbale effects as well as temperatures between 28 °C and 37 °C (Golden & Kundrot, 2003). Here, we present different strategies for the crystallization of the RNA-cleaving 10-23 DNAzyme in complex with its native RNA substrate in a catalytically relevant conformation. For this purpose, it is necessary to stabilize the RNA substrate against DNAzyme-mediated cleavage. We use RNA targets that are stabilized at the cleavage site through substitution of the 2'-hydroxyl group by a fluorine atom. Previous approaches to crystallize the 10-23 DNAzyme in complex with its RNA substrate resulted in the formation of a biologically irrelevant dimer during crystallization (Nowakowski *et al.*, 1999b). This duplex formation was due to a self-complemetary sequence within the catalytic core of the DNAzyme. However, single substitution of each nucleotide in the self-complementary sequence within the catalytic core of the 10-23 DNAzyme (5'CGATCG3'; see Figure 5.1) showed that certain nucleotides of this sequence can be exchanged while catalytic activity is retained (Zaborowska *et al.*, 2002). In this study, we used two variants of the

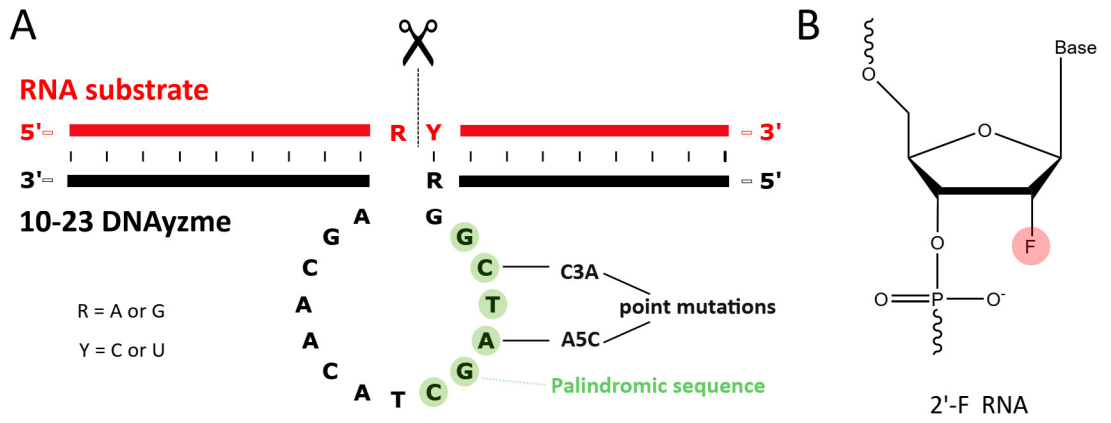


Figure 5.1: Secondary structure of the 10-23 DNAzyme. (A) Schematic representation of the 10-23 DNAzyme (black) in complex with its RNA substrate (red). The self-complementary sequence within the catalytic loop that led to the formation of a biologically irrelevant duplex formation (Nowakowski *et al.*, 1999b) is highlighted in green. The mutations C3A and A5C lead to a partialy active DNAzyme (Zaborowska *et al.*, 2002). (B) 2'-fluoro modified RNA fragment.

10-23 DNAzyme in which the self-complementary sequence is destroyed by the introduction of the point mutations C3A and A5C in order to suppress the formation of the biological irrelevant duplex conformation.

5.4 Material and Methods

5.4.1 Oligonucleotides used in this study

The DNAzyme and RNA sequences for crystallization experiments with the ternary complex were designed according to Rosenbach *et al.*. DNA and RNA oligonucleotides (Table 5.1) were acquired from biomers.net GmbH (Ulm, Germany) and BioSpring (Frankfurt a. Main, Germany). Fluorescent-labeled RNA substrates for cleavage assays are labeled wit a 6-Carboxyfluorescein (6-FAM) fluorophor on the 5'-end. The RNA substrates used for crystallography are stabilized by substituting the 2'OH group of the guanosine at the cleavage site by a fluorine atom. The oligonucleotides were dissolved in water at a concentration of 2 mM and stored at -20 °C.

Table 5.1: Oligonucleotide sequences used in this study.

Name	Sequence 5′–3′
Co-crystallization	
T839_U1A_1	ACA UGC ACC GUU ACC UCA CCU GCA UUG CAC UCC GCA G
T839_U1A_2	ACA UGC ACC GUU ACC UCA CAU CCA UUG CAC UCC GGA U
Dz839_U1A_C3A	GTG AGG TAA GGA TAG CTA CAA CGA GGT GCA TGT
Dz839_U1A_A5C	GTG AGG TAA GGC TCG CTA CAA CGA GGT GCA TGT
Helix-engineering	
6-FAM-T839_9-10	6FAM-AAA ACA UGC ACC GUU ACC UCA CC
T839_9-10	ACA UGC ACC GUU ACC UCA CC
Dz839_A5C_8-8	TGG GGT AAG GCT CGC TAC AAC GAG GTG CAT G
Dz839_A5C_8-9	TGG GGT AAG GCT CGC TAC AAC GAG GTG CAT GT
Dz839_A5C_9-8	GTG GGG TAA GGC TCG CTA CAA CGA GGT GCA TG
Dz839_A5C_9-9	GTG GGG TAA GGC TCG CTA CAA CGA GGT GCA TGT
Dz839_A5C_8-10	TGG GGT AAG GCT CGC TAC AAC GAG GTG CAT GTT
Dz839_A5C_10-8	GGT GGG GTA AGG CTC GCT ACA ACG AGG TGC ATG
Dz839_A5C_9-10	GTG GGG TAA GGC TCG CTA CAA CGA GGT GCA TGT T
Dz839_A5C_10-9	GGT GGG GTA AGG CTC GCT ACA ACG AGG TGC ATG T
Dz839_A5C_10-10	GGT GGG GTA AGG CTC GCT ACA ACG AGG TGC ATG TT
Deletion approach	
6-FAM-T839	6FAM-AAA ACA UGC ACC GUU ACC CCA A
T839	ACA UGC ACC GUU ACC CCA A
Dz839_ΔAGCT	TTG GGG TAA GGC TAC AAC GAG GTG CAT GT
Dz839_ΔT	TTG GGG TAA GGC TAG CAC AAC GAG GTG CAT GT
Tetrahedral DNAzymes	
6-FAM-T839	6FAM-AAA ACA UGC ACC GUU ACC CCA A
T839	ACA UGC ACC GUU ACC CCA A
sDz839_A	GTG TCG TTT TTG GGG TAA GGC TAG CTA CAA CGA GGT GCA TGT ACT TCA GAC TTA GGA ATG TGC TTC CCA CG
sDz839_B	GCT CTT CAC TTG GGG TAA GGC TAG CTA CAA CGA GGT GCA TGT AAA CGA CAC TAC GTG GGA ATC TAC TAT GG
sDz839_C	GTC TGA AAC TTG GGG TAA GGC TAG CTA CAA CGA GGT GCA TGT GAG AAG AGC CGC CAT AGT AAC ACA TTC CT

5.4.2 Self-assembly of tetrahedral DNAzymes

Tetrahedral DNAzymes (TDz) were prepared by self-assembly of three single stranded DNA molecules sDz839A, sDz839B, and sDz839C (Table 5.1). The single-stranded DNAzymes were incubated with a concentration of 600 nM each in 10 mM Tris/HCl pH 8.0, 20 mM $MgCl_2$, and 150 mM NaCl at 95 °C for 10 min and then slowly cooled down to 4 °C over 8 h. The assembled DNAzymes were analyzed on 10% native polyacrylamide gel electrophoresis (PAGE) and via analytical ultracentrifugation (AUC) after incubation of 3 h at 37 °C to test their stability.

5.4.3 Native PAGE

To analyze the self-assembly of TDz the samples were mixed with DNA Loading Dye (6x) (Thermo Fisher Scientific, Schwerte, Germany) and 10 µM of the complex were then analyzed on a 10% native Polyacrylamide gel for 3 h at 180 V. For visualization, the gel was incubated in Tris-borate EDTA (TBE) buffer containing a 1:10,000 dilution of the fluorescent nucleic acid dye GelRed (Merck, Darmstadt, Germany) for 1 h. Images were acquired using the ChemiDoc MP System (Bio-Rad, Hercules, CA, USA).

5.4.4 Analytical ultracentrifugation (AUC)

Sedimentation velocity (SV) analysis was performed to determine the size and shape-distribution of self-assembled TDz. Experiments were carried out in an analytical ultracentrifuge (Proteome Lab XL-A, Beckman Coulter, Brea, US) at 50,000 rpm, 20 °C and with a radial resolution of 0.002 cm in standard double sector cells made out of aluminium with an optical path length of 12 mm (Beckman Coulter, Brea, CA, USA). An An-60Ti (4-hole) rotor (Beckman Coulter, Brea, CA, USA) was used. The SV data were analyzed applying a continuous distribution Lamm equation model, $c(s)$ as implemented in the analysis software Sedfit (version 16p35) (Schuck, 2000). The wavelength for absorbance detection was set to 270 nm.

5.4.5 Cleavage assays

Cleavage assays with ternary DNAzyme:RNA:U1A complexes were performed with 10 µM unlabeled RNA substrate and 10 µM DNAzyme in 50 mM tris(hydroxymethyl)aminomethane (Tris)/HCl pH 7.5 in the presence of 10 mM Mg^{2+} for 3 h at 37 °C. Cleavage assays with the binary DNAzyme:RNA complexes were performed with 0.4 µM fluorescein-labeled RNA substrate and 0.4 µM or 0.04 µM DNAyzme in 50 mM Tris/HCl pH 7.5 in the presence of 10 mM Mg^{2+} or 100 mM Ca^{2+} for 3 h at 37 °C. Prior to cleavage the RNA and DNAzymes were denatured in buffer in the absence of Mg^{2+} or Ca^{2+} at 73 °C for 5 min, followed by an incubation for 10 min at room temperature. Subsequently, the reaction was started by adding Mg^{2+} or Ca^{2+}. For cleavage assays in the presence of the U1A protein, 10 µM of the U1A variant F56W were incubated for 30 min at room temperature to allow for formation of the ternary complex before starting the reaction with Mg^{2+}. Cleavage assays with tetrahedral DNAzymes (TDz) were performed with 0.04 µM TDz839 and 0.4 µM RNA substrate. Without prior denaturation at 73 °C. The RNA substrates were labeled with a fluorescein-label at the 5'-end to allow fluorescence-based visualization of the cleavage product. Cleavage reactions were terminated by adding 95% formamide and 25 mM EDTA. Samples were heated to 96 °C for 10 min and then cooled down on ice.

5.4.6 Denaturing PAGE

Separation of the samples from cleavage assays or dissolved crystals was carried out on 18% polyacrylamide gels with 7 M urea buffered with Tris-borate EDTA buffer (TBE) for 1.5 h at 20 W. For visualization of unlabeled nucleic acids the gel was incubated in TBE buffer containing a 1:10,000 dilution of the fluorescent nucleic acid dye GelRed (Merck, Darmstadt, Germany) for 1 h. Visualization of fluorescein-labeled RNA substrates was carried out by fluorescence detection. Images were acquired using the ChemiDoc MP System (Bio-Rad, Hercules, CA, USA).

5.4.7 Purification and isolation of the crystallization helper U1A

The crystallization helper was purified and isolated as described (Rosenbach *et al.*, 2020). Before being used in crystallization experiments, the protein was transferred into 25 mM Tris/HCl buffer pH 7.0 and the protein concentration was adjusted to 2 mM. The protein was stored at -80 °C.

5.4.8 Sample preparation for the crystallization of the ternary complex

Prior to crystallization, RNA substrate and DNAzyme were mixed in an equimolar ratio in 25 mM Tris/HCl pH 7.0 and denatured for 5 min at 73 °C, followed by an incubation for 10 min at room temperature. Then, $MgCl_2$ and KCl were added to final concentrations of 10 mM and 50 mM, respectively. DNAzyme, RNA and protein variant were sequentially incubated at room temperature for 30 min in a final concentration of 0.5 mM ternary complex in 25 mM Tris/HCl pH 7.0, 10 mM $MgCl_2$ and 50 mM KCl.

5.4.9 Sample preparation for the helix-engineering approach

Prior to crystallization, RNA substrate and DNAzyme were mixed in an equimolar ratio in 10 mM sodium cacodylate pH 7.0 and denatured for 5 min at 73 °C, followed by an incubation for 10 min at room temperature. Then, $MgCl_2$ and spermine were added to final concentrations of 10 mM and 2 mM, respectively.

5.4.10 Crystallization experiments with the ternary complex

In order to screen for a plethora of different parameters on the crystallization behavior of the ternary complexes, including type and concentration of buffer, precipitants, mono- and divalent metal ions, and other additives we used sparse matrix crystallization screens for initial screening. We used JBScreens Classic 1-8, which are commercially available from Jena Bioscience, Jena, Germany. Crystallization experiments were performed using the sitting-drop vapor diffusion method at room temperature. A sample volume of 0.3 µl with a concentration of 0.5 mM DNAzyme:RNA:protein complex in 25 mM Tris pH 7.0, 10 mM $MgCl_2$ and 50 mM KCl was mixed with an equal volume of resevoir solution. The crystallization plates were sealed with tape and

incubated for four weeks at room temperature. Tiny crystals of the ternary complex consisting of Dz839_U1A_A5C:T839_U1A_1:R70W were obtained from 10% PEG 3350, 100 mM $CaCl_2$, and 13% Glycerol; and 20 mM MES pH 5.8, 27% MPD, 400 mM NaCl, and 120 mM $CaCl_2$; and of Dz839_U1A_A5C:T839_U1A_1:F56W from 100 mM Tris/HCl pH 7.4, 15% PEG 400, 80 mM $CaCl_2$, 20 mM NaCl; and 100 mM HEPES pH 7.5, 30% PEG 400, 200 mM $CaCl_2$ as listed in Table 5.2. To obtain larger crystals, fine screening was performed based on the results from the inital screens with a sample volume of 1 µl beeing mixed with an equal volume of reservoir solution. After five weeks at room temperature, crystals were obtained. After soaking for 1 min in cryoprotecting solution (50% glycerol) the crystals were mounted on loops, flash-frozen, and stored in liquid nitrogen. Photos of the crystals were acquired using the Axiocam 105 color (Zeiss, Oberkochen, Germany).

Table 5.2: Crystallization hits for the ternary complex consisting of DNAzyme, RNA substrate and U1A protein.

Complex components	Crystallization conditions
Dz839_A5C_U1A:T839_U1A_1:F56W	100 mM Tris/HCl pH 7.4, 15% PEG 400, 80 mM $CaCl_2$, 20 mM NaCl
Dz839_A5C_U1A:T839_U1A_1:F56W	100 mM HEPES pH 7.5, 30% PEG 400, 200 mM $CaCl_2$
Dz839_A5C_U1A:T839_U1A_1:R70W	100 mM MES pH 6.0, 10% PEG 3350, 100 mM $CaCl_2$, 13% Glycerol
Dz839_A5C_U1A:T839_U1A_1:R70W	20 mM MES pH 5.8, 27% MPD, 400 mM NaCl, 120 mM $CaCl_2$

5.4.11 Crystallization experiments with binary and tetrahedral DNAzyme:RNA complexes

For the crystallization experiments with binary and tetrahedral DNAzyme(TDz):RNA complexes we used screening conditions that were designed by Nowakowski *et al.* for their combinatorial screen of paired oligonucleotides. Crystallization experiments were performed using the sitting-drop vapor diffusion method at room temperature. A sample volume of 0.5 µl with a concentration of 0.5 mM complex in 10 mM sodium cacodylate buffer pH 7.0, 10 mM $MgCl_2$, and 2 mM spermine was mixed with an equal volume of reservoir solution. The crystallization plates were sealed with tape and incubated for four weeks at room temperature. Crystals were obtained after four to seven days as listed in Table 5.3. Photos of the crystals were acquired using the Axiocam 105 color (Zeiss, Oberkochen, Germany). Since all tested cryoprotecting solutions caused visible damage to the crystals, the crystals were mounted on loops without previous soaking into cryprotectant, flash-frozen, and stored in liquid nitrogen.

Table 5.3: Crystallization hits for the binary and tetrahedral DNAzyme:RNA complexes.

Complex components	Crystallization conditions
Dz839_A5C_8-10:T839_9-10	50 mM sodium cacodylate buffer pH 6.0, 18 mM $MgCl_2$, 2.25 mM spermine, 1 mM $CuSO_4$, 9% 2-propanol
Dz839_A5C_8-10:T839_9-10	50 mM sodium cacodylate buffer pH 6.5, 18 mM $MgCl_2$, 0.9 mM spermine, 1.8 mM cobalt (III) hexamine, 9% 2-propanol
Dz839_ΔT:T839	50 mM sodium cacodylate buffer pH 6.0, 18 mM $MgCl_2$, 2.25 mM spermine, 1 mM $CuSO_4$, 9% 2-propanol
Dz839_ΔAGCT:T839	50 mM HEPES pH 7.0, 80 mM $MgCl_2$, 2.5 mM spermine
TDz839:T839	50 mM sodium cacodylate pH 6.0, 20 mM $MgCl_2$, 2.4 mM spermine, 5% PEG 4000

5.4.12 Test for X-ray diffraction

Crystals were tested for X-ray diffraction at beamline P11 at PETRA III of the Deutsches Elektronensynchrotron (German Electron Synchrotron, DESY) in Hamburg (Burkhardt *et al.*, 2016; Meents *et al.*, 2013), Germany at 100 K using radiation with an energy of 12 keV.

5.5 Results

5.5.1 Co-crystallization of the 10-23 DNAzyme in complex with its RNA substrate and the RNA-binding protein U1A

Our first approach involves the co-crystallization of the 10-23 DNAzyme variant Dz839, which specifically binds and cleaves the mRNA of the human prion protein (PrP) (Victor *et al.*, 2018), in complex with its stabilized RNA substrate and the RNA-binding protein U1A (Figure 5.2A). Besides a U1A variant that has been reported to have superior crystallization properties compared to the wildtype sequence due to the two mutations Y31H and Q36R (Oubridge *et al.*, 1994), we used two novel variants of the U1A protein, i. e. U1AF56W and U1AR70W that both contain a tryptophan residue and were reported to expand the toolbox for nucleic acid crystallization (Rosenbach *et al.*, 2020). In order to prevent the formation of catalytically irrelevant duplexes, we introduced point mutations at the positions C3 and A5 of the catalytic core region to obtain new variants of the 10-23 DNAzyme Dz839, namely Dz839_U1A_C3A and Dz839_U1A_A5C. To enable the RNA substrate T839 to simultaneously bind to the DNAzyme and the U1A protein, the RNA sequence has been elongated by the addition of the U1A specific hairpin loop. Since flexible regions in macromolecules may hinder or support crystallization of the sample, we designed two RNA substrates: For one substrate (T839_U1A_1) the RNA sequence has been designed in the most rigid conformation with the nucleotide at the very 5'-end

of the DNAzyme stacking onto the last nucleotide at the 3'-end of the elongated RNA substrate. Within the other substrate (T839_U1A_2) the nucleotide at the 5'-end of the DNAzyme is not able to stack onto the last nucleotide at the 3'-end of the RNA substrate, resulting in a more flexibel conformation. All sequences for both the DNAzyme variats as well as the RNA substrates are listed in Table 5.1. Prior to crystallization, the DNAzyme Dz839_U1A_A5C was tested for its cleavage activity in the presence of the RNA-binding protein U1AF56W. It was shown that the presence of the protein does not inhibit the cleavage of the RNA substrate T839_U1A_2 (Figure 5.2B). Nuclear magnetic resonance (NMR) experiments with the DNAzyme:RNA complex in the presence and absence of the U1A protein further demonstrate that binding of the protein has no influence on the structure of the catalytic loop region of the DNAzyme (Rosenbach *et al.*, 2020). The DNAzyme variants, RNA substrate variants and protein variants were combinatorially mixed and crystallization experiments were performed with each ternary complex.

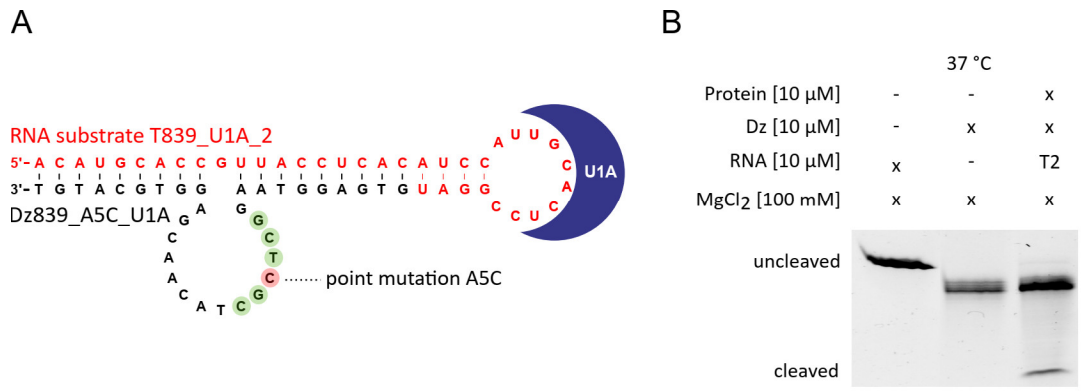


Figure 5.2: The RNA-binding protein U1A does not affect the DNAzyme activity. (A) Schematic representation of the ternary complex consisting of the 10-23 DNAzyme variant Dz839_U1A_A5C (black) in complex with the rigid RNA substrate T839_U1A_2 (red) and the RNA-binding protein U1A (blue). The point mutation at position five that was used to break down the palindromic sequence (green) is indicated in red. (B) Cleavage assay of the DNAzyme Dz839_U1A_A5C in the presence of the U1A protein variant F56W. PAGE analysis of the reaction mixtures containing DNAzyme Dz839_U1A_A5C (Dz), RNA substrate T839_U1A_2 (T2) and the DNAzyme:RNA:protein complex shows that the presence of the protein does not hinder the RNA cleavage. The samples were visualized using the nucleic acid stain GelRed. Schematic representation of the ternary complex with the rigid RNA substrate T839_U1A_1 and the corresponding cleavage assay were previously published by Rosenbach *et al.*.

Although the crystals obtained from fine screenings were single crystals with sharp edges, they did only poorly diffract. The best crystal (Figure 5.3A) was obtained from Dz839_A5C_U1A in complex with T839_U1A_1 and the U1A variant F56W. The crystal was grown in 100 mM HEPES pH 7.5, 30% PEG 400, and 200 mM $CaCl_2$ at room temperature. This crystal showed diffraction with a resolution of about 15 Å (Figure 5.3B), which is way too low for structure determination. In another study, this complex was analyzed using small angle X-ray scattering

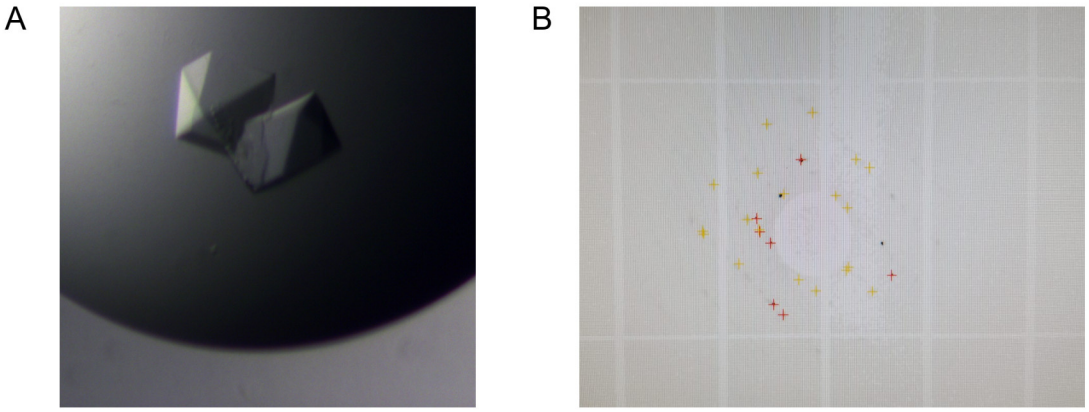


Figure 5.3: Crystals of the ternary complex led to waek diffraction patterns. (A) Single crystal of the Dz839_A5C_U1A:T839_U1A_1:F56W complex grown in 100 mM HEPES pH 7.5, 30% PEG 400, and 200 mM $CaCl_2$ at room temperature. (B) The diffraction pattern of the crystal shown in (A) only shows very weak diffraction up a maximum resolution of 15 Å.

(SAXS). The reconstitution of a low-resolution model showed that the ternary complex provides an elongated shape and has a length of about 11 nm, which is consistent with the number of baise pairs in the binding arms and the protein (Rosenbach *et al.*, 2020).

5.5.2 Crystallization experiments with DNAzyme:RNA complexes with varying helix length

In addition to co-crystallization of the DNAzyme:RNA complex with a crystallization helper, we also tested the capability of the binary DNAzyme:RNA complex to form highly-ordered crystals. For this purpose we used an approach that was previously successfully used for the crystallization of protein:DNA complexes, ribozymes, and small oligonucleotides (Anderson *et al.*, 1996; Jordan *et al.*, 1985; Scott *et al.*, 1995) and that will be referred to as the helix-engineering approach. This approach has already been applied to the wildtype 10-23 DNAzyme (hosting the self-complementary sequence within the catalytic core) and involves a combinatorial matrix consisting of nine different DNAzymes and nine different RNA substrates (Nowakowski *et al.*, 1999a) that differ from each other in length. The study by Nowakowski *et al.* resulted in well-formed single crystals obtained from 25 out of 81 different DNAyzme:RNA combinations. The best crystal was diffracting up to 2.8 Å. However, diffraction data did not lead to the structure of the 10-23 DNAzyme and its RNA substrate in a native conformation. Based on the results obtained by Nowakowski *et al.*, we designed an RNA substrate of the 10-23 DNAzyme variant Dz839_A5C that had a total length of 20 nucleotides and is termed T839_9-10. In the study by Nowakowski *et al.*, for the RNA substrate of this length crystals have been obtained in combination with any DNAzyme variant. Even after an incubation time of four weeks, we were not able to reproduce the high yield of crystals obtained by Nowakowski *et al.*. Screening only resulted in crystals of the RNA substrate T839_9-10 in complex with the DNAzyme

variant Dz839_A5C_8-10 that appeared after an incubation time of 7 days in 50 mM sodium cacodylate buffer pH 6.0, 18 mM $MgCl_2$, 2.25 mM spermine, 1 mM $CuSO_4$, and 9% 2-propanol as well as 50 mM sodium cacodylate buffer pH 6.5, 18 mM $MgCl_2$, 0.9 mM spermine, 1.8 mM cobalt (III) hexamine, and 9% 2-propanol. The binary complex consisting of Dz839_A5C_8-10 and T839_9-10 exhibits overhangs of one nucleotide each on both binding arms as demonstrated in Figure 5.4A. Activity assays performed with the unstabilized RNA substrate verified the capability of the DNAzyme Dz839_A5C_8-10 to cleave the RNA substrate T839_9-10 (Figure 5.4B).

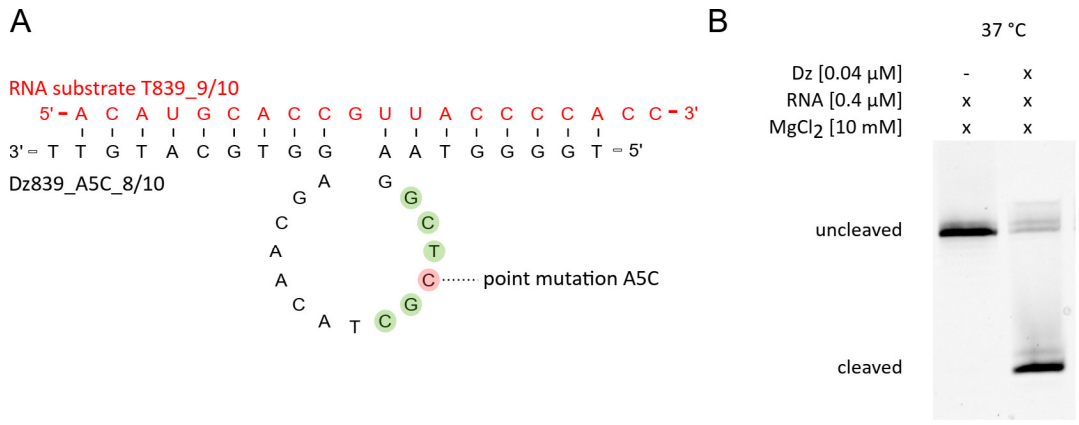


Figure 5.4: Crystallization attempts using a binary DNAzyme:RNA complex. (A) Schematic representation of the 10-23 DNAzyme variant Dz839_A5C_8-10 (black) in complex with the RNA substrate T839_9-10 (red). The point mutation at position five that was used to break down the palindromic sequence (green) is indicated in red. (B) Cleavage assay of the DNAzyme Dz839_A5C_8-10 and its fluorescent-labeled RNA substrate T839_9-10 in the presence of 10 mM $MgCl_2$ for 3 h at 37°C. The samples were visualized using fluorecence detection.

To explore the composition of the crystals of the binary complex, some crystals (Figure 5.5A) grown in the conditions listed in Table 5.3 were dissolved in RNA loading buffer and analyzed on a denaturing PAGE. Comparing the resulting bands with a standard of 20 pmol of the DNAzyme and RNA component, respectively, strongly supports the assumption that the crystals consist of intact DNAzyme and RNA sequences (Figure 5.5B). However, until now we were not able to solve the structure of the complex due to weak diffraction patterns.

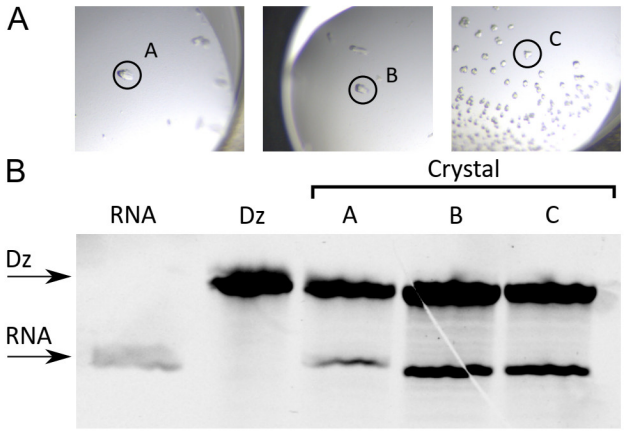


Figure 5.5: Analysis of crystals from binary DNA:RNA complexes. (A) Crystals obtained from the binary complex consisting of the DNAzyme Dz839_A5C_8-10 and its RNA substrate T839_9-10 in 50 mM sodium cacodylate buffer pH 6.0, 18 mM $MgCl_2$, 2.25 mM spermine, 1 mM $CuSO_4$, 9% 2-propanol (crystals A and B) and 50 mM sodium cacodylate buffer pH 6.5, 18 mM $MgCl_2$, 0.9 mM spermine, 1.8 mM cobalt (III) hexamine, 9% 2-propanol (crystal C). (B) Analysis of the crystal composition by denaturing PAGE (18%). As a standard 20 pmol DNAzyme and RNA substrate were used. The samples were visualized using the nucleic acid stain GelRed.

5.5.3 Crystallization experiments with 10-23 DNAzyme with reduced flexibility due to nucleotide deletion within the catalytic core

Flexible regions can tremendously hamper the crystallization of macromulecules. With 15 nucleotides, the 10-23 DNAzyme provides a large catalytic core, which is supposed to be highly flexible. In contrast to the 8-17 DNAzyme, the 10-23 DNAzyme does not provide any intramolecular basepairing within the catalytic core. Stabilizing the catalytic core or even minimizing the conserved region to obtain a more rigid conformation might be a great benefit for structural studies. In 2000, Okumoto & Sugimoto reported on a mutant DNAzyme that was derived from the 10-23 DNAzyme but lacks the nucleotides five to eight within the catalytic core, therefore providing a flexible region with eleven instead of 15 nucleotides. This new DNAzyme variant showed activity in the presence of Ca^{2+}, but was nearly inactive in the presence of Mg^{2+}. Five years later, in 2005, Zaborowska *et al.* performed deletion experiments where they stepwise deleted nucleotides within the catalytic core. It was shown that deletion of the thymidine at position eight did not have a severe effect on the catalytic activity of the 10-23 DNAzyme. To broaden our repertoire of biological samples for the structure elucidation of the 10-23 DNAzyme we performed crystallization experiments with both mutants. Based on the nucleotides that have been deleted from the catalytic core region, these DNAzymes will be referred to as Dz839_ΔT and Dz839_ΔAGCT (Figure 5.6A and B). Prior to crystallization experiments, both DNAzymes were tested for their cleavage activity. While the DNAzyme Dz839_ΔT was tested for RNA-cleaving activity in the presence of Mg^{2+}, the shorter variant Dz839_ΔAGCT was tested in the presence of Ca^{2+} (Figure 5.6C).

Crystallization experiments were performed under conditions described by Nowakowski *et al.* and that already have been used for crystallization experiments on the binary complexes with

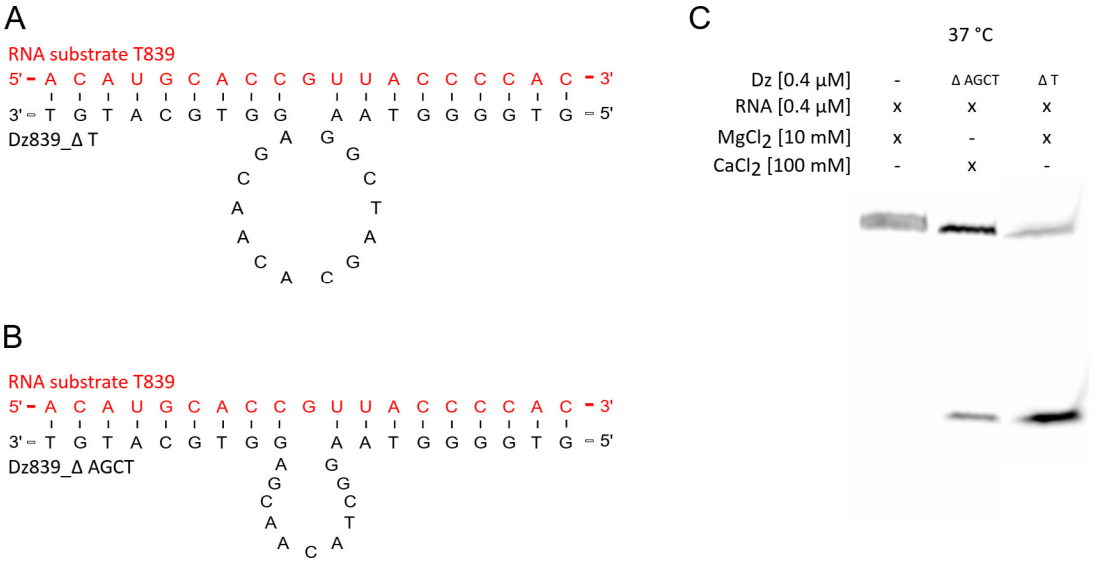


Figure 5.6: DNAzymes with reduced flexibility show RNA-cleavage activity in the presence of divalent metal ions. Schematic representation of the reduced DNAzyme variants Dz839_ΔT (A) and Dz839_ΔAGCT (B) with the DNAzyme in black and the RNA substrate in red. (C) Cleavage assay of the DNAzyme Dz839_ΔT and Dz839_ΔAGCT and their fluorescent-labeled RNA substrate T839_9-10 in the presence of 10 mM $MgCl_2$ and 100 mM $CaCl_2$ for 3 h at 37 °C. The samples were visualized using fluorescence detection.

varying helix length. We obtained very fragile crystals of the Dz839_ΔT:RNA complex from 50 mM sodium cacodylate buffer pH 6.0, 18 mM $MgCl_2$, 2.25 mM spermine, 1 mM $CuSO_4$, and 9% 2-propanol and of the Dz839_ΔAGCT:RNA from 50 mM HEPES pH 7.0, 80 mM $MgCl_2$, and 2.5 mM spermine (Figure 5.7A and B). However, so far, it was not possible to reproduce these crystals.

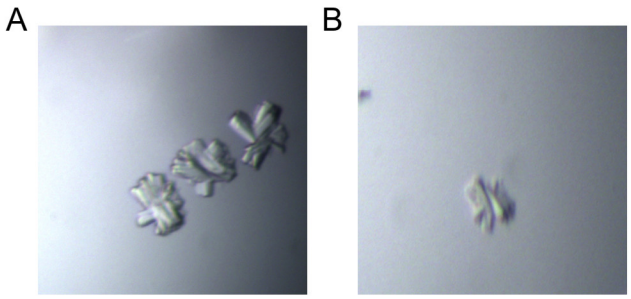


Figure 5.7: Crystallization experiments with DNAzymes with reduced flexibility led to crystals, which could not be reproduced. (A) Crystals obtained from crystallization experiments with the binary complex Dz839_ΔT:RNA grown in 50 mM sodium cacodylate buffer pH 6.0, 18 mM $MgCl_2$, 2.25 mM spermine, 1 mM $CuSO_4$, and 9% 2-propanol at room temperature. (B) Crystals obtained from crystallization experiments with the binary complex Dz839_ΔAGCT:RNA in 50 mM HEPES pH 7.0, 80 mM $MgCl_2$, and 2.5 mM spermine at room temperature.

5.5.4 Crystallization experiments with tetrahedral nanostructures built from 10-23 DNAzymes with their RNA substrates

In 2018, Thai *et al.* reported on the design of tetrahedral nanostructures with enhanced intracellular gene-silencing activity. These tetrahedral DNAzymes (TDz) were derived from three single-stranded 8-17 DNAzyme variants during self-assembly. The TDz-based nanostructures hold promise for enhanced cellular uptake and higher stability in biological fluids, since DNA tetrahedrons have already been successfully employed as nanocarriers for the delivery of small molecule drugs (Goodman *et al.*, 2005; Kim *et al.*, 2013b). The 8-17 DNAzyme-based tetrahedral nanostructures by Thai *et al.* were designed for intracellular cleavage of the mRNA of the enhanced green fluorescent protein (eGFP). The TDz were tested for their *in vitro* activity. Here, a very low cleavage activity could be observed in the presence of 0.5 mM Mg^{2+}. However, the duration of the cleavage assays is not mentioned. Furthermore, the TDz by Thai *et al.* were tested for their cellular uptake properties. For this purpose NIH3T3 cells were treated with fluorescence-labeled TDz in the absence of transfection agents and the uptake efficiency was then measured using flow cytometry. Here, Thai *et al.* found that the uptake of TDz is much more efficient than the uptake of their single-stranded counterparts. This observation was traced back to the enhanced serum stability of TDz. Thai *et al.* also reported on an enhanced *in vivo* gene-silencing activity. However, control experiments to exclude that the observed downreagulation of the mRNA and protein levels is solely due to antisense effects are missing. Nevertheless, the tetrahedral nanostructures provide a defined three-dimensional structure, which makes the TDz promising canditates for crystallization experiments. We designed TDz that contain the catalytic motif of the 10-23 DNAzyme (TDz839, Figure 5.8A) and thus, differ from the TDz reported by Thai *et al.*, which provide the catalytic motif of the 8-17 DNAzyme. However, the sequences, which are required for the self-assembly process are adapted from the previousely reported TDz (Thai *et al.*, 2018). Self-assembly of the TDz was induced by incubating three single-stranded DNAzymes (sDzA, sDzB and sDzC), each of it capable of binding the RNA substrate T839, in an equimolar ration of 600 nM in 10 mM Tris/HCl pH 8.0, 20 mM $MgCl_2$, and 150 mM NaCl. The DNA sequences were denatured at 95°C for 10 min, before cooled down to 4°C over about 8 h. The success of the assembly process was varified using native PAGE analysis (Figure 5.8B, line 4) and analytical ultracentrifugation (AUC) (Figure 5.8C). The tetrahedral complex was then incubated with three equivalents of the stabilized RNA substrate T839 to allow specific binding of the RNA to the three pairs of binding arms within the DNA tetrahedron and the binding was monitored using a native PAGE (Figure 5.8B, line 5). Unstabilized RNA substrates were used to test the TDz for its RNA-cleavage activity in the presence of 1 mM $MgCl_2$ (Figure 5.8D).

We also tested the tetrahedral DNAzyme TDz839 for its ability to cleave the mRNA of the human PrP *in vivo* (data not shown). For this purpose, WAC2 cells were transfected with the TDz839 using Lipofectamine® 3000. After 5 days, cell lysates were prepared and analyzed for their PrP levels using Western blot. No differences in PrP levels could be observed between cells that were treated with the TDz839 and untreated cells.

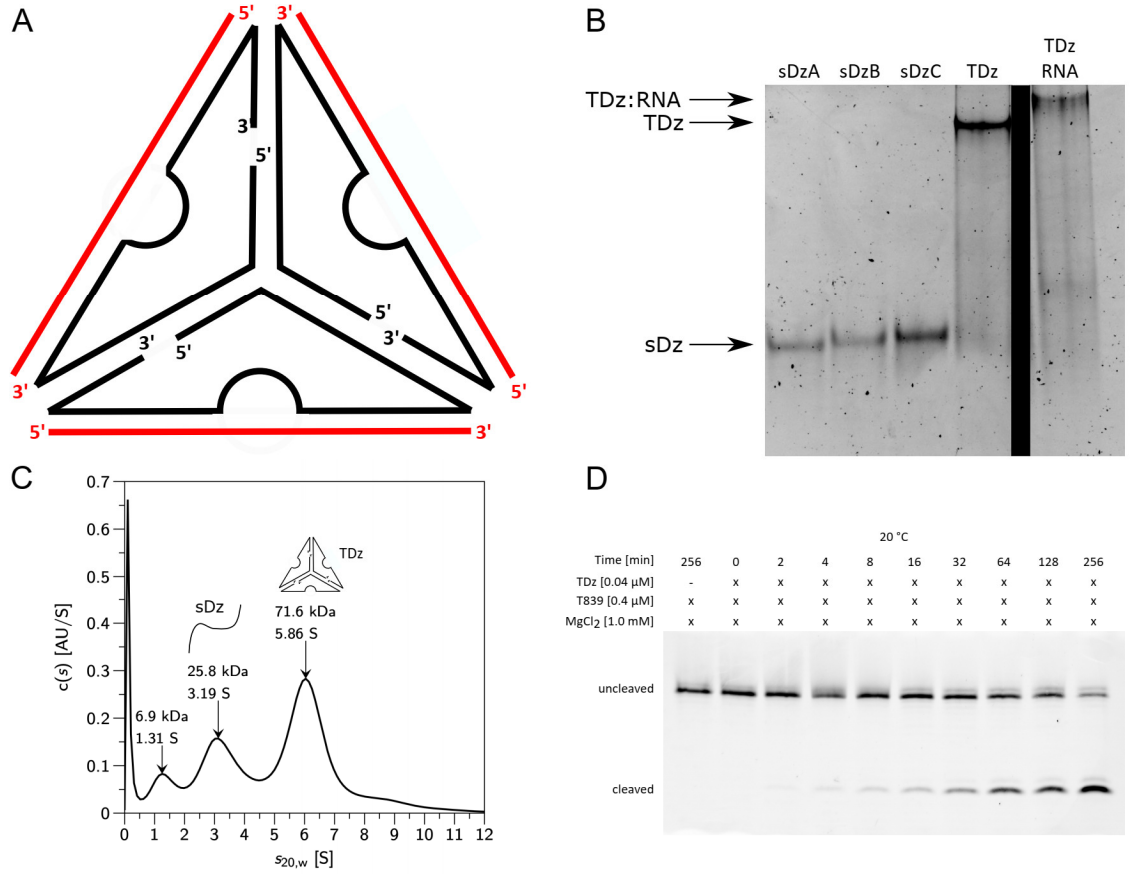


Figure 5.8: Tetrahedral nanostructures built from the 10-23 DNAzyme. (A) Schematic representation of the tetrahedral DNAyzme TDz839 (black) in complex with the RNA substrates T839 (red). (B) Native PAGE visualizes the assembly of the TDz from single-stranded DNA molecules (sDzA, sDzB, sDzC) as well as the binding of the RNA substrate T839 to the TDz (TDz:RNA). (C) Size distribution derived from AUC experiments. (D) Cleavage assay of the TDz839 with fluorescent-labeled RNA substrates T839 in the presence of 1 mM $MgCl_2$ at different time points after incubation at 20 °C.

Nevertheless, TDz839 in complex with the stabilized T839 RNA substrates was used for crystallization experiments using the screening conditions reported by Nowakowski *et al.*. So far, only one crystal was obtained after one week in 50 mM sodium cacodylate pH 6.0, 20 mM $MgCl_2$, 2.4 mM spermine, and 5% PEG 4000 at room temperature. However, this crystal subsequently resolved.

5.6 Discussion

Structure elucidation of nucleic acids is often impeded by their negatively charged and regularly ordered phosphate backbone. Although several crystal structures of catalytic RNAs have been solved, until today, only two crystal structures of DNA catalysts in a biologically relevant conformation exist. One crystal structure shows the post-catalytic state of the RNA-ligating DNAzyme 9DB1 (Ponce-Salvatierra *et al.*, 2016) and one crystal structure shows the RNA-

cleaving 8-17 DNAzyme in complex with a non-native DNA substrate (Liu *et al.*, 2017). Besides the detrimental surface properties of nucleic acids, the 10-23 DNAzyme provides an additional feature, which may be disadvantageous for crystallization experiments: The DNA catalyst comprises a highly flexible catalytic core region, which consists of 15 nucleotides. In this study, we present four different approaches for the structure elucidation of the 10-23 DNAzyme in complex with its RNA substrate using X-ray crystallography. All approaches aim at finding the optimal properties of the biological sample, rather than optimizing the screening conditions. The U1A protein has been successfully used for X-ray crystallography resulting in several crystal structures (Adams *et al.*, 2004; Cochrane *et al.*, 2007; Ferré-D'Amaré *et al.*, 1998). However, in 2020 we have been the first reporting on the use of the U1A protein for structural studies of a ternary complex comprising DNA, RNA and the protein itself. In the present study we used this approach for crystallization experiments, resulting in poorly diffracting crystals. One challenge that may arise with the co-crystallization of the protein with a binary nucleic acid complex is the homogeneity of the sample, since the three components need to be combined in exact equimolar ratios to provide a highly pure sample. One opportunity to circumvent this challenge may be the synthesis of a chimeric RNA:DNA strand that comprises both the DNAzyme moijety as well as the RNA subtrate. However, synthesis of a chimeric nucleic acid of such a length might be difficult and cost-intense. In analogy to the helix-engineering approach used for the binary complex, the length, as well as the sequence of the 3'-end of the DNAzyme or the 5'-end of the RNA could be modified, as well as the stacking interaction between the 5'-end of the DNAzyme and the 3'-end of the RNA hairpin.

The most promising results were obtained from the crystallization experiments with the binary complexes with varying helices based on the combinatorial screen by Nowakowski *et al.*. Analysis of the crystals using denaturing PAGE clearly indicates that the crystals consist of the DNAzyme and its RNA target. However, we were not yet able to solve the structure of the complex due to poor diffraction of the crystals indicating a poor long-range order within the crystal lattice. Introduction of the point mutation A5C on the one hand prevents the formation of a catalycally irrelevant duplex formation as it was reported by Nowakowski *et al.*, on the other hand, the omission of the self-complementary sequence reduces the amount of intramolecular interactions, probably leading to a higher flexibility of the catalytic core. However, due to the inital success that the obtained crystals contain the intact RNA substrate and DNAzyme, this approach should be followed up focussing on the optimization of the crystallization conditions, including optimization of the incubation temperature for the crystal plates, since nucleic acids favour higher temperatures between 28 °C and 37 °C for crystallization (Golden & Kundrot, 2003). The deletion of nucleotides from the catalytic core is one approach to tackle the high flexibility of the catalytic core region. We used 10-23 DNAzyme variants that were reported to lack one or four nucleotides within the flexible loop, but still showed activity in the presence of divalent metal ions (Okumoto & Sugimoto, 2000; Zaborowska *et al.*, 2005). Since the reduction of the catalytic core has so far not been leading to highly-ordered crystals of the DNAzyme:RNA complex, a combination of both, the helix-engineering approach and the deletion of nucleotides from the catalytic core might be a step into the right direction.

The TDz that were reported to be promising in terms of enhanced *in vivo* activity (Thai *et al.*, 2018), were used as a starting point for the design of novel tetrahedral DNAzymes based on the sequence motif of the 10-23 DNAzyme. A combined analysis of the self-assembly process via native PAGE and AUC provide evidence that the TDz are formed in high purity (Figure 5.8B and C), although AUC profiles point towards the presence of single-stranded DNA molecules within the sample. In contrast to the TDz designed by Thai *et al.* the 10-23 DNAzyme based tetrahedrons did not show any *in vivo* activity, although *in vitro* activity was comparably high. Since the TDz are stable for several hours at 37°C and have shown an enhaced stability in blood serum, it is unlikely that the lack of *in vivo* activity is due a disassembly of the tetrahedrons. Nevertheless, the intracellular milieu provides conditions that significatly differ from *in vitro* conditions. Therefore, aspects like high salt concentrations and molecular crowding effects that may impede the catalytic activity of the DNA catalysts should not be disregarded. The three-dimensional fold of the DNA tetrahedrons makes these DNAzyme variants attractive for crystallization experiments. Nevertheless, we did not test the influence of changes in salt concentration or pH that may have an impact on the tetrahedron stability after the assembly. Thus, we have no evidence for the stability of the TDz under the tested crystallization conditions. The analysis of the stability of the assembled TDzs provides a crucial test for the suitability of these nanostructures of crystallization experiments. Another approach that can be taken under consideration is the co-crystallization of TDz with the RNA-binding protein U1A to further improve the crystallization properties of the biological sample. Therefore, the TDz needs to be incubated with pre-assembled RNA:U1A complexes after the self-assembly of the tetrahedrons.

Taken together, the crystallization of the 10-23 DNAzyme remains a challenging task. One critical aspect is the design of the biological samples for crystallization experiments. Further approaches may involve the introduction of chemically modified nucleotides or the introduction of locked nucleic acids (LNA), since it was reported that LNA modifications enhance the stability of DNA:RNA complexes (Vester & Wengel, 2004). Solving the structure of the 10-23 DNAzyme will provide essential insights into the reaction mechanism, which is crucial for understanding DNAzyme-mediated catalysis. Structural and mechanistic information provide the basis for further optimization and developing DNAzyme applications.

6 General discussion on DNA-mediated catalysis

In this thesis, the RNA-cleaving 10-23 DNAzyme has been the center of attention. Several mutational studies and activity analyses have been reviewed with special regard to the importance of specific molecular features and functional groups as well as the role of different metal ions in the catalytic process that will enable the development of improved strategies to elucidate the three-dimensional structure of the 10-23 DNAzyme and to understand the mechanism by which the DNAzyme performs RNA-cleavage (Chapter 2). We discussed the experimental setups that have been used in numerous studies to explore the influence of nucleotide modifications on the catalytic properties of the DNAzyme. In a majority of these studies, DNAzyme and RNA substrate were used in non-catalytic ratios with the DNAzyme in 100- to 200-fold excess over the RNA substrate (Nawrot *et al.*, 2007; Räz & Hollenstein, 2015) or with enormous amounts of $MgCl_2$ (Wang *et al.*, 2010a) (see Table 2.3). To obtain comparable insights into structural and functional properties of the 10-23 DNAzyme, activity and mutagenesis studies should be performed under catalytic conditions, i. e. reasonable DNAzyme:RNA ratios and moderate concentrations of divalent cations to mimick physiological conditions. Nevertheless, the analyses of the mutational and substitutional studies allowed for the identification of some functional groups that play a key role in 10-23 DNAzyme-mediated catalysis (Section 2.5).

The investigation, understanding and prediction of M^{n+} binding to nucleic acids is a highly complex and difficult task. Two main problems exist: (i) all four nucleotides offer possible coordination sites, including the negatively charged phosphate backbone as well as atoms of the sugar and nucleobase moieties (Figure 2.7, Chapter 2). (ii) M^{n+} ions that are usually found to be associated with nucleic acids such as Na^+, K^+ or Mg^{2+}, and also Mn^{2+} are kinetically labile and show a mixture of inner- and outer-sphere contacts. As a consequence, a precise localization of the binding sites is very challenging. The analysis of kinetic parameters of the 10-23 DNAzyme in the presence of different sodium ion levels and different divalent metal ions revealed novel aspects that have previously not been under consideration: The lack of *in vivo* activity of the 10-23 DNAzyme has been previously traced back to the low abundance of free Mg^{2+} ions inside the cell (Victor *et al.*, 2018). In this study a simple two-state model was assumed in which the DNAzyme:RNA complex is either completely free of Mg^{2+} or fully bound to all three Mg^{2+}. Our new calorimetric data show that in in the absence as well as presence of NaCl a model assuming multiple Mg^{2+}-binding sites in an *n*:1 binding mode provides the best fit (Figure 3.1E and 3.1F). EPR measurements further support the observation that more than one divalent metal ion is interacting with the DNAzyme:RNA complex in the presence of all tested NaCl concentrations. In the absence of NaCl our EPR data is best fit to a simple model assuming a non-cooperative binding of six Mn^{2+} to the DNAzyme:RNA complex. Increasing NaCl concentrations induce a cooperative binding behavior with a Hill coefficient of > 1 (Figures 3.3B and 3.3C, S3.4). A Hill coefficient $n > 1$ indicates positive cooperativity, meaning that once one Mn^{2+} is bound to the DNAzyme:RNA complex, its affinity for another Mn^{2+} in-

creases. In comparison to the binding behavior in the absence of NaCl (Figure 3.3A) our data indicate that the presence of NaCl makes binding of the first Mn^{2+} ions less favorable. Possible explanations for the role of NaCl in this behavior include competition for the same binding site, shielding of electrostatic interactions and/or stabilization of different complex conformations.

That increasing NaCl concentrations induce a cooperative binding behavior of Mn^{2+} to catalytically active nucleic acids has been further reported for a DNA:RNA hybrid consisting of a DNA substrate and the hammerhead ribozyme (Horton *et al.*, 1998). In the abscence of NaCl, the obtained data was best fit to a single set of 16 Mn^{2+} sites with an averange K_D of 40 µM. In contrast to experiments that were performed at low monovalent cation concentrations, Mn^{2+} binding to the hammerhead:DNA hybrid in the presence of 1 M NaCl resulted in a sigmoidal binding isotherm, suggesting a cooperative binding behavior. In 1 M NaCl, a single Mn^{2+} binds at low concentrations of added Mn^{2+}. This is contrary to the the behavior observed in 100 mM NaCl, where a steep rise to three to four bound Mn^{2+} per hybrid was observed at similar concentrations of added Mn^{2+}. In 1 M NaCl, however, at higher concentrations of added Mn^{2+} a transition occurs resulting in the population of additional metal sites. Fitting the data obtained at low Mn^{2+} concentrations, an apparent K_D of $\leqslant$ 10 µM was obtained for the initially occupied, tight Mn^{2+} site in the presence of 1 M NaCl.

Using NMR spectroscopy and Mg^{2+} titration experiments, we were able to shed more light on the structural changes induced by mono- and divalent metal ions. For the interaction of Mg^{2+} with the binding arms of the DNAzyme, fast exchange interactions could be observed. Analysis of Mg^{2+}-induced CSPs of selected peaks allowed for the precise determination of a K_D value of 2.2 mM. Interaction of Mg^{2+} with the catalytic loop resulted in intermediate exchange rates that lead to the disappearance of the affected peaks. Here, we were not able to determine an accurate K_D value for this region. Nevertheless, our data suggest that the affinity of nucleotides in the catalytic loop could be about 10-fold higher for Mg^{2+} as compared to residues in the binding arms (Figure 3.4C). In addition, we compared the influence of the addition of 1 mM $MgCl_2$ or addition of 400 mM NaCl on the complex conformation and exchange dynamics. Both conditions induced comparable peak shifts (Figure 3.4), indicating that both conditions have similar structural effects. Nevertheless, the DNAzyme is inactive even in the presence of high Na^+ concentrations of 1 M (Figure S3.7), indicating that the intrinsic properties of Mg^{2+} ions are cruicial for DNAzyme activity. Our new results obtained by FRET-based cleavage assays as well as NMR spectroscopy provide evidence, that the sodium ions compete with Mg^{2+} ions for the same binding sites within the DNAzyme:RNA complex and that therefore high levels of monovalent metal ions inside the cell have an inhibitory effect on the DNAzyme activity. At monovalent cation concentrations of physiological relevance a decrease of the cleavage rate of approximately 40% was observed for each Mg^{2+} concentration (Figure 3.5C). For simplicity, we assumed a model in which one Mg^{2+} ion replaces one Na^+ ion ((3.4) and (3.5) in Chapter 3), although ITC measurements as well as NMR experiments speak for the binding of further Mg^{2+} ions. However, due to a too large range of parameters that can be assumed for this model, it was not possible to draw definite conclusions with regard to binding constants for the different cations (Figure S3.8 and Table 3.2). A completely opposite effect effect on the cleavage rate has

been observed for Mn^{2+}. Here, the cleavage rates increase with increasing NaCl concentrations at constant Mn^{2+} concentrations (Figure 3.6). We assumed a model in which Na^{+} induces a folding of the DNAzyme:RNA complex that is favorable for the binding of Mn^{2+} ((3.6) and (3.7) in Chapter 3). Combining our findings with the intrinsic properties of the different metal ions (see Table S3.2) it can be speculated that Na^{+} can rather adequately mimick the hydration shell contacts of divalent ions which are responsible to induce a cleavage-competent conformation of the pre-catalytic DNAzyme:RNA complex. This pre-catalytic state, however, requires at least one additional divalent cation for catalytic activity. Especially with regard to *in vivo* appplications or applications as biosensors of RNA-cleaving DNAzymes the observation made for the metal ion dependency of the DNAzyme:RNA complex are of great interest, since they do not only reveal another aspect that may hinder the *in vivo* activity of DNAzymes, but also point towards the complexity of DNA-mediated catalysis. The different behavior of the 10-23 DNAzyme with regard to catalytic rates and metal binding reveal that one has to carefully consider the use of the 10-23 DNAzyme as metal sensors, before the RNA-cleavage mechanism is fully understood.

Although our kinetic analyses already provide deep insights into the metal ion dependency of the 10-23 DNAzyme, in the absence of structural data, the mechanism by which the DNAzyme performs the cleavage reaction is still not fully understood. The same is true for the exact metal binding sites within the DNAzyme:RNA complex. We put great effort in screening for samples with enhanced properties for crystallization. The surface of nucleic acids is dominated by negatively charged and regularly ordered phosphate groups that often lead to crystals with a poor long-range order. Several strategies have been developed to overcome these difficulties, including the design of several matrix screens that have been exclusively designed for the crystallization of nucleic acids. Analyses of the crystallization conditions from 1450 nucleic acid crystal structures that are deposited in the Protein Data Bank (PDB) revealed that 2-methyl-2,4-pentanediol (MPD) as a precipitant and $MgCl_2$ and spermine as additives are the most successfully used components for the crystallization of nucleic acids (Viladoms & Parkinson, 2014). For our crystallization experiments we basically used the commercially available nucleic acid screen JBScreen Nuc-Pro (JenaBioscience, Jena, Germany) or a matrix screen that has been previously reported by Nowakowski *et al.* (1999a) that are both designed based on rational combination of additives and buffers that have successfully been used for nucleic acid crystallization. Nevertheless, the screen reported by Nowakowski *et al.* (1999a), basically used cacodylate buffer. While the advantage of cacodylate buffer is its extrem pH stability at different temperatures, the buffer contains arsenic. Its presence can result in drowning out a fluorescence emission signal during a fluorescence scan for selenium or bromine atoms, masking the peak and inflection points of the anomalous scatters (Stepanov *et al.*, 2011; Viladoms & Parkinson, 2014). Although that is not relevant for this study, since we do not use nucleic acid sequences with modified bases with anomalous scattering atoms, it should be kept in mind, since such modifications are ideal for *ab initio* methods for structure elucidation. Further matrix crystallization screens that have been used in this study include the JBScreens Classic 1-10 from JenaBioscience (Jena, Germany) as well as the PEG and $AmSO_4$ Suits from Qiagen (Hilden, Germany). These screens were primary tested for crystallization experiments that in-

volve the co-crystallization of nucleic acids in complex with an RNA-binding protein. Besides the crystallization conditions, the biological sample has a great impact on the success of crystallization experiments with nucleic acids. Therefore, we put great effort in optimizing the samples for crystallization experiments. One approach that previously has been successfully used for the crystallization of nucleic acids such as ribozymes (Ferré-D'Amaré *et al.*, 1998), riboswitsches (Cochrane *et al.*, 2007) or introns (Adams *et al.*, 2004) involves the co-crystallization of the large nucleic acid in complex with the RNA-binding domain of the human U1 small nuclear ribonucleoprotein A (U1A-RBD). This protein provides several properties that make it a suitable crystallization helper: (i) The U1A domain comprises residues 1-98 and adopts a globular shape with a compact fold, in presence or absence of RNA (Nagai *et al.*, 1990; Oubridge *et al.*, 1994). (ii) The crystal structures of U1A:RNA complexes reveal that contacts between RNA and protein comprises both polar and non-polar interactions, which are independent of ionic strength. Therefore, complex formation is not strongly affected by extreme crystallization conditions, such as solutions with very low or high salt concentrations (Ferré-D'Amaré, 2010). (iii) The protein U1A binds to the specific RNA loop with high affinity, the dissociation constant K_D of the complex was determined to be approximately 10^{-11} M (van Gelder *et al.*, 1993). Nevertheless, no tryptophan residues are present in the amino acid sequence of the protein U1A as it has been used as a crystallization module, which prohibits the distinction between salt and protein crystals by fluorescence techniques. Therefore, we designed and successfully isolated new variants of the U1A protein, A2W, H10W, F56W, R70W and K98W, with each of them exhibiting a tryptophan residue for fluorescence detection. Analysis of the crystals that where obtained from three of the variants in the absence of nucleic acids reveal that the variants F56W, R70W and K98W all crystallize in different space groups (see Figure 4.3, Chapter 4), while the overall structure and the RNA-binding properties are not affected. Hereby, the U1A variant F56W was in particular catching our eyes, since the arrangement of the molecules within the asymmetric unit was particulary interesting for soaking experiments with nucleic acids (Figure S4.9). While most U1A-RBD variants in animals lack tryptophan residues, the homologs from *Viridiplantae* contain a tryptophan residue at position 56 (Figure S4.1). This amino acid residue is involved in RNA-binding (Shiels *et al.*, 2002). However, since the nucleotide sequence of the U1 RNA hairpin sequence is highly conserved in both *Animalia* and *Viridiplantae*, the mutation F56W should not have an impact on the RNA-binding properties of the protein. In crystals obtained from the protein variant F56W, the distance between two molecules was determined to be 19 Å with the RNA-binding pockets facing towards the bulk solvent of the crystal and that no close crystal-packing contacts are present around the RNA-binding pocket. These properties of the preformed protein crystals allowed for soaking of the specific, 21 nucleotides comprising hairpin loop into the crystals resulting in crystals that diffracted up to 2.37 Å. The structure of this protein:RNA complex did not significantly differ from the structure that was obtained from co-crystallization experiments of the U1A variant F56W with the specific RNA hairpin (Figure 4.6). Furthermore, the presence of a tryptophan residue within the RNA-binding pocket allows for the monitoring of RNA-binding by fluorescence detection (Figure 4.4). In the absence of RNA, the protein crystal shows tryptophan fluorescence when exposed to UV light, while binding of

the specific RNA hairpin leads to a decrease in the fluorescence signal, resulting from stacking interactions between the RNA and the tryptophan residue. This allows for *in crystallo* detection of RNA-binding to the protein in the crystalline form, that has not previously been reported. Although soaking of nucleic acid sequences into preformed protein crystals is so far not a common technique and only a few crystal structures of nucleic acid:protein crystals are reported that have been obtained by soaking (Horn *et al.*, 2004; Valegård *et al.*, 1994; Wojtas *et al.*, 2012), this method has several advantages over the co-crystallization of nucleic acid:protein complexes: (i) Since especially RNA is a very fragile macromolecule that can be easily degraded by RNAses, soaking provides a method that preserves the RNA. The incubation time of the RNA or DNA within the crystallization solution is drastically reduced. (ii) Furthermore, the amount of RNA that is needed for soaking experiments is significatly lower than for co-crystallization experiments, since one does not have to search for new crystallization conditions once protein crystals are obtained. That makes the soaking approach also more cost-effective. Nevertheless, soaking causes the risk of irreversibly disrupting the packing interactions of preformed protein crystals. In the worst case, this can lead to the complete dissolution of the crystal, a phenomenon that we observed several times, until the optimal soaking conditions, including RNA concentration and composition of the crystallization conditions for the protein have been evaluated. Nevertheless, we are convinced that our protocol is applicable for structure elucidation of RNA:protein or DNA:RNA:protein complexes that differ from the reported nucleic acids and protein variants. While the novel protein U1A F56W, that we successfully used for the soaking of the U1 RNA hairpin sequence into preformed protein crystals is the only variant derived from the sequence of the previously reported U1A crystallization helper by the introduction of tryptophan mutations that forms soakable crystals, the other U1A variants expand the toolbox for co-crystallization experiments. The U1A mutants A2W, H10W, R70W, and K98W are all capable of binding the specific RNA hairpin sequence (Figure S4.3). We used two of the novel U1A variants, F56W and R70W, for co-crystallization experiments with the 10-23 DNAzyme and a stabilized and elongated RNA substrate, after the formation of a homogeneous ternary complex has been varified by using SEC, native PAGE and NMR spectroscopy (Figures S4.4, 4.2A, and 4.2C). These attempts led to poorly diffracting crystals with a quality not high enough for structure elucidation (Chapter 5.5.1, Figure 5.3B). Ternary complexes with the U1A variants A2W, H10W and K98W have not been screened for their crystallization properties simply due to the high amount of nucleic acid samples that are needed for this purpose and that would have exceeded the financial framework of this thesis. The search for crystals of the ternary complex with a sufficient long-range order can therefore be expanded to at least complexes involving the U1A variants A2W and K98W. Analysis of crystals derived from the parental, tryptophan-lacking U1A protein holds evidence that the residue at position 10 is crucial for intermolecular crystal contacts and therefore mutation of this residue may hamper the crystallisation properties (Nagai *et al.*, 1990). An additional option that can be explored to further improve the crystallization properties of the ternary complex is the modification of the protein surface by methylation of the amino group of solvent-exposed lysine residues. This method has been routinely used as a rescue strategy in cases where initial crystallization screens are not successful (Walter

et al., 2006). Besides the attempts to crystallize a ternary complex consisting of the DNAzyme, the stabilized RNA substrate and the RNA-binding protein U1A, we performed several crystallization experiments in the absence of a crystallization module. These experiments were solely based on the variation of the sequence and length of the DNAzyme and the RNA substrate. For this purpose, we used an approach that has previousely been reported as the helix engineering-approach (Jordan *et al.*, 1985). This approach involves the variation of the length of the binding arms of the DNAzyme and of the RNA substrate. Here, we obtained crystals that, as verified by a denaturing PAGE (Figure 5.5), consist of the DNAzyme variant Dz839_A5C_8-10 and the RNA substrate T839_9-10 in the presence of cacodylate buffer, $MgCl_2$ and spermine as well as additional additives. However, so far we have not been able to solve the structure of this complex due to weak diffraction. We have also been able to produce crystals of the RNA substrate T839_9-10 and a DNAzyme variant that lacks either one or four nucleotides within the catalytic core. Previously, it was shown that the deletion of the nucleotide T at position 8 only slightly affects the activity of the DNAyzme, while the deletion of the nucleotides at position 5 to 8 changes the metal ion specificty of the DNAzyme from Mg^{2+} to Ca^{2+} (Okumoto & Sugimoto, 2000; Zaborowska *et al.*, 2005). However, we were not able to reproduce these initially obtained crystals for structure elucidation. As another strategy, we designed tetrahedral DNAzyme structures that consist of three DNAzymes that form a tetrahedron. Such tetrahedral DNAzyme complexes were previously reported to show an enhanced *in vivo* activity (Thai *et al.*, 2018). However, the tetrahedral DNAzymes were assembled from single-stranded DNAzymes derived from the 8-17 DNAzyme, while we used single-stranded 10-23 DNAzyme for the tetrahedron assembly. For these nanostructures, we were not able to detect any *in vivo* activity for gene silencing of the human Prion protein. Furthermore, we were not able to obtain reliable crystals from this sample (Chapter 5.5.4).

Besides the great effort that we put in attempts to crystallize the 10-23 DNAzyme in complex with its RNA substrate I was involved in a collaborative project that aims at solving the structure of the DNAzyme:RNA complex by solution NMR. Here, I assisted with kinetic measurements of the 10-23 DNAzyme Dz839_A5C to provide functional data that support the structural data. These results have not been included to this thesis, nevertheless they will shed a new light on the structural aspects that DNA-mediated RNA-cleavage is based on.

In this thesis we did not only reveal new aspects of metal ion dependency of the DNAzyme, but we also introduced new methods in the field of structural elucidation of DNA catalysts that will serve as a starting point for future experiments on this interesting and fascinating molecular tool. High-resolution structural information will allow for a better understanding of the reaction mechanism, which provides the basis for introducing chemical modifications to obtain an improved activity, especially at low Mg^{2+} concentrations. Such an optimization was for example reported by Fedoruk-Wyszomirska and co-workers (Fedoruk-Wyszomirska *et al.*, 2009) who introduced a structural element into minimalist hammerhead ribozymes that stabilizes the catalytic core in its active conformation and therefore maked the ribozyme more active at low Mg^{2+} concentrations. This modification was only possible due to a detailed understanding of the structural and mechanistic properties of the ribozyme.

7 Iron-sulfur clusters in pro- and eukaryotes

Iron-sulfur (Fe–S) clusters are ubiquitous cofactors present in all domains of life. They are capable of catalyzing diverse reactions and are associated with different enzymes that are involved in crucial cellular processes. Reactions catalyzed by proteins harboring Fe–S clusters range from simple one-electron transfer, over dinitrogen reduction to complex organic reactions promoted by radical *S*-adenosylmethionine (SAM) enzymes (Beinert *et al.*, 1997; Frey *et al.*, 2008; Hoffman *et al.*, 2013). Despite the wide structural range of Fe–S clusters within proteins, the biological synthesis of all Fe–S clusters starts with the assembly of [2Fe–2S] and [4Fe–4S] clusters (Figure 7.1). These clusters serve as a starting point for the assembly of more complex metallocofactors such as FeMo cofactors of nitrogenases or the H-cluster in hydrogenases. Besides the rhombic [2Fe–2S] clusters as found in ferredoxins (Morales *et al.*, 1999) and cubic [4Fe–4S] clusters as found in the quinolinate synthetase (Cicchillo *et al.*, 2005), [3Fe–4S] clusters as found for example in the enyzmatically inactive form of the aconitase (Robbins & Stout, 1989) as well as the rubredoxin-like [1Fe] cofactors (Bachmayer *et al.*, 1967; Maher *et al.*, 2004; Watenpaugh *et al.*, 1972) belong to the most widly spread iron-containing metallostructures. In rubredoxin-like [1Fe] centers, the iron atom is coordinated by four cysteine residues in a tetrahedral geometry. In [2Fe–2S], [3Fe–4S], and [4Fe–4S] complexes, protein bound Fe ions are linked to each other through sulfide bridges. Fe–S clusters are predominantly ligated by cysteine residues, but also histidine, arginine, serine, tyrosine or glutamate residues where reported to be involved in the coordination of Fe–S clusters (Freibert *et al.*, 2018). However, the experimental characterization of Fe–S clusters is not trivial due to Fe–S clusters being prone to react with oxygen, resulting in the damage or loss of the Fe–S clusters upon aerobic isolation of the proteins. That is why mainly Fe–S cluster containing proteins are isolated in their apo form, especially, when they are over-expressed in *Escherichia coli* (*E. coli*). However, the clusters of many simple Fe–S proteins can be reconstituted under anaerobic conditions using either chemical (Meyer *et al.*, 1986) or semi-enzymatic strategies (Gao *et al.*, 2013).

In vivo assembly of Fe–S clusters is a highly regulated process. This is most likely due to the toxicity of free iron and sulfides (Müller *et al.*, 2015; Munday, 1989). Several machineries have been identified that are involved in the assembly of Fe–S clusters: (i) The Iron Sulfur Cluster (ISC) machinery is responsible for housekeeping cluster assembly under standard conditions in bacteria and eukaryotic mitochonria (reviewed in Lill, 2009)). (ii) The Sulfur Utilization Factor (SUF) is used under stress conditions in bacteria and chloroplasts of eukaryotes. Furthermore, SUF is the only machinery for Fe–S cluster assembly in archae, cyanobacteria, and many Gram-positive, thermophilic, and pathogenic bacteria (for review see Pérard & de Choudens, 2018) (iii) The NItrogen Fixation (NIF) system was found in nitrogen-fixing bacteria such as *Azotobacter vinelandii* and is responsible for nitrogenase maturation (Hu & Ribbe, 2016; Py & Barras, 2015; Zheng *et al.*, 1998). (iv) In eukaryotes, the Cytosolic Iron-Sulfur Cluster Assembly (CIA) system was found for cytosolic and nuclear Fe–S cluster assembly (for review see Sharma *et al.*, 2010). (v) The Cysteine Sulfinate Desulfinase (CSD) from *E. coli* was reported to have components similar to the ISC or SUF system, but lacks a scaffold protein (Loiseau *et al.*, 2005).

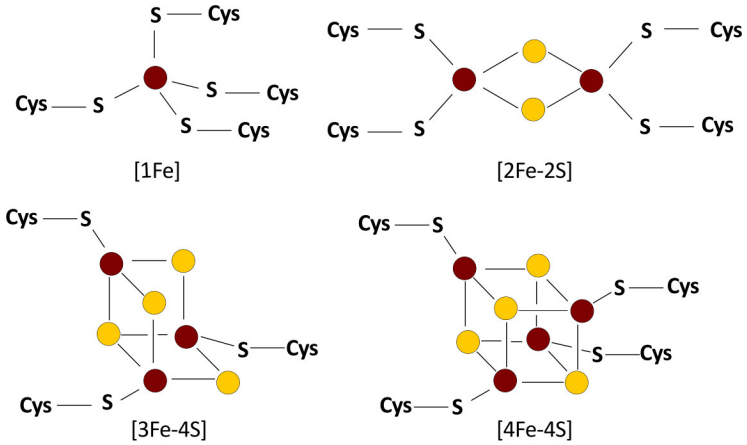


Figure 7.1: Schematic structural represetation of the most common Fe–S clusters. The simplest cluster is the [1Fe] cluster as found in rubredoxins. Here the iron is coordinated in a tetrahedral manner by four cysteine residues. The most prominent clusters are the [2Fe–2S] and [4Fe–4S] clusters. [3Fe–4S] clusters often result from the degradation of [4Fe–4S] clusters.

NIF, ISC, SUF, and CIA machinery all follow the same overall biogenesis process that can be split into two main steps: The first step involves the *de novo* synthesis of an Fe–S cluster on a scaffold protein. The second step involves the transfer of the assembled Fe–S cluster from the scaffold protein to the target apo protein. The general priciples that are involved in the assembly and insertion of Fe–S clusters into target apo proteins are similar for all different machineries. However, the concrete proteins have not been identified for all machineries. Nevertheless, the main actors and key steps in the Fe–S cluster assembly and insertion process can be split into five categories (Lill, 2009).

(1) A **sulfur donor**, a cysteine desulfurase, is required for the release of sulfur by the conversion from cysteine to alanine. In bacteria, the cysteine desulfurase is termed NifS (Zheng *et al.*, 1993), IscS (Urbina *et al.*, 2001), or SufS (Bühning *et al.*, 2017). During the transformation process persulfide is formed on a conserved cysteine residue of the cysteine desulfurase as an intermediate, which is then transferred to an assisting protein such as SufE (Ollagnier-de Choudens *et al.*, 2003a) or directly to scaffold proteins, for example IscU (Wu *et al.*, 2002), SufU (Albrecht *et al.*, 2010) or NifU (Johnson *et al.*, 2005).

(2) Since iron is not free in cells, an **iron donor** is required for the accurate and efficient delivery to the scaffold protein. Such iron donors were for example reported to be the bacterial CyaY (Layer *et al.*, 2006) and the human frataxin (Yoon & Cowan, 2003) as ISC components. These proteins bind iron, the desulfurase as well as the scaffold protein such as IscU.

(3) Another crucial step in the assembly of Fe–S clusters is the **electron transfer** for the reduction of S^0 as present in cysteine to S^{2-} as present in Fe–S clusters. The electron transfer is performed by ferredoxin reductases or ferredoxin (Fdx) (Mühlenhoff *et al.*, 2003; Yan *et al.*, 2015) in the ISC machinery. In the NIF system, this role is taken by the ferredoxin-like domain of NifU (Lill, 2009).

(4) The *de novo* assembly of Fe–S clusters from iron and sulfide takes place on so-called **scaf-**

fold proteins. These proteins labilely bind Fe–S clusters over conserved cysteine residues, to guarantee the transfer to the target protein. In bacteria, the most conserved scaffold proteins are reported to be IscU and SufU, while in eukaryotes Isu1 (Wu *et al.*, 2002) performs this function.
(5) The final step in the maturation of Fe–S cluster containing proteins is the transfer of the assembled Fe–S cluster to the target protein. This step is mediated by **cluster transfer proteins.** These proteins are required for the dissociation of the cluster from the scaffold protein, to mediate the transfer to the apo target Fe–S cluster protein and to induce the correct assembly and insertion of the cluster at the correct acceptor side in the apo protein. Bacterial examples of those cluster transfer proteins are the chaperones HscA and HscB (Chandramouli & Johnson, 2006) of the ISC machinery.

The general pathway of Fe–S cluster assembly performed by ISC, SUF, NIF, or CIA is shown in Figure 7.2.

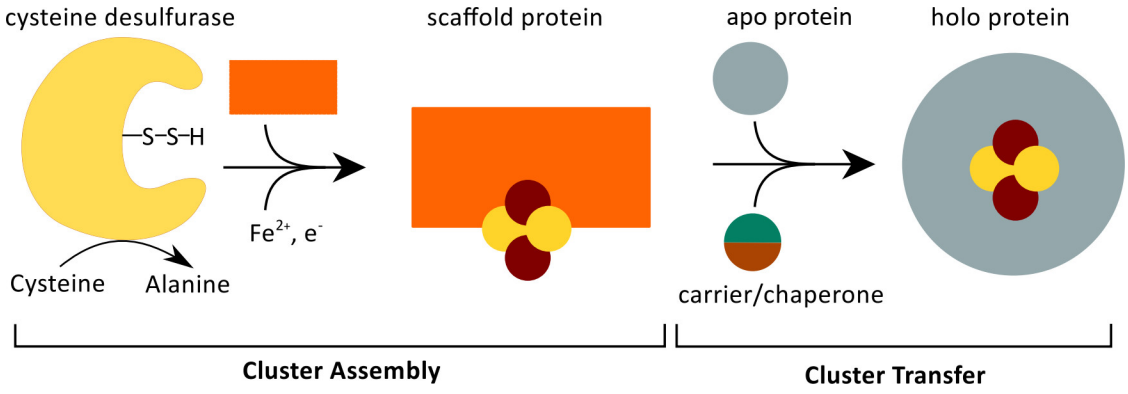


Figure 7.2: Schematic representation of Fe–S cluster assembly as performed by the ISC, SUF, NIF and CIA machineries. In a first step, a cysteine desulfurase (yellow) converts a substrate cysteine to alanine to form persulfide. Persulfide and iron provided from an iron donor as well as electrons from an electron donor are transferred to a scaffold protein (orange). On the scaffold protein, the Fe–S cluster is assembled. It is still under debate, whether the physiological form of the scaffold protein is monomeric or dimeric and whether the scaffold carries one or two [2Fe–2S] clusters or a [4Fe–4S] cluster, altough most evidence is provided for the presence of one [2Fe–2S] cluster *in vivo*. For the assembly of [4Fe–4S] clusters additional electrons are required for the reductive dimerisation of [2Fe–2S] clusters to form one [4Fe–4S] cluster (not shown). The assembled Fe–S cluster is finally transferred with the support of chaperones or carrier proteins (brown/green) to the apo target protein (grey). Figure based on Blanc *et al.* (2015); Lill (2009).

7.1 Iron-Sulfur Cluster (ISC) machinery

Under standard growth conditions the ISC machinery is responsible for the housekeeping cluster assembly in bacteria and eukaryotic mitochondria. In addition to several open reading frames of unknown function, the *iscRSUA-hscBA-fdx* (*isc*) operon encoding the proteins of the ISC machinery contains genes for molecular chaperones (*hscA* and *hscB*), an electron transferring [2Fe–2S] ferredoxin (*fdx*), and three *isc* genes, *iscS*, *iscU*, and *iscA* as well as the gene encoding for the transcriptional regulator of this operon, *iscR* (Blanc *et al.*, 2015; Johnson *et al.*, 2005). During the ISC-mediated Fe–S biosynthesis process in *E. coli* the cysteine desulfurase IscS provides the sulfur to form a labile Fe–S cluster on the scaffold protein IscU. A complex formed by the chaperone proteins HscB and HscA with IscU facilitates the cluster transfer to the apo target protein. In an alternative way, the A-type carrier IscA supports the transfer of the Fe–S cluster from IscU to the apo target protein. A [2Fe–2S] ferredoxin that is also encoded on the *isc* operon promotes the electron transfer in the reduction of S^0 to S^{2-} and the transcription factor IscR functions as a transcriptional regulator of the operon.

In the following, the most important game players in iron-sulfur cluster biosynthesis that are part of the *isc* operon (Figure 7.3) will be briefly discussed.

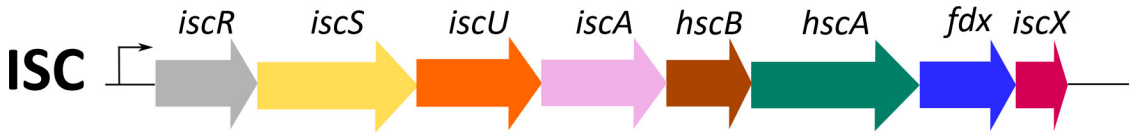


Figure 7.3: Operon organization of the genes encoding for the proteins of the ISC machinery in Gram-negative bacteria such as *E. coli* (Zheng *et al.*, 1998). The figure is color-coded according to the function of the corresponding gene products: transcription regulator iscR in gray, sulfur-activating enzyme IscS in yellow ,scaffold protein IscU in orange, iron donor or cluster carrier in light pink, ATPase HscA in green, co-chaperone HscB in brown, electron donor ferredoxin (Fdx) in blue, and a protein of unknown function, IscX, in light red. Adapted from Blanc *et al.* (2015).

The transcriptional regulator IscR is a [2Fe–2S] cluster protein that controls the expression of more than 40 genes in *E. coli* (Giel *et al.*, 2006). Although it was found that transcription of the *isc* operon is repressed by IscR (Schwartz *et al.*, 2001), this autoregulatory mechanism is not fully understood and how Fe–S biogenesis is regulated *in vivo* still remains to be elucidated completely. However, it was shown that two different classes of IscR binding sites exist on DNA, which are reffered to as Type 1 and Type 2 (Giel *et al.*, 2006). While [2Fe–2S] bound IscR is capable of binding to the Type 1 site with higher affinity than IscR without any Fe–S cluster, both IscR forms bind to the Type 2 site with the same affinity (Rajagopalan *et al.*, 2013). It has been proposed that under aerobic conditions more Fe–S clusters are required due to an increased rate of Fe–S cluster turnover, leading to a higher amount of apo proteins. In turn, this leads to a competition between these apo proteins and the IscR protein for Fe–S clusters. As a

result, the occupancy of [2Fe–2S] in IscR is low, and IscR does not repress the expression of the *isc* operon. In contrast, under anaerobic conditions, the rate of cluster turnover is reduced, the IscR [2Fe–2S] occupancy is increased and IscR represses the expression of the *isc* genes by binding to the Type 1 DNA binding site (Giel *et al.*, 2013). The proposed mechanism for the regulation of the *isc* operon by the transcriptational regulator IscR in the presence and abscence of oxygen is depict in Figure 7.4.

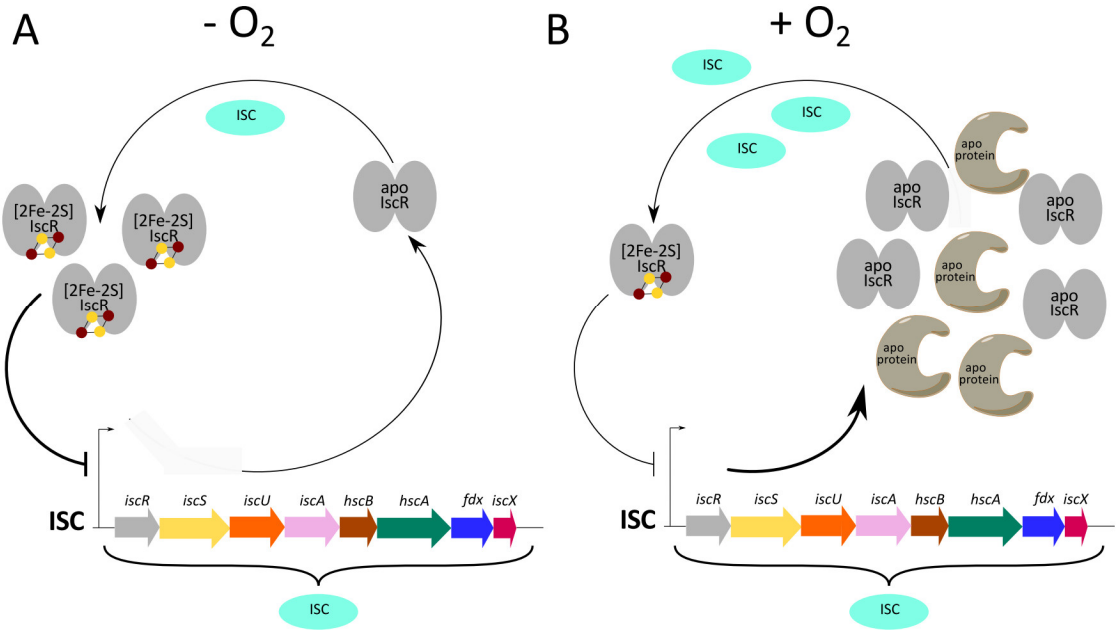


Figure 7.4: Proposed model for the IscR regulated Fe–S cluster biosynthesis under (A) anaerobic and (B) aerobic conditions. In the presence of oxygen, a high amount auf Fe–S clusters is needed due to an increased rate of Fe–S cluster turnover resulting in elevated levels of apo target proteins. This results in a competition between the apo target proteins and IscR proteins in acquiring Fe–S clusters from the ISC machinery, resulting in low IscR [2Fe–2S] cluster occupancy and a reduced repression of the *isc* operon. Under anaerobic conditions, the demand for Fe–S clusters is reduced compared to aerobic conditions, due to a decreased rate of Fe–S custer turnover. This leads to an increased IscR [2Fe–2S] cluster occupancy, and a higher repression of the ISC pathway. Figure based on Giel *et al.* (2013).

Two of the nine genes encoded by the *isc* operon, *iscS* and *iscU*, encode for homodimeric proteins. The gene products of *iscS* and *iscU* have a high similarity with the extensively studied proteins NifS and NifU (Dean *et al.*, 1993; Fu *et al.*, 1994; Zheng & Dean, 1994; Zheng *et al.*, 1994). IscS is a highly conserved cysteine desulfurase that catalyzes the transformation from L-cysteine to alanine and sulfur. The sulfur is stored as a persulfide at a conserved cysteine residue of the protein. In solution, IscS is a homodimer with a pyridoxal-phospahte (PLP) cofactor (Flint, 1996; Zheng *et al.*, 1998). IscS delivers sulfur to numerous sulfur-accepting proteins such as the primary scaffold protein for the assembly of Fe–S cluster proteins, IscU (Agar *et al.*, 2000a), ThiI, TusA, and MoaD/MoeB that commit the sulfur to different metabolic

pathways, including iron-sulfur cluster assembly, thiamine and biotin synthesis, tRNA modifications, or molybdopterin biosynthesis (Johnson *et al.*, 2005; Kessler, 2006; Zhang *et al.*, 2010a). In addition, IscS interacts with several other proteins such as the bacterial homolog of human frataxin, CyaY (Adinolfi *et al.*, 2009; Layer *et al.*, 2006), a possible adaptor protein, IscX, whose gene is also located in the *isc* operon, but whose function is currently not well understood,(Pastore *et al.*, 2006; Tokumoto *et al.*, 2002), and rhodanese RhdA (Forlani *et al.*, 2005). CyaY/frataxin has been postulated as an iron chaperone (Bou-Abdallah *et al.*, 2004), an iron donor for Fe–S cluster assembly (Ding *et al.*, 2007; Layer *et al.*, 2006; Li *et al.*, 2009), or a regulator of Fe–S cluster formation (Adinolfi *et al.*, 2009). However, in contrast to the catalytic mechanism of the cysteine desulfurase, which has been intensively studied (Tirupati *et al.*, 2004; Zheng *et al.*, 1994), little is known about how transfer of persulfide sulfur to an acceptor protein occurs. The crystal structure of the IscS dimer from *E. coli* revealed that the catalytic cysteine residue (Cys328) is located in a partially ordered active site loop, pointing away from the active site pocket of IscS where the PLP cofactor as well as the bound cysteine substrate are found (Cupp-Vickery *et al.*, 2003). This active site loop in IscS was reported to be significantly longer than in the cysteine desulfurases NifS or SufS, resulting in an open conformation in IscS in contrast to the closed conformation in NifS and SufS (Shi *et al.*, 2010). Since the active residue Cys328 has a dual role in the sulfur transfer by being directly involved in the attack of the substrate cysteine as well as by taking part in the transfer of persulfide to the accpetor scaffold protein (Urbina *et al.*, 2001; Zheng *et al.*, 1994), it has been assumed that the relatively open orientation and the longer loop sequence in which the active Cys328 is located are crucial for the function of Cys328 to target different acceptor proteins (Shi *et al.*, 2010) that are not exclusively related to Fe–S cluster biosynthesis. In contrast, in the proteins NifS and SufS, the catalytic cysteine is located in closer proximity to the PLP cofactor, a structural feature that is in agreement with direct sulfur loading from the cysteine substrate (Blanc *et al.*, 2015).

IscU is the scaffold protein of the ISC machinery and it is supposed to be one of the most conserved proteins in nature (Hwang *et al.*, 1996). Its function has been studied both *in vivo* and *in vitro* (Agar *et al.*, 2000a; Raulfs *et al.*, 2008; Tokumoto & Takahashi, 2001). *In vitro* studies on dimeric *E. coli* IscU under anaerobic conditions reveal that the protein can be reconstituted with iron and sulfur to generate a holo enzyme. This reconstitution occurs within hours in a sequential manner: At first, one [2Fe–2S] cluster is incorporated per dimer, followed by a second [2Fe–2S] cluster. In a third step, the two [2Fe–2S] cluster form one [4Fe–4S] cluster (Agar *et al.*, 2000a). Under aerobic conditions, the [4Fe–4S] cluster is converted back to a [2Fe–2S] cluster. Several studies support the hypothesis that in its physiological form IscU harbors one [2Fe–2S] cluster per dimer. An X-ray structure of trimeric IscU from *Aquifex aeolicus* was solved in the presence of a [2Fe–2S] cluster (Shimomura *et al.*, 2008) and by *in vitro* experiments it could be shown that the rate of Fe–S transfer from IscU with one [2Fe–2S] cluster per dimer to apo target proteins could be enhanced in the presence of chaperones, whereas this effect could not be observed for IscU proteins harboring either two [2Fe–2S] cluster or one [4Fe–4S] cluster per dimer (Shakamuri *et al.*, 2012; Unciuleac *et al.*, 2007). Additionally, it has been observed that one of the [2Fe–2S] clusters in the IscU form with two [2Fe–2S] clusters per dimer can

be selectively removed in the presence of EDTA (Agar *et al.*, 2000a). Spectroscopic studies on IscU in either the one or two [2Fe–2S] cluster-containing form provide evidence that the [2Fe–2S] clusters have both very similar protein environments with at least one and possibly two oxygen or nitrogen ligands at one iron site (Agar *et al.*, 2000a). Since IscU in *E. coli* and *A. vinelandii* contains three conserved cysteine residues (Zheng *et al.*, 1998) it is likely that therefore each IscU monomer can harbor one [2Fe–2S] cluster, but it cannot be excluded that both clusters bridge the subunits. Discussing all these properties of the dimeric isolated IscU, one has to keep in mind that IscU does not form a complex with the cysteine desulfurase IscS in its dimeric form. Isothermal titration calorimetry (ITC) experiments showed that IscU interacts with IscS in a 1:1 stoichiometry (Prischi *et al.*, 2010). Two crystal structures of the IscS-IscU complex have been solved so far (Marinoni *et al.*, 2012; Shi *et al.*, 2010) and both provide evidence that one IscU monomer binds near the C-terminal portion of an IscS dimer.

The protein IscA is member of the so-called A-type carrier family. Studies in *E. coli* and *Saccharomyces cerevisiae* using gene disruptions reveal that the protein is important but not essential in Fe–S cluster biosynthesis (Jensen & Culotta, 2000; Pelzer *et al.*, 2000; Tokumoto & Takahashi, 2001). IscA has three conserved cysteine residues. Spectroscopic analysis of IscA in *A. vinelandii* suggest the coordination of either a [2Fe–2S] cluster or a [4Fe–4S] cluster by these cysteine residues to be more likely than the coordination of a Fe(II) or Fe(III) (Krebs *et al.*, 2001; Ollagnier-de Choudens *et al.*, 2001). Based on these studies, a role for IscA in the delivery of iron to cluster assembly proteins seems unlikely. However, *in vitro* studies of *E. coli* IscA in the presence of oxygen with Fe^{2+} and Dithiothreitol (DTT) or the thioredoxin reductase system revealed that IscA binds one iron atom per homodimer with high affinity (Ding & Clark, 2004; Ding *et al.*, 2005). The different characterizations of IscA as either an Fe–S cluster or an Fe-binding protein led to the proposal of different roles of this protein as a scaffold protein under aerobic or oxidative stress conditions, an Fe–S carrier protein that mediates the delivery of an Fe–S cluster from IscU to the apo target protein, or as a chaperone for the delivery of iron to IscU (Blanc *et al.*, 2015)

HscA and HscB are two proteins encoded by the *isc* operon. HscA is a member of the Hsp70 chaperone family, while HscB belongs to the so-called auxiliary J-type co-chaperones. In addition to their role in stress responses, Hsp70s are involved in housekeeping functions, including the folding of newly synthesized proteins, translocation of proteins across membranes, cellular trafficking, and control of regulatory proteins (Mayer *et al.*, 2001). Initial studies on the properties of HscA and HscB of *E. coli* have shown that HscA exhibits a low intrinsic ATPase activity (Vickery *et al.*, 1997), as it is typical for other members if the Hsp70 family. In the presence of the co-chaperone HscB, the ATPase activity is significantly enhanced. It was shown that HscA recognizes a specific amino acid sequence motif on IscU (LPPVK). This interaction is regulated by HscB, which interacts with IscU through hydrophobic residues (Füzéry *et al.*, 2008; Hoff *et al.*, 2003). However, the exact mechanism by which HscA-HscB promotes the Fe–S cluster assembly is still controversial. Roles in stabilizing the Fe–S clusters on IscU were assumed based on *in vitro* studies (Wu *et al.*, 2005), while *in vivo* studies suggest a contribution in cluster transfer from IscU to the apo form of the target protein (Mühlenhoff *et al.*, 2003).

However, as already discussed in the section dealing with IscU it was found that *in vitro* the HscA–HscB system enhances the rate of [2Fe–2S] cluster transfer from the IscU dimer, but that it has no effect on the rate of [4Fe–4S] cluster transfer from IscU dimer to the apo target protein. These observations suggest the presence of a specific conformation of IscU harboring a [2Fe–2S] cluster that is able to interact with chaperones (Chandramouli & Johnson, 2006; Shakamuri *et al.*, 2012; Unciuleac *et al.*, 2007). However, it is still under debate, whether these properties observed *in vitro* truly reflect the physiological properties. Nevertheless, based on *in vitro* analyses, two mechanisms have been proposed for the interaction of HscA-HscB with IscU in the transfer process of [2Fe–2S] clusters that are depict in Figure 7.5 (Bonomi *et al.*, 2011; Kim *et al.*, 2012b).

Ferredoxins are small electron transfer proteins that are found in biological redox systems. They harbor either a [4Fe–4S], a [3Fe–4S], or a [2Fe–2S] cluster with highly negative reduction potentials (Beinert, 1990; Holden *et al.*, 1994; Matsubara & Saeki, 1992). Ferredoxins with one [2Fe–2S] cluster is found in plants, animals, and bacteria (Matsubara & Saeki, 1992; Müller *et al.*, 1999). They have four cysteine residues that coordinate the Fe–S cluster, however, their amino acid sequences differ depending on the biological process in which they are involved. Deletion studies in *E. coli* have shown that ferredoxins are cruicial in the assembly process of Fe–S clusters (Tokumoto & Takahashi, 2001). It was revealed that ferredoxin (Yah1p in yeast) is involved in the formation of the Fe–S cluster on the scaffold protein IscU by the reduction of persulfide (Lange *et al.*, 2000). An additional function of ferredoxin is supposed to be the reductive coupling of two [2Fe–2S] clusters to one [4Fe–4S] on the scaffold protein IscU (Chandramouli *et al.*, 2007). Furthermore, pull-down assays in *E. coli* suggest interactions of ferredoxin with IscS in the Fe–S cluster assembly step and with HscA and IscA in the transfer step of the Fe–S cluster to the target apo-protein (Ollagnier-de Choudens *et al.*, 2001; Tokumoto *et al.*, 2002).

IscX is a small protein and its function has not been satisfactorily understood. Studies in which the *iscX* gene has been deleted do not reveal any specific phenotype, except a small decrease of the activity of the transcriptional regulator IscR (Prischi *et al.*, 2010; Tokumoto *et al.*, 2002). Structural studies reveal a motif that is typically present in DNA-binding proteins (Pastore *et al.*, 2006; Shimomura *et al.*, 2005). However, it was shown that IscX can inhibit the formation of Fe–S clusters on the scaffold protein IscU, but does not affect the desulfurase activity of IscS (Kim *et al.*, 2014). These results strongly suggest the function of IscX as an adaptor that can modulate Fe–S biosynthesis, rather than being a transcription factor as suggested by its structure.

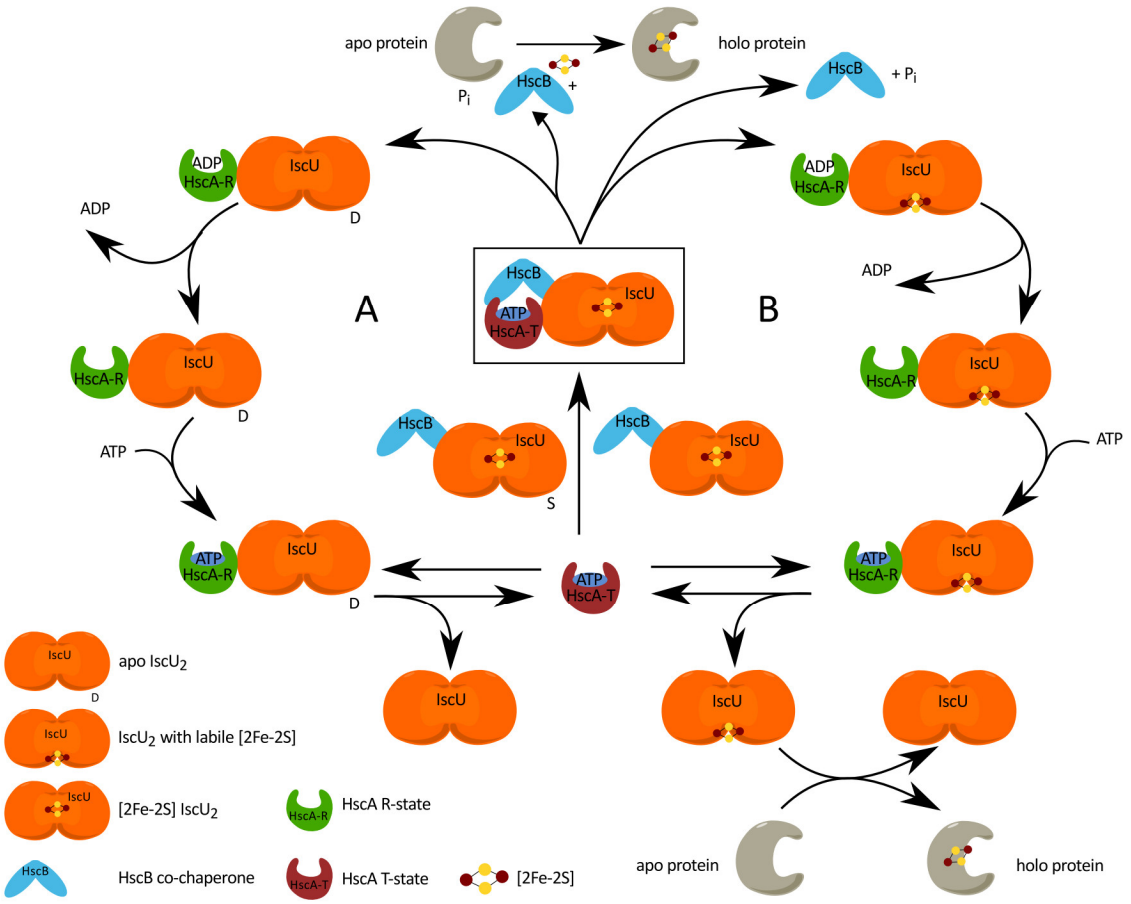


Figure 7.5: Proposed models for chaperone-mediated transport of Fe–S clusters from IscU dimers ($IscU_2$) to the target apo proteins. (A) The co-chaperone HscB is bound to $IscU_2$ in its structured state (S-form) and targets the ATP-bound form of the chaperone HscA (T-state with low substrate affinity), resulting in a HscA-HscB-holo-$IscU_2$ complex. Interaction with an Fe–S cluster acceptor apo target protein leads to the hydrolysis of ATP to ADP and P_i and the conformation of HscA changes from T-state to ADP-bound form (R-state with high substrate affinity), which stabilizes the disordered form (D-form) of $IscU_2$. This leads to the release of the Fe–S cluster and of HscB. HscA remains bound to the disordered $IscU_2$ after the release of ADP, but subsequent binding of ATP induces a conformational change of HscA to the T-state. HscA in its T-state releases $IscU_2$ in its disorderes apo form (Kim *et al.*, 2012b). (B) The second model involves two isomers of [2Fe–2S]–bound $IscU_2$ that differ by their affinity for the [2Fe–2S] cluster due to different metal coordination (Bonomi *et al.*, 2008). HscB is released from a HscA-HscB-holo-IscU after ATP hydrolysis and the [2Fe–2S] shifts to the IscU isomer in which the cluster is less tightly bound and which is stabilized by ADP bound HscA or HscA in the absence of ADP or ATP. The transition of HscA to the ATP-bound state leads to the release of the [2Fe–2S]–$IscU_2$ protein then able to transfer its Fe–S cluster to the apo target protein (Bonomi *et al.*, 2011). Figure adapted from Blanc *et al.* (2015).

7.2 Sulfur Utilization Factor (SUF)

While the ISC pathway is widely spread across almost all domains of life, the SUF system is less wide distributed. Genes encoding for the SUF machinery are located in the *suf* operon (Figure 7.6), which is in *E. coli* under the control of the iron-dependent Fur regulator (Fontecave *et al.*, 2005). The SUF machinery is used under stress conditions and it is the sole machinery for Fe–S cluster biogenesis in archae, cyanobacteria, and many Gram-positive, thermophilic and pathogenic bacteria, but it is not found in yeast or animals (Bai *et al.*, 2018).

The specific process of Fe–S cluster assembly via the SUF pathway involves the cysteine desulfurase SufS, a PLP-dependent enyzme, which catalyzes the transformation of cysteine to alanine to form persulfide that is bound to a specific cysteine residue of the enzyme to form a stable intermediate (Bai *et al.*, 2018). The reactivity of SufS was found to be dramatically increased in the presence of SufE, a sulfur transfer protein that receives the persulfide from SufS via a conserved cysteine residue (Loiseau *et al.*, 2003; Ollagnier-de Choudens *et al.*, 2003a; Outten *et al.*, 2003). SufE transfers the sulfur to the scaffold complex formed by $SufBC_2D$, the iron donor remains unknown. The Fe–S cluster is preassembled on the scaffold complex and then either directly transferred to the apo target protein or via carrier proteins (Bai *et al.*, 2018).

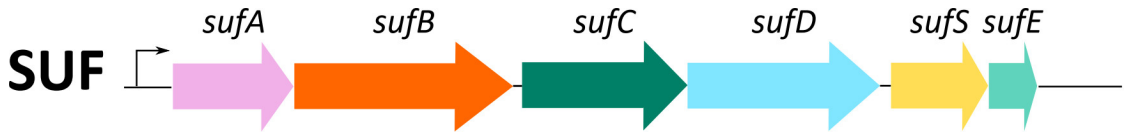


Figure 7.6: Operon organization of the genes encoding for the proteins of the SUF machinery in Gram-negative bacteria such as *E. coli* (Takahashi & Tokumoto, 2002). The figure is color-coded according to the function of the corresponding gene products: sulfur-activating enzyme SufS in yellow ,scaffold protein SufB in orange, iron donor or cluster carrier SufA in light pink, ATPase SufC in green, a sulfur acceptor protein SufE in light green, and a protein of unknown function, SufD, in cyan. Figure adapted from Blanc *et al.* (2015).

In the following, the different gene products encoded by the *suf* operon will be briefly discussed.

In *E. coli* the protein SufA is highly similar to the protein IscA, both in structure and function (Bilder *et al.*, 2004; Wada *et al.*, 2005). Deletion studies show that the proteins seem to have redundant functions (Lu *et al.*, 2008; Vinella *et al.*, 2009). However, as for IscA the exact function is unknown. Spectroscopic analysis of *E. coli* SufA isolated under anaerobic conditions revealed that the protein hosts a [2Fe–2S] cluster *in vivo*. Additionally it was shown that SufA can transfer its [2Fe–2S] cluster to apo [2Fe–2S] and [4Fe–4S] proteins (Gupta *et al.*, 2009; Ollagnier-de Choudens *et al.*, 2003b), supporting the theory that SufA functions as a carrier for a subset of Fe–S proteins.

SufB contains several conserved cysteine residues that are capable of assembling Fe–S clusters (Ayala-Castro *et al.*, 2008; Fontecave & Ollagnier-de Choudens, 2008). The protein plays a key role in the assembly of Fe–S clusters, since it functions as the core protein of the $SufBC_2D$ scaffold complex. In the SUF pathway of Fe–S biosynthesis, SufB receives persulfide from the cysteine desulfurase via the sulfur transfer protein SufE (Layer *et al.*, 2007). In the stable $SufBC_2D$ complex, SufB interacts tightly with the proteins SufC, an ATPase, and SufD (Nachin *et al.*, 2003; Outten *et al.*, 2003). This complex was shown to be capable of transfering Fe–S clusters to SufA (Chahal *et al.*, 2009); cluster transfer to aconitase B was reported to occur without the requirement of SufA (Wollers *et al.*, 2010).

In addition to the $SufBC_2D$ complex, a stable $SufB_2C_2$ complex has been found *in vivo* (Saini *et al.*, 2010). Semi-enzymatic reconstitution of SufB leads to the formation of a [4Fe–4S] cluster, which is degraded to a [2Fe–2S] cluster when exposed to oxygen (Layer *et al.*, 2007). The formation of a [4Fe–4S] cluster in SufB has also been observed, when the protein was part of the $SufBC_2D$ scaffold complex (Wollers *et al.*, 2010). The co-expression of *sufB* with the genes *sufCDSE* led to the isolation of the complexes $SufBC_2D$ and $SufB_2C_2$ with [3Fe–4S] and [4Fe–4S] clusters (Saini *et al.*, 2010).

The $SufBC_2D$ complex also binds one $FADH_2$ molecule per complex, which has been proposed to function in the reduction of Fe^{3+} to Fe^{2+} or persulfide (S^0) to sulfide (S^{2-}) (Saini *et al.*, 2010; Wollers *et al.*, 2010).

The protein SufC is another component of the $SufBC_2D$ scaffold complex. The protein was reported to be a member of ABC-type ATPases that are mostly found to be associated with membrane proteins, forming a complex that allows translocation across membranes (Holland & Blight, 1999). The crystal structure of *E. coli* SufC revealed that despite low sequence similarity with other ABC-type ATPases, the secondary structures and overall architectures are similar (Kitaoka *et al.*, 2006). However, no such function was found for SufC. As previously mentioned, SufC complexes interact with SufB and SufD to form a scaffold complex for the assembly of Fe–S clusters. When ATP is bound to SufC, two SufC proteins form a dimer. This process was reported to induce a conformational change within the $SufBC_2D$ complex that enables the formation of an Fe–S cluster (Hirabayashi *et al.*, 2015; Yuda *et al.*, 2017).

The biological function of the protein SufD is still not understood. However, it has been proposed to be important for the acquisition of iron during the assembly process (Saini *et al.*, 2010). The protein SufD is homologous to SufB and both proteins share the same domain organization (Yuda *et al.*, 2017).

The protein SufS is the cysteine sulfurase within the SUF pathway. This homodimeric protein is PLP-dependent and catalyzes the transformation of cysteine to alanine to form persulfide that is required for the assembly of Fe–S clusters (Dos Santos, 2017b; Mihara & Esaki, 2002). The activity of SufS is more than tenfold enhanced when another homodimeric protein, SufE, binds to the protein (Lill, 2009). The interaction between SufS and SufE also allows the persulfide bound to a cysteine residue in SufS to be transferred to a conserved cysteine residue on SufE.

From SufE, the persulfide is then transferred to the scaffold protein SufB (Layer *et al.*, 2007). Surprisingly, the structure of SufE is very similar to the structure of U-type scaffold proteins such as IscU. Although, SufE does not function as a scaffold protein (Liu *et al.*, 2005). The structurally characteristic loop that hosts the active cysteine in IscS was not found in SufS. For this reason, sulfur transfer from SufS typically involves the participation of a sulfur acceptor protein, such as SufE in *E. coli*, which serves as an intermediate sulfur-carrier to a scaffold protein (Blauenburg *et al.*, 2016; Tirupati *et al.*, 2004).

7.3 Fe–S Cluster assembly in eukaryotes

The early steps of Fe–S cluster assembly in eukaryotes such as *Saccharomyces cerevisiae* (*S. cerevisiae*) are very similar to what has been discussed for the ISC machinery in *E. coli*. In *S. cerevisiae* the process begins within the mitochondria with the components of the ISC machinery as already described for *E. coli* (Section 7.1). The early proteins of the ISC machinery include the Isu1 scaffold protein requiring sulfide from the cysteine desulfurase complex Nfs1-Isd11-Acp1 (Braymer & Lill, 2017; Van Vranken *et al.*, 2016), electrons from the transfer complex NADPH-Arh1 and the ferredoxin Yah1 (Webert *et al.*, 2014), and the regulator and/or iron donor Yfh1 (Parent *et al.*, 2015). All these proteins are required for the synthesis of [2Fe–2S] clusters on the scaffold protein Isu1 (Mühlenhoff *et al.*, 2003). The [2Fe–2S] cluster is then transferred to the monothiol glutaredoxin Grx5 (Uzarska *et al.*, 2013). This transfer is assisted by the chaperone Ssq1, an ATPase, and the co-chaperone Jac1 (Kampinga & Craig, 2010; Majewska *et al.*, 2013). From Grx5, the [2Fe–2S] cluster is either directly transferred to an apo target protein or it is transported to the late proteins of the ISC machinery, namely the protein complex Isa1-Isa2-Iba57, to assemble [4Fe–4S] clusters (Mühlenhoff *et al.*, 2011; Sheftel *et al.*, 2012). In addition to the [2Fe–2S] cluster proteins of the early ISC machinery an uncharacterized X-S component is generated that is required for the transport of sulfur to the components of the CIA machinery for the assembly of cytosolic or nuclear Fe–S cluster containing proteins. The final trafficking and insertion of [4Fe–4S] clusters into the apo target proteins was reported to require the proteins Nfu1, Inc1 and Bol. However, this last step is not well understood (Braymer & Lill, 2017).

7.4 Fe–S cluster assembly in Gram-positive bacteria

The formation of Fe–S clusters in Gram-positive bacteria such as *Bacillus subtilis* (*B. subtilis*) is far less understood than the according processes in Gram-negative bacteria such as *E. coli* or in the mitochondria of eukaryotes. Nevertheless, the basic steps in Fe–S cluster assembly are proposed to be the same as in *E. coli*, including sulfur activation, assembly of the Fe–S cluster, and the transfer of the pre-assembled cluster onto the apo target protein (Figure 7.2) (Dos Santos, 2017b). According to Dos Santos (2017a), three different pathways of Fe–S cluster assembly have been discovered so far: The Clostridia-ISC partway, the Actinobacteria-SUF pathway, and the Bacillii-SUF pathway. A remarkable difference between Gram-positive and Gram-negative bacteria is that Gram-positive bacteria usually only host one set of genes encoding for an Fe–S cluster assembly machinery, while the genome of Gram-negative bacteria usually contains at least two Fe–S cluster assembly machineries (Huet *et al.*, 2005; Kobayashi *et al.*, 2003). Furthermore, Gram-positive bacteria have been reported to have a variety of different cysteine desulfurases encoded on different loci on the genome, each participating in the synthesis of different sulfur-containing cofactors (Black & Dos Santos, 2015). In Gram-negative bacteria, one cysteine desulfurase such as IscS is involved in sulfur acquisition for various sulfur-containing cofactors (Dos Santos, 2017a; Schwartz *et al.*, 2000).

The genes encoding for the components of the SUF machinery in Gram-positive are located in the operon *sufCDSUB* (Figure 7.7) that slightly differs from the *suf* operon found in Gram-negative bacteria (Figure 7.6) (Dos Santos, 2017b). The same is true for the *isc* operon, however, the ISC machinery of Gram-positive bacteria will not be the subject of this thesis, since it is only very rarely distributed. The *suf* operon from *B. subtilis* lacks the genes encoding for the carrier proteins SufE and SufA, but instead harbors the gene encoding for SufU. SufU was reported to be an U-type scaffold protein that shares high similarities with IscU (Riboldi *et al.*, 2009). As IscU, SufU contains three conserved cysteine residues that are involved in Fe–S cluster coordination, but lacks the characteristic LPPVK amino acid sequence. This is in agreement with the absence of the chaprone proteins HscA and HscB in *B. subtilis*. Together with the cysteine desulfurase SufS, the protein SufU catalyzes the sulfur mobilization (Selbach *et al.*, 2010). SufC is an ATPase as described for Gram-negative bacteria. Binding of SufD to SufC can enhance the ATP hydrolysis approximately 30-fold. SufBCD is proposed to have the same scaffold function as the complex $SufBC_2D$ in *E. coli* (Dos Santos, 2017a).

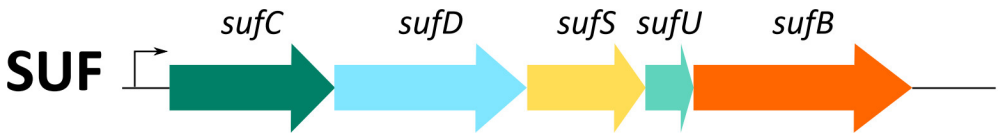


Figure 7.7: Operon organization of the genes encoding for the proteins of the SUF machinery in Gram-positive bacteria such as *B. subtilis* (Takahashi & Tokumoto, 2002). The figure is color-coded according to the function of the corresponding gene products: sulfur-activating enzyme SufS in yellow, scaffold protein SufB in orange, ATPase SufC in green, a sulfur acceptor protein SufU in light green, and a protein of unknown function, SufD, in cyan. Figure adapted from Blanc *et al.* (2015).

7.5 Aim of this thesis

Fe–S cluster assembly machineries in pro- and eukaryotes have been studied for many years and many components within the pathways and their functions have been deciphered. However, the recombinant production of Fe–S cluster containing proteins for spectroscopic and structural studies is still challenging. Isolation of recombinant holo proteins from *E. coli* is highly desirable for the collection of high quality functional and structural data. The objective of this work is to systematically investigate the synthesis of different Fe–S proteins using different strategies. We use three target proteins that all host a [4Fe–4S] cluster and that have been extensively studied in the past. These proteins are the aconitase B (AcnB), 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (IspH) and quinolinate synthase A (NadA), all from *E. coli*. The influence of different *E. coli* cell strains and expression systems on the expression levels and protein yields, as well as the Fe–S cluster occupancy will be studied. Here, the influence of the

co-expression of plasmids harboring the genes of either the ISC or the SUF machinery, as well as the knockdown of the transcriptational regulator of the ISC machinery, IscR, will be investigated. In addition to different expression systems, we compare the results of the as-isolated protein samples with protein samples obtained by chemical or semi-enzymatic reconstitution. Furthermore, we investigate the expression levels, protein yields and Fe–S cluster occupancy of the radical S-adenosylmethionine (SAM) enzyme ThnB after *in vivo* and *in vitro* cluster incorporation. The protein is involved in thioether bond formation in the sactipeptide ThnH from *B. thuringiensis* and hosts two [4Fe–4S] cluster. Previously, ThnB was isolated in the absence of the Fe–S clusters and spectroscopic analyses were performed on a chemically reconsituted sample (Wieckowski *et al.*, 2015). However, *in vitro* reconstitution of Fe–S clusters can lead to incomplete cluster formation (Archer *et al.*, 1994; Leartsakulpanich *et al.*, 2000) or iron-containing precipitate that can lead to falsified results in spectroscopic analyses (Freibert *et al.*, 2018). Furthermore, excess iron and sulfur that are added to a protein sample within the reconstitution process may have a negative impact on the crystallization properties of a protein. Therefore, the isolation of Fe–S cluster containing proteins in its holo form is important to get a realistic and unadulterated view on the protein.

A second objective of this thesis is to characterize the pyrophosphatase domain of the bifunctional pyrophosphatase-kinase Asp1 from *Saccharomyces pombe*. Previously, it was reported that chemical reconstitution of the pyrophosphatase domain of Asp1 ($Asp1^{365\text{-}920}$) leads to the formation of a [2Fe–2S] cluster (Wang *et al.*, 2015b). However, the biological function and physiological relevance of this cluster is still under debate. This thesis aims at characterizing the Fe–S cluster of "as-isolated" $Asp1^{365\text{-}920}$ samples that do not require a reconstitution process after the protein isolation from *E. coli* as well as *S. pombe*. We will use optical spectroscopy, electron paramagnetic resonance (EPR) spectroscopy, and X-ray absorption spectroscopy to study the nature, environment and redox state of $Asp1^{365\text{-}920}$. Studies performed with the "as-isolated" protein may also lead to a deeper understanding of the biological function of the protein.

All methods and protocols that will be established or used for the work on this thesis are not exclusively directed to the proteins that are under investigation. Much more this thesis aims at obtaining a deeper understanding on how different proteins can be individually addressed for structural and functional analysis. Special focus will be set on methods for the recombinant production and isolation of holo Fe–S proteins that do not require additional reconstitution steps, since chemical or semi-enzymatic reconsitution procedures may entail the risk of the formation of aggregates or biologically irrelevant clusters.

8 Comparing the maturation of [4Fe–4S] cluster containing proteins *in vivo* and *in vitro*

This chapter reflects content of the following publication.

8.1 Publication information

Hannah Rosenbach, Steffen Mielenbrink, Melissa Jansing, Sabine Metzger, and Ingrid Span
To be submitted to: **ChemBioChem**

8.2 Abstract

Fe–S clusters are ubiquitous cofactors that are essential in many biological processes. The biosynthesis of Fe–S clusters is strictly regulated in all organisms. Several strategies have been reported for the maturation of recombinantly produced proteins, however, a systematic comparison of different maturation protocols for the same target proteins has not yet been reported. In this work, we use three well-characterized [4Fe-4S] proteins, including aconitase B, 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (IspH), and quinolinate synthase (NadA), to investigate the influence of different maturation strategies. These include the maturation of the apo protein using chemical or semi-enzymatic reconstitution, co-expression with two different plasmids containing the iron-sulfur cluster (*isc*) or sulfur formation (*suf*) operon, and a cell strain lacking IscR, the transcriptional regulator of the ISC machinery. Our results show that the maturation of the Fe-S proteins *in vivo* is more specific and efficient compared to reconstitution. Surprisingly, we observed that the differences between the *in vivo* maturation systems are pronounced, although ISC and SUF should be capable of maturating the same proteins. We observed that it is difficult to predict which strategy results in successful maturation, thus, the optimal cell strain has to be determined experimentally. We also provide a rationale for preferences for SUF versus ISC pathway. Furthermore, we show that the *in vivo* strategies can be extended to more complex enzymes, such the radical SAM protein ThnB, which was previously only maturated by reconstitution. Our results shed light on the differences of *in vitro* and *in vivo* Fe–S cluster maturation and points out the pitfalls of chemical reconstitution.

8.3 Introduction

Fe–S clusters are ancient and ubiquitous cofactors in proteins. They are involved in many essential biological processes, including such as oxidative respiration, photosynthesis, hydrogen production, nitrogen fixation, and DNA replication/repair. The most common species found in proteins are the rhombic [2Fe–2S] and cubic [4Fe–4S] forms and they are predominantly coordinated by cysteine residues, however, other residues can be involved as ligands as well (Freibert *et al.*, 2018). Fe–S proteins are predominantly involved in electron transfer reactions. Furthermore, they also play important roles in providing stability to protein structures, regulation of gene expression, non-redox catalysis, repair and processing of nucleic acids, regulation of cellular processes, and iron homeostasis (Beinert, 2000; Beinert *et al.*, 1997; White & Dillingham, 2012).

Fe–S cofactors have played a significant role in the metabolism of the primordial cells, which were formed in an atmosphere with an extremely low oxygen concentration. Under these conditions, oxidation of Fe(II) to Fe(III), which is insoluble in water, by molecular oxygen leading to the degradation of the cluster was not a major issue. The versatility and stability of the cluster lead to a dominant role in primordial catalysis. With rising oxygen concentration in the atmosphere, some organisms have evolved to utilize copper proteins instead of Fe-S proteins, since both oxidation states of copper are soluble in aqueous solutions. Despite the need to protect the cofactors from oxygen, Fe–S proteins are still wildly used cofactors in nature and their key role in many biological processes makes them the focus of a large research community.

The sensitivity of Fe–S proteins towards oxygen requires special attention during gene expression, protein isolation and purification. Gene expression occurs in the cytosol with a reducing environment and virtually no free molecular oxygen, so in most cases the cells can be cultivated under aerobic conditions without an impact on the Fe–S proteins. However, cultivation of cells under shaking with low speed as well as expression under anaerobic conditions have been shown to have a beneficial effect on Fe–S protein production (Kuchenreuther *et al.*, 2010). *Escherichia coli* (*E. coli*) is a facultative anaerobic bacterium, thus, it can also be cultivated at low oxygen concentrations. After cell lysis, the Fe–S proteins are exposed to molecular oxygen and their cofactor decomposes rapidly in most cases due to the insolubility of ferric ions. Some members of the Fe–S protein family are capable of stabilizing protein-bound ferric ions, such as rubredoxins, or the clusters are protected from the environment by the protein, such as some ferredoxins, resulting in no or slow degradation of the cluster under aerobic conditions. In most cases though, isolation and purification of Fe–S proteins lead to protein in apo form or with partial cluster loading.

Another important consideration is the assembly and transfer of the cofactors to recombinantly produced proteins using the Fe–S bioassembly machinery of the host (see Figure 8.1 for a schematic representation of the components of Fe–S assembly machineries in general). While the recombinant gene is expressed to an abnormally and excessively high level, the level of the Fe–S assembly proteins remains unchanged and the machinery is not capable of keeping up with the production of the cofactor. The Fe–S cluster content in partially loaded proteins

can be enhanced by chemical reconstitution, which is performed by treating the protein with an inorganic iron and sulfide source in the presence of a reductant (Malkin & Rabinowitz, 1966). A more sophisticated variation of the reconstitution reaction is a semi-enzymatic approach that involves the protein cysteine desulfurase and L-cysteine instead an inorganic sulfur source (Figure 8.2A) (Gao *et al.*, 2013).

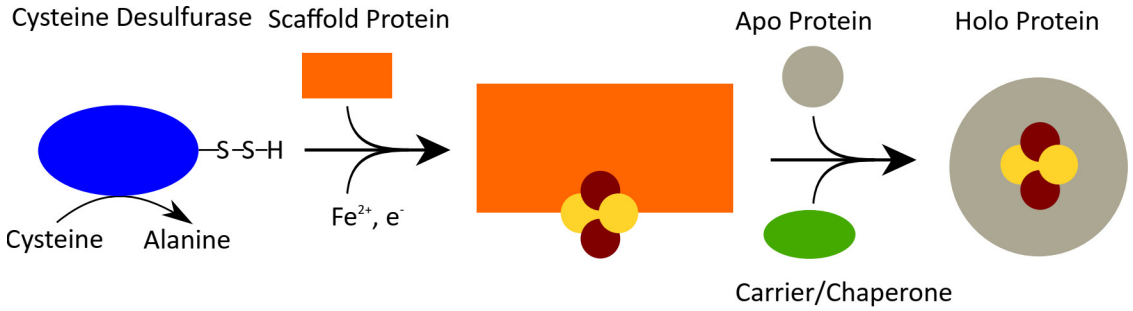


Figure 8.1: Schematic representation of iron sulfur cluster assembly machineries such as the ISC and SUF machinery. Adapted from Lill (2009).

In cases where the nature of the native Fe–S cluster has not been ascertained, chemical reconstitution can result in incomplete cluster formation or assembly of artifacts (Archer *et al.*, 1994; Leartsakulpanich *et al.*, 2000). Biosynthesis of the Fe–S clusters *in vivo* by co-expression of the genes encoding for the Fe–S cluster assembly machinery ensures the incorporation of biologically relevant cluster forms. While the Iron Sulfur Cluster (ISC) machinery is used for housekeeping cluster assembly under standard growth conditions in bacteria and eukaryotic mitochondria (reviewed byLill), the Sulfur Formation (SUF, also termed Sulfur Utilization Factor) machinery is used under stress conditions in bacteria and chloroplasts of eukaryotes (reviewed by Bai *et al.*). SUF is the sole machinery for Fe–S cluster biogenesis in archaea, cyanobacteria, and many Gram-positive, thermophilic and pathogenic bacteria (Pérard & de Choudens, 2018). Additional machineries responsible for Fe-S cluster assembly are the Nitrogen Fixation (NIF) system that plays a specialized role in the maturation of Fe–S cluster proteins in nitrogen fixing organisms such as *Azotobacter vinelandii* (Pérard & de Choudens, 2018), the Cysteine Sulfinate Desulfinase (CSD) from *E. coli* that has components similar to the ISC or SUF system, but lacks a scaffold protein (Loiseau *et al.*, 2005), and the Cytosolic Iron-Sulfur Cluster Assembly (CIA) system, that is responsible for cytosolic and nuclear Fe–S cluster assembly in eukaryotes (reviewed Sharma *et al.*). All Fe-S biosynthesis pathways follow the same pattern and include enzymes with similar function (Figure 8.1). The cysteine desulfurase converts the substrate L-cysteine into L-alanine and provides the sulfur to the scaffold protein. The requirement of an iron donor protein or the specific iron source are not clear (Iannuzzi *et al.*, 2011; Layer *et al.*,

2006; Roche *et al.*, 2015). For the bacterial pathway, it has been shown that a complex of cysteine desulfurase and scaffold protein is formed that assembles [2Fe–2S] and possibly also [4Fe–4S] clusters in the presence of an electron source (Agar *et al.*, 2000a,b; Marinoni *et al.*, 2012). Previous studies have shown that co-expression of genes encoding for Fe–S proteins with the *isc* or *suf* operon have resulted in higher protein yields and increased cofactor content (Figure 8.2B) (Gräwert *et al.*, 2004; Hänzelmann *et al.*, 2004). Furthermore, it was reported that the deletion of the gene encoding for the transcriptional regulator of the ISC machinery, iscR, in *E. coli* BL21(DE3) cells results in an enhanced expression of the *isc* operon providing as an alternative to plasmid-based overexpression (Figure 8.2B) (Schwartz *et al.*, 2001). Each method has been successfully used for one or several proteins (Akhtar & Jones, 2008; Gao *et al.*, 2013; Gräwert *et al.*, 2004; Hänzelmann *et al.*, 2004; Malkin & Rabinowitz, 1966), however, systematic studies comparing several different strategies for the same protein have not been reported. The obvious approach would be to use the biosynthesis machinery that is biological relevant, which is only possible for prokaryotic proteins. In addition, the co-expression of two plasmids may result in an additional stress for the host cells and, therefore, reduce the productivity of protein synthesis.

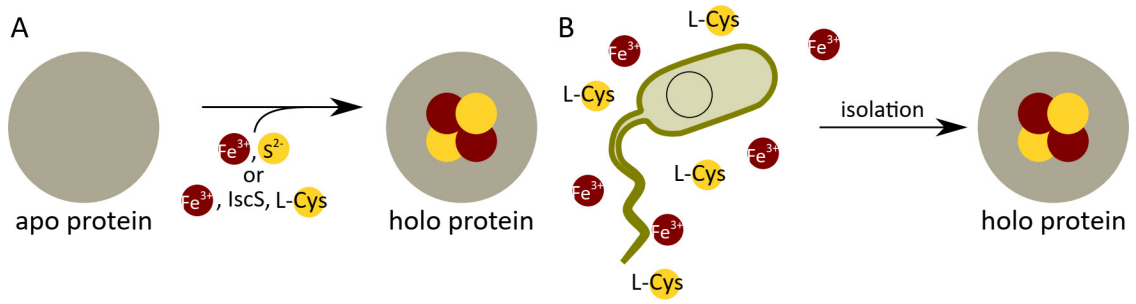


Figure 8.2: Assembly of Fe–S clusters in proteins. (A) Schematic overview on the workflow for chemical and semi-enzymatic *in vitro* reconstitution of Fe–S cluster into the apo form of a protein. While for chemical reconstitution an inorganic sulfur source is required, semi-enzymatic reconstitution requires the cysteine desulfurase IscS and L-cysteine as an organic sulfur source. Both methods require an inorganic iron source. (B) Schematic overview of *in vivo* assembly for the isolation of Fe cluster proteins in their holo form using the co-expression of the target protein with the plasmids pSUF or pISC in BL21(DE3) as well as the expression of the target protein in BL21(DE3) Δ*iscR*. Cells are cultivated in growth media supplemented with inorganic iron and L-cysteine as a sulfur source.

In this study, we investigate six different strategies for the production of [4Fe–4S] proteins using four different target proteins. Our study shows that selection of the appropriate strategy for cluster incorporation not only influences the cofactor content in proteins, but also significantly alters the expression levels of the recombinant genes. We use electronic absorption spectroscopy (EAS) and circular dichroism (CD) spectroscopy to assess the amount of Fe–S cluster bound to the proteins as well as the specificity of the method. In addition, we quantified

the Fe–S cluster content of the samples using inductively coupled plasma mass spectroscopy (ICP-MS).

We chose *E. coli* aconitase B (AcnB), *E. coli* 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (IspH), *E. coli* quinolinate synthase A (NadA) and *Bacillus thuringiensis* Thurincin B (ThnB) as target proteins, since all proteins bind [4Fe–4S] clusters that are coordinated by three cysteine residues. ThnB harbors a second [4Fe–4S] cluster that is coordinated by four cysteines. We have chosen three proteins that have previously been studied extensively and the properties of their cofactor is well established. In addition, we included a radical SAM enzyme in our study, where we have observed a dramatic impact of the expression system.

In the presence of iron, AcnB from *E. coli* functions as an enzyme in the essential citric acid cycle and the glyoxylate cycle, which assembles a [4Fe–4S] cluster. Under oxidative stress conditions and iron starvation, the Fe–S cluster disassembles and the enzymatically inactive apo protein functions as a nucleic acid binding protein, which binds to its own mRNA for positive regulation (Beinert & Kennedy, 1993; Tang & Guest, 1999; Williams *et al.*, 2002). The crystal structure of an inactive porcine heart mitochondrial aconitase hosting a [3Fe–4S] cluster reveals that the cluster is ligated by three cysteine residues (Robbins & Stout, 1989). The fourth coordination partner in the active form of the aconitase with a [4Fe–4S] cluster is proposed to be the citrate substrate (Beinert & Kennedy, 1993). IspH from *E. coli* produces the universal precursors of terpenes in the methylerythritol phosphate isoprenoid biosynthesis using its $2H^+$ /$2e^-$ reductase activity (Span *et al.*, 2012). The protein hosts a [4Fe–4S] cluster (Span *et al.*, 2012), but enzymatic activity has also been discovered in the presence of a [3Fe–4S] cluster (Gräwert *et al.*, 2009). In both cases, the cluster is coordinated by three cysteine residues, as known from the aconitase. NadA from *E. coli* plays an important role in the synthesis of nicotinamide adenine dinucleotide (NAD), an essential cofactor in numerous important redox reactions (Begley *et al.*, 2001). It harbors a [4Fe–4S] cluster coordinated by 3 cysteine residues (Cicchillo *et al.*, 2005). The protein ThnB from *B. thuringiensis* is a member of the so-called radical SAM enzymes, that are reported to be involved in the maturation process of sactipeptides (Wieckowski *et al.*, 2015). Sactipeptides represent a new class of natural products featuring promising antimicrobial activities. However, the role of these peptides within the producer strain remains unknown (Flühe & Marahiel, 2013). The first step in the biosynthesis of these sactipeptides is the formation of the characteristic cysteine-sulfur to α-carbon thioether linkage in the precursor peptide (Flühe & Marahiel, 2013; Lohans & Vederas, 2014). It has been shown that this reaction is catalyzed by enzymes that incorporate two Fe–S clusters and a SAM cofactor. These enzymes share a characteristic CX3CX2C motif, which is capable of binding one catalytic [4Fe–4S] cluster that is required for the reductive cleavage of SAM into methionine and a 5'-deoxyadenosyl (5'-dA) radical (Broderick *et al.*, 2014; Roach, 2011). The function of a second [4Fe–4S] cluster, hosted by radical SAM enzymes and studied for example in the radical SAM enzyme family members subtilisin A (AlbA) and sporulation killing factor (SkfB) is supposed to be electron transfer during the process of thioether bond formation (Flühe *et al.*, 2013, 2012; Flühe & Marahiel, 2013).

The proteins AcnB, IspH and NadA studied in that work were all taken from *E. coli*, whereas the radical SAM enzyme ThnB was taken from *B. thuringiensis*. While *E. coli* hosts both the ISC and the SUF machinery (Mettert & Kiley, 2014), in for Gram-positive bacteria such as *B. subtilis*, only the SUF machinery is thought to be the sole prevalent pathway (Selbach *et al.*, 2014). Therefore, we set special focus on how the co-expression with the ISC or SUF machinery from *E. coli* influences the expression of the radical SAM protein from *B. thuringiensis*.

Several spectroscopic techniques, including the optical absorption methods EAS and CD spectroscopy, can be used for the biophysical characterization of Fe–S cluster containing proteins. Among the broad variety of spectroscopic methods including electron paramagnetic resonance (EPR) (Jasniewski *et al.*, 2019), Mössbauer (Dunham *et al.*, 1994), X-ray absorption (Ward *et al.*, 2014) or resonance Raman spectroscopy (Spiro & Czernuszewicz, 1995), EAS and CD spectroscopy are the most widespread techniques that require only µM protein concentrations and small volumes of the samples.

Fe–S clusters usually show broad but characteristic absorption bands in the EAS These characteristic features can be used to roughly estimate the type of Fe-S cluster present in the protein. Generally spoken, proteins containing a [2Fe–2S] show more complex absorption bands between 410 nm and 430 nm, in the range of 470 nm, as well as a relatively broad maximum between 550 nm and 600 nm, while proteins containing a [4Fe–4S] cluster show a characteristic peak between 400 nm and 420 nm (Freibert *et al.*, 2018).

8.4 Material and Methods

8.4.1 Gene expression and protein isolation

For gene expression in *E. coli* BL21(DE3) overnight starting cultures containing one of the plasmids hosting the genes *acnB*, *ispH*, *nadA* or *thnB* were used to inoculate 2x Yeast extract (2YT) medium at 1% (v/v). 2YT medium was supplemented with a final concentration of 100 µg/ml ampicillin and the cells were cultivated under shaking at 37°C and 160 rpm until the optical density measured at 600 nm (OD_{600}) reached 2. Gene expression was induced by adding 0.5 mM Isopropyl-β-D-thiogalactopyranosid (IPTG). For the gene expression supported by the operons of the Fe–S cluster assembly machineries, overnight starting cultures of *E. coli* BL21(DE3) Δ*iscR* (Akhtar & Jones, 2008) containing the plasmids encoding for the target protein as well as *E. coli* BL21(DE3) containing one plasmid encoding for the target protein as well as a second plasmid pACYC*iscS-fdx* (pISC) (Gräwert *et al.*, 2004) or pACYC-Duet-1-*suf* (pSUF) (Hänzelmann *et al.*, 2004) were used to inoculate 2YT medium at 1% (v/v). TB medium was supplemented with kanamycin (50 µg/ml), ampicillin (100 µg/ml) or chloramphenicol (25 µg/ml) as needed, as well as ferric ammonium citrate (2 mM final concentration). Cells were cultivated aerobically at 37°C and 160 rpm until the OD_{600} reached 2. Subsequently, the gene expression was induced using 0.5 mM IPTG. To facilitate Fe–S cluster assembly 2 mM L-cysteine were added. All cultures were incubated at 25°C and 140 rpm for 20 h following induction. Cells were harvested by centrifugation for 10 min at 6,000 x *g* and 4°C. For cell lysis

cells were resuspended in 50 mM Tris pH 8.0, 150 mM NaCl containing EDTA-free cOmplete^{TM} Protease Inhibitor Cocktail Tablets (Roche, Basel, Switzerland) were added according to the manufacturer's instructions. After stirring at room temperature (RT) for 20 min under anaerobic conditions, the samples were covered with argon to maintain anaerobic conditions and the cell suspension was sonicated for 20 min with an amplitude of 60% and a pulse of 1 s every 3 s using VS70/T sonotrode (Bandelin electronic, Berlin, Germany). Argon covered lysates were clarified by centrifugation at 40,000 x *g*.

8.4.2 Anaerobic protein purification

Protein purification and all following procedures were carried out under strictly anaerobic condition in an anaerobic chamber with $<$ 2 ppm O_2. Isolation and purification of the protein was performed using immobilized metal affinity chromatography. Therefore, protein lysates were applied to a Protino Ni-NTA Agarose column (Macherey-Nagel, Düren, Germany) that was equilibrated in wash buffer (50 mM Tris pH 8.0, 150 mMNaCl) using an ÄKTA start system (GE Healthcare, Little Chalfont, UK). The column was then washed with wash buffer containing 50 mM imidazole to remove unspecifically bound proteins and the target protein was eluted with 250 mM imidazole. Fractions containing the target protein were pooled, dialyzed against wash buffer, and their purity was analyzed by SDS-PAGE. Protein concentrations of the protein samples were determined using the Bradford method (Bradford, 1976).

8.4.3 Chemical reconstitution

For the first step of the chemical reconstitution reaction, the reduction of the as-isolated protein with dithiothreitol (DTT), 100 µM protein solution were incubated with 1 mM DTT for 1 h on ice and then supplemented with 600 µM ferric ammonium citrate (FAC) and 600 µM Li_2S for the proteins with one [4Fe–4S] cluster. In case of ThnB, which contains two [4Fe–4S] clusters 800 µM FAC and 800 µM Li_2S were added. After an incubation time of 1 h the protein was applied to a PD-10 Desalting Column (GE Healtthcare, Little Chalfont, UK) that was equilibrated with 50 mM HEPES, pH 7.5; 100 mM KCl to separate the protein from excess iron and sulfide. Finally, the volume of the protein solution was adjusted to 500 µl.

8.4.4 Semi-enzymatic reconstitution

For the semi-enzymatic reconstitution reaction 100 µM protein solution were incubated on ice with 1 mM DTT for 1 h. Then 20 µM *E. coli* cysteine desulfurase IscS, 5 mM L-cysteine, and 600 µM FAC were added to the protein followed by an incubation of 1 h at RT. Subsequently, excess reagents were removed using a PD-10 Desalting Column (GE Healtthcare, Little Chalfont, UK) equilibrated with 50 mM HEPES, pH .5; 100 mM KCl. Finally, the volume of the protein solution was adjusted to 500 µl.

8.4.5 Determination of metal content

To obtain the metal-to-monomer ratio of the isolated samples, we determined the protein concentration using the Bradford method (Bradford, 1976) and measured the iron content using inductively coupled plasma mass spectroscopy. Samples were prepared by precipitating 10 µM of freshly, anaerobically isolated protein with 3% trace-metal grade nitric acid. Precipitate was removed by centrifugation for 30 min at 15,000 × *g* and the sample was then transferred to a metal-free centrifugation tube (VWR, Radnor, PA, USA). Measurements of the iron content were performed using an Agilent 7700 ICP-MS (Agilent Technologies, Waldbronn, Germany) in the Biocenter MS-Platform at the University of Cologne. The measurements were done by strictly following the manufacturer's instructions using He in the collision cell mode to minimize spectral interferences. Measurements were performed in technical triplicates and the presented data has an r^2 value of 0.999.

8.4.6 Electronic absorption spectroscopy (EAS)

Electronic absorption spectra were collected using a Cary-60 spectrophotometer (Agilent Technologies, Ratingen, Germany) with 1 nm bandwidth, a scanning speed of 120 nm/min and a 1-cm-path-length quartz cuvette at RT.

8.4.7 Circular dichroism (CD) spectroscopy

CD spectra were collected using a Jasco J-815 Circular dichroism spectrometer (Jasco Germany, Pfungstadt, Germany) with a scanning speed of 100 nm/min and a bandwidth of 5 nm in a 2 mm path length quartz cuvette. An average of 20 scans was collected at 20 °C with 1 nm resolution. The spectra were processed with the Spectra Analysis software using the Adaptive-Smoothing function with the following parameters: convolution width 15, noise deviations 1. CD spectra of HEPES and Tris buffer (Figure S8.3) were subtracted from protein spectra prior to the calculation of the molar ellipticity and plotting the data. Samples for CD spectroscopy were in a concentration range between 150 and 300 µM.

8.5 Results and discussion

The production of correctly assembled Fe–S cluster proteins is a major challenge in the field. We have compared six different strategies for protein production in this work: (i) expression in BL21(DE3); (ii) expression in BL21(DE3) and chemical reconstitution; (iii) expression in BL21(DE3) and semi-enzymatic reconstitution; (iv) co-expression in BL21(DE3) with pISC; (v) co-expression in BL21(DE3) with pSUF; and (vi) expression in BL21(DE3) Δ*iscR*. Hereby, gene expression was always performed under aerobic conditions, cell lysis and all further steps under anaerobic conditions. The target proteins *E. coli* AcnB, *E. coli* IspH, and *E. coli* NadA contain one [4Fe–4S] cluster ligated by three cysteine residues. We also studied cluster incorporation into the radical SAM enzyme *B. thuringiensis* ThnB that contains two Fe–S clusters, one of

them with an open coordination site for the SAM cofactor. We studied the impact of the protocol on expression levels, Fe–S cluster content, and specificity of cluster incorporation.

8.5.1 Expression levels

We compared the protein levels obtained in different cell lines, including XL1-Blue, BL21(DE3), and BL21 (DE3) Δ*iscR* utilizing SDS-PAGE and Western blot analysis (Figure 8.3). Common BL21(DE3) cells harbor both, the ISC and SUF Fe–S pathways (Barras *et al.*, 2005; Johnson *et al.*, 2005). However, when expressing genes to an abnormally high level, the endogenous enzymes involved in biosynthesis are not capable of producing and transferring a sufficient amount of Fe–S clusters. BL21(DE3) cells containing an additional plasmid encoding for the Fe–S assembly machineries ISC or SUF produce an increased level of ISC or SUF proteins, respectively, which may result in increased levels of Fe-S cluster content in the target protein. XL1-Blue cells are routinely utilized for cloning, yet they have been successfully applied for co-expression of IspH with the ISC proteins (Gräwert *et al.*, 2004), therefore, we included expression tests with XL1-Blue cells harboring pISC or pSUF in our study. Another approach to increase the level of Fe–S biosynthesis proteins is to delete the regulator of the ISC pathway, IscR, which is responsible for the attenuation of the ISC proteins. The *E. coli* BL21(DE3) Δ*iscR* cell strain shows an upregulation of *isc* operon expression, leading to higher levels of the ISC proteins and the ability to maturate abnormally high levels of Fe–S proteins. Notably, recent studies indicate a more general role of IscR in the regulation of Fe–S cluster biogenesis, therefore the deletion of IscR may also influence the SUF protein levels (Giel *et al.*, 2006).

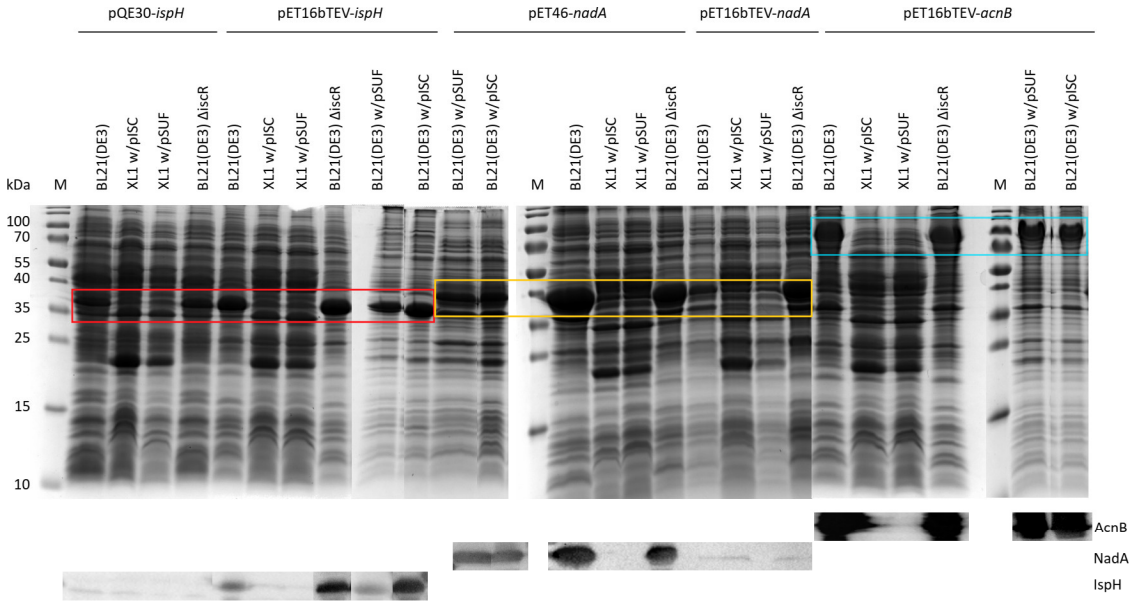


Figure 8.3: Analysis of *ispH*, *nadA* and *acnB* expression in different systems. The cell lysate was analyzed by SDS-PAGE on a 15% Tris-Glycine gel and proteins were visualized by Coomassie staining (top) and the corresponding a Western blot analysis using a conjugated anti-Histidine antibody (bottom).

The constructs for IspH, NadA, and AcnB were designed to result in proteins with an N-terminal Histidine affinity tag followed by the recognition site for the TEV protease to allow for cleavage of the Histidine affinity tag. In addition, we have also investigated the expression of the constructs pQE30-*ispH* (Gräwert *et al.*, 2004) with a Histidine affinity tag and no protease recognition site as well as pET46-*nadA* with a Histidine affinity tag and an enterokinase cleavage site. These constructs were previously used for protein production, however, the Histidine affinity tag enhances the purity after the first purification step and is, thus, more suitable for our study. We expressed the five different constructs in six different cell lines, except pQE30-*ispH*, which did not co-transform with neither the pISC, nor the pSUF plasmid.

Western blot analysis indicated that pQE30-*ispH* does not express in either cell strain, which is surprising, as previous studies report high protein yields for pQE30-*ispH* using the XL1-Blue cells harboring the pISC plasmid (Gräwert *et al.*, 2004). The pISC plasmid used for protein biosynthesis in this study is different from the one previously used, since it was discovered that the originally used pISC plasmid contained three point mutation in the *iscS* and *iscU* genes (Span, 2012). Interestingly, correction of these mutations resulted in a dramatic reduction of *ispH* expression in *E. coli* XL1-Blue cells. SDS-PAGE analysis shows protein bands for expression in BL21(DE3) and BL21(DE3) Δ*iscR*, suggesting that IspH protein is synthesized. The Histidine affinity tag of this construct is shorter, which could influence the Western blot analysis. Another possibility is that the Histidine affinity tag is not accessible, possibly due to dimer formation, which was also observed in the crystal structure (Gräwert *et al.*, 2009).

We also observe no detectable expression of either protein in the XL1-Blue cell line independent of the co-transformed plasmid. Furthermore, we obtained high amounts of protein for all other samples and Western blot analysis confirmed that the proteins carry a Histidine affinity tag and are most certainly our target proteins. Notably, the intensity of the signal on the Western blot does not necessary correspond to the amount of protein, since there are several factors that influence staining. Thus, we base our decision of which expression system to select for protein production not fully on SDS-PAGE or Western blot analysis, but on the resulting protein yield after the initial purification step.

The protein AcnB from *E. coli* is a [4Fe–4S] cluster protein that plays a key role in the citric acid cycle under conditions in which sufficient iron is present, whereas in the absence of iron, the protein loses its enzymatic activity and functions as a nucleic acid binding protein, stabilizing its own mRNA (Beinert & Kennedy, 1993; Tang & Guest, 1999; Williams *et al.*, 2002). Western blot analysis of cell lysates after the expression of the protein in *E. coli* BL21(DE3) provide evidence that additional co-expression of the *acnB* gene with genes encoding for the SUF or ISC machinery do not have an influence on the protein levels of AcnB. The same is true for the expression of *acnB* in BL21(DE3) Δ*iscR* (Figure 8.3). In all cases, expression was directed from a pET16bTEV expression plasmid resulting in a target protein with an N-terminal Histidine affinity tag.

Purification of AcnB has only slightly been influenced by the choice of the expression system. Isolation of AcnB from BL21(DE3) with the additional plasmid pSUF and from BL21(DE3) Δ*iscR* resulted in 500 µl protein solution with a concentration of approximately 300 µM (75 nmol, 7 mg)

obtained from 1 l expression culture using immobilized metal affinity chromatography (IMAC). Isolation of AcnB from BL21(DE3) with and without the additional plasmid pISC resulted in 500 µl protein solution with a slightly lower concentration of 250 µM (63 nmol, 6 mg) (Table 8.1, Figure S8.1).

The protein IspH from *E. coli* is an enzyme catalyzing the conversion of 1-hydroxy-2-methyl-2-(E)-butenyl diphosphate into isopentenyl diphosphate and dimethylallyldiphosphate in the nonmevalonate isoprenoid biosynthesis pathway. The protein hosts a [4Fe–4S] cluster (Span *et al.*, 2012), but enzymatic activity can also be discovered in a variant hosting a [3Fe–4S] cluster (Gräwert *et al.*, 2009). In this study, we expressed the *ispH* gene fused to an N-terminal Histidine affinity tag from a pET16bTEV expression vector in BL21(DE3) and BL21(DE3) Δ*iscR* cells in the presence and absence of the plasmids pSUF and pISC and compared our results with previously reported spectroscopic data on IspH. Previous studies on IspH used either the *E. coli* strain M15[pREP4] or XL1-Blue with the *ispH* gene hosted in a pQE30 expression vector (Gräwert *et al.*, 2004; Wolff *et al.*, 2003). In the latter study using *E. coli* XL1-Blue cells for gene expression, the pISC plasmid was co-expressed with the pQE30 expression vector hosting the target protein, however the growth medium was not supplemented with additional iron or sulfur sources. Isolation of the protein was carried out under strictly anaerobic conditions resulting in a protein fraction that shows a significant peak at 410 nm with a molar extinction coefficient of 11.9 mM^{-1} cm^{-1} for a monomeric protein structure (Gräwert *et al.*, 2004). Purification of IspH isolated from *E. coli* M15[pREP4] was performed aerobically and the Fe–S cluster was then anaerobically reconstituted using Na_2S, $FeCl_3$ and DTT. Here, a molar extinction coefficient at 410 nm with a value of 18.75 mM^{-1} cm^{-1} has been observed for a dimeric protein structure (Wolff *et al.*, 2003). In this study, we explored that the co-expression of the *ispH* gene with the plasmids pSUF and pISC had a positive effect on the protein level in BL21(DE3) (Figure 8.3). For expression in BL21(DE3) in the absence of pSUF and pISC as well as the expression in BL21(DE3) Δ*iscR* reduced protein levels have been observed. Additionally, we tested the expression of *ispH* directed from the expression vector pET16bTEV in *E. coli* XL1-Blue with the additional plasmids pISC and pSUF respectively, but Western blot analysis revealed that the protein production using these expression systems was not successful. The same is true for the expression of *ispH* directed from the expression vector pQE30 in BL21(DE3) with and without the plasmids pSUF and pISC as well as in BL21(DE3) Δ*iscR* (Figure 8.3). The protein yields obtained after the purification process are in agreement with the observations for the expression levels of *ispH*. After the IspH isolation from BL21(DE3) with pSUF or pISC 500 µl with a concentration of 150 µM (38 nmol, 1.4 mg) were obtained, while the concentrations were higher in samples obtained from the expression in BL21(DE3) and from BL21(DE3) Δ*iscR*(500 µl with a concentration of 250 µl, i.e. 63 nmol, 2.3 mg) (Table 8.2 and Figure S8.1).

Quinolinic acid is an intermediate in the synthesis of NAD that is an important cofactor in numerous redox reactions (Begley *et al.*, 2001). *E. coli* and most other prokaryotes generate quinolinic acid via a condensation reaction between dihydroxyacetone phosphate and iminosuccinate. For this purpose, two proteins are required: the quinolinate synthetase, NadA, and the aspartate oxidase, NadB (Begley *et al.*, 2001). In this study we are focusing on the

[4Fe–4S] cluster protein NadA. In a previous study, the presence and nature of the Fe–S cluster has been explored after the expression of *nadA* on the expression plasmid pET-28A in *E. coli* with an additional arabinose-inducible plasmid pDB1282 harboring the genes encoding for the proteins of the ISC machinery (Cicchillo *et al.*, 2005). Spectroscopic analysis to verify the presence of the [4Fe–4S] cluster was carried out either on the as-isolated protein using EAS and EPR spectroscopy or with a chemically reconstituted sample using Mössbauer spectroscopy. In the present study, we tested two constructs for the expression of the gene *nadA*: In analogy to our studies on AcnB, IspH and ThnB we generated a construct in which the gene was incorporated into the expression plasmid pET16bTEV. However, using Western blot analysis, we could not detect any recombinant protein production neither in *E. coli* BL21(DE3), BL21(DE3) Δ*iscR*, nor XL1-Blue (Figure 8.3). Interestingly, the NadA protein levels could be significantly increased using a pET46-*nadA* construct that also allows for the expression of a fusion protein with an N-terminal Histidine affinity tag. Here, the highest expression levels could be detected in BL21(DE3) in the presence of pSUF in BL21 Δ*iscR*, while the expression levels were moderate when the genes encoding for the ISC machinery were co-expressed in BL21(DE3) as well as in BL21(DE3) without co-expression. Purification via IMAC resulted in highly pure protein samples. In a reduced volume of 500 µl after the isolation process a concentration of 250 µM (63 nmol, 2.5 mg) was detected for the samples obtained from BL21(DE3) Δ*iscR*, and BL21(DE3) with the additional plasmid pSUF, while a significantly lower concentration of 150 µM (38 nmol, 1.5 mg) has been detected in the sample obtained from BL21(DE3) without the co-expression of Fe–S cluster assembly machineries and with additional co-expression of pISC (Table FeSTable3 and Figure S8.1).

8.5.2 Analysis of iron content

The first protein that we investigated in terms of Fe–S cluster maturation is *E. coli* AcnB, which contains an [3Fe–4S] cluster in the inactive form that can be activated by conversion to the [4Fe–4S] form. Our spectroscopic analysis of AcnB produced using the previously mentioned strategies reveals significant differences in the cluster occupancy of the different samples. Expression in BL21(DE3) cells in the absence or presence of an additional plasmid encoding for the genes for the ISC or SUF machinery led to the isolation of the apo protein without significant amounts of iron as indicated by an electronic spectra (Figure 8.4A). The absorption bands at 325 nm and in the range of 400–420 nm with low intensities and molar extinction coefficients below 3 $mM^{-1}cm^{-1}$ (Table 8.1) are a strong indicators for low Fe–S cluster content. Protein obtained from *E. coli* BL21(DE3) Δ*iscR* cells shows a molar extinction coefficient of 7.6 $mM^{-1}cm^{-1}$ at 420 nm, suggesting an improved incorporation of Fe–S clusters into the protein. According to the literature, values of 10 $mM^{-1}cm^{-1}$ are common for fully loaded Fe–S proteins (Jacquot, 2017), leading to the conclusion that the majority of AcnB protein obtained from *E. coli* BL21(DE3) Δ*iscR* cells contains an Fe–S cluster. We were able to increase the amount of Fe–S cluster in protein isolated in the apo form from *E. coli* BL21(DE3) cells using chemical or semi-enzymatic reconstitution. Both protocols led to samples with a molar extinction coefficient

of > 18 $mM^{-1}cm^{-1}$ for the major absorption band. Surprisingly, the molar extinction coefficients of the major signal are significantly higher than previously reported for *E. coli* AcnB obtained by expression in the *E. coli* strain W3110 and reconstituted with $(NH_4)_2Fe(SO_4)$, Na_2S, and DTT after (8 $mM^{-1}cm^{-1}$ to 10 $mM^{-1}cm^{-1}$) (Jordan *et al.*, 1999). Interestingly, we also observe a shift of the most prominent signal from 420 nm to 400 nm in the protein sample obtained by chemical reconstitution. Furthermore, we observe a large absorption feature at wavelengths higher than 650 nm, indicating impurities that are not bound to the protein but cannot be separated using a desalting column.

To investigate the sample in more detail, we measured CD spectra in the visible region (Stephens *et al.*, 1978). The chiral nature of Fe–S clusters inside proteins makes Fe–S proteins amenable for CD spectroscopy. While all Fe–S proteins exhibit characteristics CD signals, the CD spectra of [4Fe–4S] cluster containing proteins are less defined compared to [2Fe–2S] proteins and often of low intensity (Freibert *et al.*, 2018). Low intensity or absence of a signal in the CD spectrum is an indicator of either low concentration of a species or a highly symmetric environment. The advantage of CD spectroscopy over EAS is that the absorption bands split up to maxima and minima, therefore, facilitating the detection of subtle changes, and bands deriving from absorbing impurities do not contribute to the spectrum. The latter includes metal ions in solution, as well as metal complexes formed with buffer substances or other reagents, such as reducing substances.

The electronic spectra of the AcnB samples show bands at 325 nm and in the region of 400–420 nm, which are characteristic for protein-bound [4Fe–4S] clusters, however, all absorption features are quite broad with a high background. The bands in the CD spectra show minima at 325 nm and 400 nm as well as maxima at 310 nm, 360 nm, 450 nm, 590 nm, and 640 nm. When taking a closer look at the 400-450 nm range of the CD spectrum it becomes evident that the absorption bands for both reconstituted samples are at 410 nm, whereas the predominant band of the samples from the BL21(DE3) Δ*iscR* cells is located at 450 nm. This shift most likely indicates differences in the oxidation states of the [4Fe–4S] cluster, similar to the observations for the [4Fe–4S] ferredoxin from *Bacillus stearothermophilus* (Mullinger *et al.*, 1975). The electronic spectrum of the oxidized cluster in ferredoxin resembles the spectrum of AcnB protein. The maximum at 410 nm in the ferredoxin spectrum was assigned to the oxidized form of the [4Fe–4S] cluster, whereas the maximum at 450 nm corresponds to the reduced form. These findings suggest that the [4Fe–4S] cluster is in a different oxidation state when matured *in vitro*, which has a dramatic impact on the characterization of proteins by spectroscopy.

Absorption features in CD spectra can be more easily distinguished from each other than in electronic spectra, especially the features beyond 550 nm. The baseline was corrected for the buffer used for the samples, however, the baseline of some of the AcnB samples is unexpectedly high above 600 nm. In addition, the noise levels increase significantly beyond 600 nm. Therefore, it is difficult to compare the molar ellipticity values analogous to molar extinction coefficients. Despite the shifted baselines, the CD spectra indicate that the samples obtained by chemical and semi-enzymatic reconstitution do not contain significantly more Fe–S clusters than the *in vivo* maturated samples isolated from *E. coli* BL21(DE3) Δ*iscR* cells. In addition,

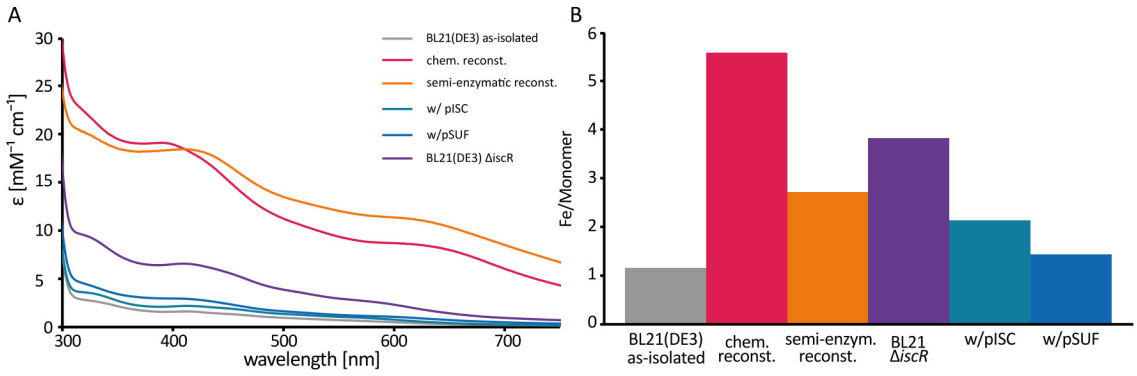


Figure 8.4: Analysis of Fe-S content of AcnB protein obtained by different strategies. (A) Electronic absorption spectra reveal the Fe–S content of the AcnB protein. Corresponding CD spectra are shown in Figure S8.2A. (B) Iron content per monomer of the samples in (A) determined by ICP-MS measurements. AcnB protein as-isolated from *E. coli* BL21(DE3) is shown in grey, samples obtained by subsequent chemical or semi-enzymatic reconstitution are displayed in pink or orange, respectively. Samples co-expressed with either pISC or pSUF are shown in green or blue, respectively, and protein isolated from BL21(DE3) Δ*iscR* is presented in purple.

the broad features at 590 nm and 640 nm support the presence of impurities consisting most likely of Fe–S aggregates.

To determine the amount of iron bound to the AcnB protein, we adjusted the protein concentration of all samples and measured the iron content utilizing ICP-MS. Hence, we were able to determine the iron per monomer ratio for AcnB obtained by different protocols. Our results show that the samples isolated form *E. coli* BL21(DE3) contain approximately one iron per monomer, while the samples isolated from *E. coli* BL21(DE3) co-expressed with either the pISC or pSUF plasmid contain up to two iron per monomer (Figure 8.4B). AcnB as isolated from *E. coli* BL21(DE3) Δ*iscR* binds close to four iron per monomer, pointing to a fully loaded Fe–S protein. In line with the spectroscopic results, the protein obtained by chemical reconstitution binds more than five irons per monomer, suggesting adventitious metal binding. Surprisingly, the samples obtained by semi-enzymatic reconstitution show an iron content lower than 3, which is significantly lower than expected based on the high absorption at 420 nm in the EAS. We speculate that the discrepancy between this ICP-MS result and the EAS may be caused by inaccuracy during determining the protein concentration, which has previously been reported to be the most error-prone step (Galle *et al.*, 2018). In summary, our results show that reconstitution of Fe–S cluster using the chemical or semi-enzymatic approach lead to high amount of Fe–S cluster content, however, these samples contain also a larger amount of impurities, presumably consisting of Fe–S aggregates. Assembly of Fe–S clusters *in vivo* is more specific and efficient, as the reconstitution step as well as subsequent purification steps are not required. The best result for AcnB protein was achieved by isolating the protein from BL21(DE3) Δ*iscR*.

As previously mentioned, the ISC machinery is responsible for the production of Fe–S clusters in E. coli under normal conditions, while the SUF machinery fulfils the same function under stress conditions. Under oxidative stress or iron depletion condition, AcnB enhances its own

Table 8.1: Overview on different expression strategies and reconstitution methods for the Fe–S cluster protein AcnB.

Expression strategy	Significant band [nm]	Extinction coefficient [$mM^{-1}cm^{-1}$]	Fe/monomer	Yield [nmol] per l cell culture	Purity [%]
BL21(DE3)	420	1.6	1.1	60	>90
BL21(DE3) chemic. reconst.	400	18.9	5.5	53	–
BL21(DE3) semi-enzym. reconst.	420	18.3	2.8	55	–
BL21 Δ*iscR*	420	7.6	3.9	76	>90
BL21(DE3) w/pISC	420	2.0	2.1	63	>90
BL21(DE3) w/pSUF	420	2.9	1.2	75	>90

production by binding to its own mRNA and inducing a positive regulation mechanism (Tang & Guest, 1999). AcnB plays a key role in the citric acid cycle as well as the glyoxylate cycle. Nevertheless, deletion of this important protein in *E. coli* can be compensated by a second aconitase, AcnA (Gruer *et al.*, 1997), although the sequence identity between both proteins is only 17% (Bradbury *et al.*, 1996). Whereas AcnB is adapted for the major catabolic function (Jordan *et al.*, 1999), AcnA is responsible for maintenance of the citric acid cycle activity during oxidative stress (Cunningham *et al.*, 1997). A previous study has shown that the $SufBC_2D$ complex is capable of inserting an Fe–S cluster into AcnB, however, we speculate that the interaction of AcnB with the SUF machinery is not biological relevant, since the function of AcnB is compensated by AcnA during oxidative stress conditions. Therefore, the ISC machinery is most likely responsible for the maturation of AcnB inside the cell, which is in line with our observation that the *E. coli* BL21(DE3) Δ*iscR* cells lead to the highest yield of Fe–S containing AcnB protein.

Surprisingly, the co-expression of the *acnB* gene with the additional plasmid pISC in *E. coli* BL21(DE3) cells leads to incomplete cluster assembly in the as-isolated protein. This effect may be a result of additional stress for the host cell caused by the expression of a second plasmid.

The second protein that we investigated is the [4Fe–4S] cluster containing protein IspH. The EAS of the protein samples obtained by different methods are dominated by an absorption band at 415 nm (Figure 8.5A), which is in agreement with previous studies (Gräwert *et al.*, 2004; Wolff *et al.*, 2003). In general, we observe similar trends as for AcnB, even though the differences between as-isolated a reconstituted samples are not as pronounced as for AcnB. IspH protein isolated from *E. coli* BL21(DE3) cells that was then maturated using the semi-

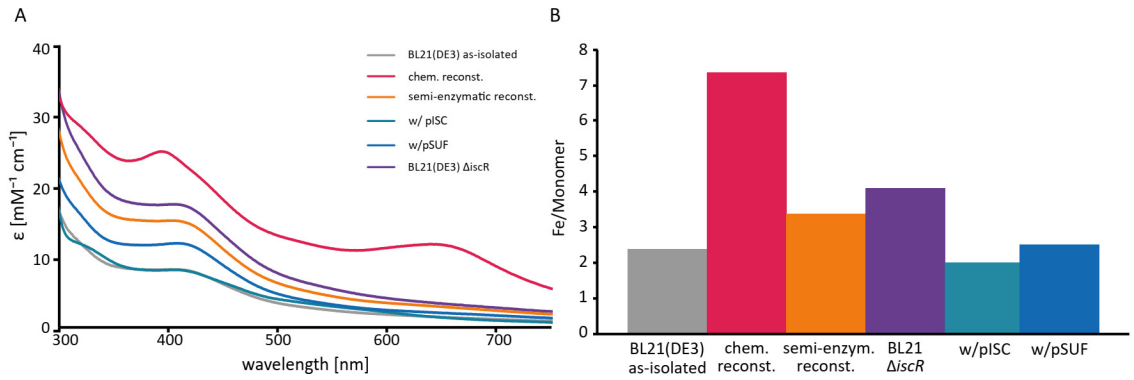


Figure 8.5: Analysis of Fe-S content of IspH protein obtained by different strategies. (A) EAS of IspH protein from different cell strains and after Fe-S reconstitution reactions. Corresponding CD spectra are shown in Figure S8.2B. (B) Iron content per monomer of the samples in (A) determined by ICP-MS measurements. IspH protein as-isolated from *E. coli* BL21(DE3) is shown in grey, samples obtained by subsequent chemical or semi-enzymatic reconstitution are displayed in pink or orange, respectively. Samples co-expressed with either pISC or pSUF are shown in green or blue, respectively, and protein isolated from BL21(DE3) Δ*iscR* is presented in purple.

enzymatic method shows a lower Fe–S content compared to AcnB, while the co-expression in presence of the pSUF plasmid yields a larger portion of Fe–S containing protein. As for AcnB, we observe a shift of the band at 410 nm to 400 nm for the chemically reconstituted IspH sample in addition to a broad feature at 650 nm. The large feature beyond 600 nm in the electronic spectrum of chemically reconstituted IspH indicates again the presence of large amount of Fe–S aggregates, which is also supported by the CD spectra (Figure S8.2B). In line with these observations, ICP-MS measurements show that this protein sample contains an average of seven iron per monomer (Figure 8.5B). Overall, the molar extinction coefficients of all IspH samples are higher compared to AcnB (Table 8.2), but the iron per monomer ratios follow the same trend as for AcnB (Figure 8.5B). The absorption feature at 415 nm in the CD spectrum shown for the protein samples obtained by chemical reconstitution and as-isolated from *E. coli* BL21(DE3) Δ*iscR* cells (Figure S8.2B) indicates that the majority of the cluster is in the oxidized form.

The amount of iron per monomer is similar for the protein isolated from *E. coli* BL21(DE3) in absence or presence of an additional plasmid encoding for the ISC or SUF machinery. While the yield of IspH protein obtained in BL21(DE3) cells is similar to AcnB, 61 nmol/l and 60 nmol/l cell culture, respectively, the yields of IspH protein obtained from co-expression with the plasmids encoding for the biosynthesis pipelines is reduced by approximately 50% (Tables 8.1 and 8.2). Semi-enzymatic reconstitution results in an iron/monomer ratio of 3, which is consistent with the molar extinction coefficient. The highest Fe-S cluster content of as-isolated protein was again achieved in *E. coli* BL21(DE3) Δ*iscR* cells, with a molar extinction coefficient of 15 $mM^{-1}cm^{-1}$ at 415 nm and 4 Fe/monomer.

Several expression strategies have been reported previously for IspH from *E. coli* as well as *Aquifex aeolicus*. For this purpose, expression was performed in *E. coli* M15, *E. coli* TOP10,

and *E. coli* BL21(DE3) cells (Rekittke *et al.*, 2008; Wang *et al.*, 2010b; Wolff *et al.*, 2003). These studies performed by three different groups all reported very low Fe-S content, with very low A_{280}/A_{410} ratios, which they increased by chemical reconstitution. A different study reported the co-expression of *E. coli* IspH protein with the pISC plasmid in *E. coli* XL1-Blue cells (Gräwert *et al.*, 2004). As mentioned previously, these experiments were carried out using a pISC plasmid with three point mutation that were only discovered afterwards (Span, 2012). We have tried to reproduce these results using a pISC plasmid that was corrected using site-directed mutagenesis, however, no expression was detectable by Western Blot analysis. IspH protein obtained by chemical reconstitution results in a molar extinction coefficient of 2 $mM^{-1}cm^{-1}$ at 410 nm and 7 Fe/monomer. The broad absorption feature at 650 nm and the high iron content suggest that significant amounts of impurities have been formed during the chemical reconstitution reaction. To investigate if the excess iron is specifically bound to the protein or if the impurities consist of Fe–S aggregates, we performed size exclusion chromatography (SEC) on this sample (Figure S8.4). The electronic spectrum of the IspH protein after the SEC reveals that the absorption feature at 650 nm is completely removed by this additional purification step. The A_{650}/A_{410} ratio is reduced from 0.50 to 0.16 by SEC, showing that these impurities consist of Fe–S aggregates, which can be separated from the protein. The additional purification step improves the purity and homogeneiety of the protein sample significantly, however, it requires a fast protein liquid chromatography (FPLC) system in an anaerobic chamber. If this instrumental setup is available, then isolation of the mature Fe–S protein under anaerobic conditions is the more efficient method to obtain active protein, since chemical reconstitution followed by SEC are more time-consuming. Taken together, our results show that expression and maturation in *E. coli* BL21(DE3) Δ*iscR* cells is the most efficient strategy to obtain [4Fe–4S] IspH, as previously observed for AcnB.

The third [4Fe–4S] protein that was studied here is the quinolinate synthetase NadA. Our spectroscopic analysis of the proteins reveals that the trends observed for AcnB and IspH are not reproduced in these experiments. As previously observed for AcnB and IspH, protein maturated by chemical reconstitution using FAC and Li_2S in the presence of DTT yields a high absorption band at 415 nm and a broad feature beyond 600 nm (Figure 8.6A). In combination with the high iron content determined by ICP-MS (Figure 8.6B), the data indicate the formation of unspecific Fe–S aggregates, which is consistent with our previous observations. Unlike AcnB and IspH, the expression of *nadA* (vector pET22b+) in *E. coli BL21(DE3)* Δ*iscR* resulted in the isolation of protein with insufficient cluster loading as shown by the EAS and low Fe/monomer ratio. Expression in *E. coli* BL21(DE3) as well as *E. coli* BL21(DE3) Δ*iscR* result in molar extinction coefficients of 4-6 $mM^{-1}cm^{-1}$ for the band at 415 nm, which indicate that the majority of the protein is in the apo form (Figure 8.6A). NadA protein obtained by semi-enzymatic reconstitution or by co-expression with the *isc* operon yield slightly better results, but overall a low amount of Fe–S cluster containing protein (Table 8.3). The most promising result for NadA protein was obtained by co-expression with the pSUF plasmid in *E. coli* BL21(DE3) cells. The CD spectra of chemically reconstituted NadA and protein obtained by co-expression with the SUF plasmid (Figure S8.2C) show a maximum at 415 nm, supporting an [4Fe–4S] cluster

Table 8.2: Overview on different expression strategies and reconstitution methods for the Fe–S cluster protein IspH.

Expression strategy	Significant band [nm]	Extinction coefficient [$mM^{-1}cm^{-1}$]	Fe/monomer	Yield [nmol] per l cell culture	Purity [%]
BL21(DE3)	415	8.1	2.4	31	>95
BL21(DE3) chemic. reconst.	400	24.6	7.2	23	–
BL21(DE3) semi-enzym. reconst.	415	15.2	3.2	30	–
BL21 Δ*iscR*	415	17.2	4.0	63	>90
BL21(DE3) w/pISC	415	8.0	2.0	63	>95
BL21(DE3) w/pSUF	415	11.5	2.2	38	>95

in the oxidized form. The spectroscopic results are in good agreement with the Fe/monomer ration obtained by ICP-MS (Figure 8.6B). The Fe/monomer rations for proteins isolated from *E. coli* BL21(DE3) as well as *E. coli* BL21(DE3) Δ*iscR* are below 1, proteins obtained by semi-enzymatic reconstitution and co-expression with pISC yield a value of approximately 2, and co-expression with pSUF results in 4 Fe/monomer. As previously mentioned, chemical reconstitution leads to excess iron binding to NadA. These results are in good agreement with the presence of an [4Fe–4S] cluster in NadA protein and show that specific and efficient maturation of NadA was achieved using the co-expression strategy with the pSUF plasmid (Rousset *et al.*, 2008).

Our results on the biosynthesis and maturation of AcnB and IspH show that expression in the BL21(DE3) Δ*iscR* cells yields the highest Fe–S cluster content, as observed by EAS, without the binding of excess iron or formation of Fe–S aggregates, as deducted from the ICP-MS analysis. This lead to the initial conclusion that the deletion of the IscR regulator is a superior approach compared to introducing an additional plasmid into the cells, which may result in additional stress for the host cell. The maturation of NadA demonstrates that the co-expression with the SUF plasmid leads to increased Fe–S cluster incorporation. This raises the question whether NadA interacts specifically with the SUF machinery inside the cell. To our knowledge, the interaction of NadA with the SUF proteins have not yet been investigated in detail. Previous studies of *E. coli* NadA with an *E. coli* strains lacking the *iscS* gene have shown that NAD biosynthesis is impaired, leading to the conclusion that *iscS* is essential for the maturation of NadA (Ollagnier-de Choudens *et al.*, 2005). Apparently, *sufS* was not able to replace *iscS* in these experiments and maturate NadA protein. This hypothesis has not been further

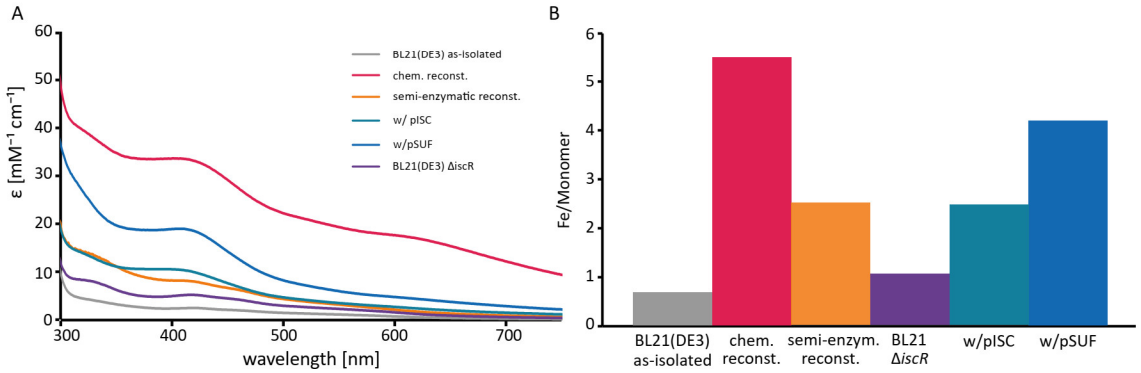


Figure 8.6: Analysis of Fe–S content of NadA protein obtained by different strategies. (A) EAS of NadA samples as-isolated from different cell strains or obtained by different reconstitution methods. Corresponding CD spectra are shown in Figure S8.2C. (B) Iron content per monomer of the samples in (A) determined by ICP-MS measurements. NadA protein as-isolated from *E. coli* BL21(DE3) is shown in grey, samples obtained by subsequent chemical or semi-enzymatic reconstitution are displayed in pink or orange, respectively. Samples co-expressed with either pISC or pSUF are shown in green or blue, respectively, and protein isolated from BL21(DE3) Δ*iscR* is presented in purple.

investigated on a molecular level. Controversially, another study reports a specific connection between NadA and a potential new system for Fe–S assembly in *E.coli* referred to as CSD (Bolstad *et al.*, 2010; Kurihara *et al.*, 2003; Loiseau *et al.*, 2005). The role of this system in Fe–S cluster biosynthesis has not been established yet and recent review articles on this topic only mention two Fe–S biosynthesis pathways in *E. coli* (Lill, 2009). The only evidence for the interaction of NadA and a member of the SUF machinery was provided in *Arabidopsis thaliana* (Narayana Murthy *et al.*, 2007). Hereby, a characterization of two novel chloroplast SufE-like proteins from *A. thaliana* was reported and the results revealed that the mature SufE3 contains two domains, one SufE-like and one with similarity to the bacterial quinolinate synthase, NadA. They showed that SufE3 displayed both SufE activity and quinolinate synthase activity, thus, they concluded that SufE3 is the NadA enzyme of *A. thaliana*. Our results on the maturation of *E. coli* NadA protein corroborate a close relation of NadA and the SUF machinery. These findings are particularly interesting, since in *E. coli* both systems are co-existing and supposedly responsible for the maturation of the same set of proteins under different conditions. While the ISC pathway is predominantly used for Fe–S biosynthesis under standard growth conditions, the SUF pathway is essential under stress conditions. As a consequence, both systems should be able to maturate the same proteins with a high efficiency. The different preferences of the proteins for one system versus the other may be related to the role of the protein in the cell, or possibly relevance of the protein for the survival of the organism.

Taken together our results from the systematic analysis of six different strategies for the maturation of three different [4Fe–4S] proteins, we observe that *in vivo* maturation results in more specific and effective maturation of [4Fe–4S] proteins. For each protein one of the *in vivo* maturation strategies yielded fully assembled Fe–S clusters without excess iron species. However, the suitable biosynthesis method has to be determined for each protein. In all three cases,

Table 8.3: Overview on different expression strategies and reconstitution methods for the Fe–S cluster protein NadA.

Expression strategy	Significant band [nm]	Extinction coefficient [$mM^{-1}cm^{-1}$]	Fe/monomer	Yield [nmol] per l cell culture	Purity [%]
BL21(DE3)	415	3.5	0.7	19	>95
BL21(DE3) chemic. reconst.	415	36.4	5.6	17	–
BL21(DE3) semi-enzym. reconst.	415	10.1	2.7	18	–
BL21 Δ*iscR*	415	5.5	1.0	63	>95
BL21(DE3) w/pISC	415	12.2	2.4	38	>95
BL21(DE3) w/pSUF	415	20.2	4.2	63	>95

chemical reconstitution leads to high Fe–S cluster contents in the protein, but also to the formation of high amounts of Fe–S aggregates, which require an additional SEC purification step to separate them from the protein sample. In one case, we have also observed that maturation by reconstitution results in a different oxidation state of the [4Fe–4S] cluster compared to *in vivo* maturation. CD spectroscopy has been a very useful tool to obtain more detailed information about the Fe–S cluster and provided insights on the oxidation states of the cluster. So far, we have investigated three proteins, AcnB, IspH, and NadA, which contain only one [4Fe–4S] cluster and have previously been characterized. Next, we aimed at transferring the new insights on Fe–S cluster maturation on the radical SAM enzyme ThnB, which has not yet been maturated *in vivo*.

8.5.3 Maturation of the radical SAM enzyme ThnB

The radical SAM enzyme ThnB from *B. thuringiensis* has been reported to host two [4Fe–4S] clusters and is required for thioether bond formation during the maturation of the sactipeptide ThnH (Wieckowski *et al.*, 2015). Previous spectroscopic analyses of the enzyme were performed on samples obtained from recombinant production in *E. coli* BL21(DE3) directed from the expression plasmid pET28a(+) resulting in a fusion protein with an N-terminal Histidine affinity tag. Using this construct, ThnB could be isolated with high purity by IMAC . EAS and EPR spectroscopy on the chemically reconstituted sample provide evidence for the presence of at least one [4Fe–4S] cluster. Characteristic signals have been detected at 413 nm with a molar extinction coefficient of approximately 30 $mM^{-1}cm^{-1}$ and at 325 nm with a molar extinc-

tion coefficient of approximately 40 $mM^{-1}cm^{-1}$. However, determination of the iron and metal content indicate the presence of more than one [4Fe–4S] cluster (Wieckowski *et al.*, 2015). In the present study, we directed the expression of the *thnB* gene from the expression plasmid pET16bTEV in *E. coli* BL21(DE3) in the presence and absence of the plasmid pISC or pSUF as well as in BL21(DE3) Δ*iscR*. Analysis of the protein levels using Western blot show that the co-expression of *thnB* with the pSUF plasmid results in the highest amount of protein (Figure S8.5). A slightly lower expression level has been detected for the co-expression with pISC and the expression in BL21(DE3) Δ*iscR*. The lowest expression level has been detected for the expression in BL21(DE3) in the absence of pISC or pSUF. These results show that specific expression systems tailored for Fe–S proteins not only increase the fraction of holo protein, but also lead to higher protein levels. In some cases where expression has not resulted in detectable protein levels in common BL21(DE3), using BL21(DE3) Δ*iscR* cells improved the expression levels dramatically and large amounts of Fe–S proteins could be isolated (unpublished data).

The electronic spectra show that co-expression with pISC, expression in BL21(DE3) Δ*iscR* cells and BL21(DE3) cells leads to a large fraction of apo ThnB. This is in line with a low metal content of $<$ 1.0 Fe/monomer (Figure 8.7B, Table 8.4). Maturation of ThnB by chemical reconstitution was previously reported to result in considerable amounts protein containing two correctly assembled [4Fe–4S] clusters (Wieckowski *et al.*, 2015). Unfortunately, we were not able to reproduce the high Fe–S content in ThnB. The absorption band at 413 nm in the electronic spectrum of the ThnB protein has a molar extinction coefficient of 12 $mM^{-1}cm^{-1}$, which is considerably lower than in the earlier study. Controversially, these samples contain 13.6 Fe/monomer, which is much higher than expected for a protein binding two [4Fe–4S] clusters. The additional iron most likely corresponds to excess reagent that was not removed by the desalting column or Fe–S aggregates that are often observed for this method.

When using IscS and L-cysteine as sulfur source, the reconstituted ThnB protein reveals a higher Fe–S content compared to the chemical reaction with a molar extinction coefficient of 24 $mM^{-1}cm^{-1}$ at 413 nm. ICP-MS measurements shows that the protein binds 2.5 Fe/monomer. The most efficient Fe–S cluster assembly was achieved by co-expression of the *thnB* gene with the pSUF. The electronic spectra show a high absorption feature at 413 nm with a molar extinction coefficient of 32 $mM^{-1}cm^{-1}$ (Figure 8.7A, Table 8.4), which is similar to the previously reported values for chemically reconstituted ThnB protein. The metal content was determined by ICP-MS and reveals that the protein binds 5.3 Fe/monomer (Figure 8.7B, Table 8.4). The lower metal content could derive from incomplete cluster assembly or an inaccurate protein concentration determination, as reported previously (Galle *et al.*, 2018). The protein concentration was performed using the Bradford method, which so far has resulted in reliable values, however, the purity of the ThnB samples was lower compared to AcnB, IspH, and NadA. Based on SDS-PAGE analysis we estimate the purity of the samples to be between 60% and 80% (Figure S8.1). While the impurities do not seem to interfere with the spectroscopic analysis, the values for the concentration of ThnB are overestimated, resulting in a lower Fe/monomer ratio.

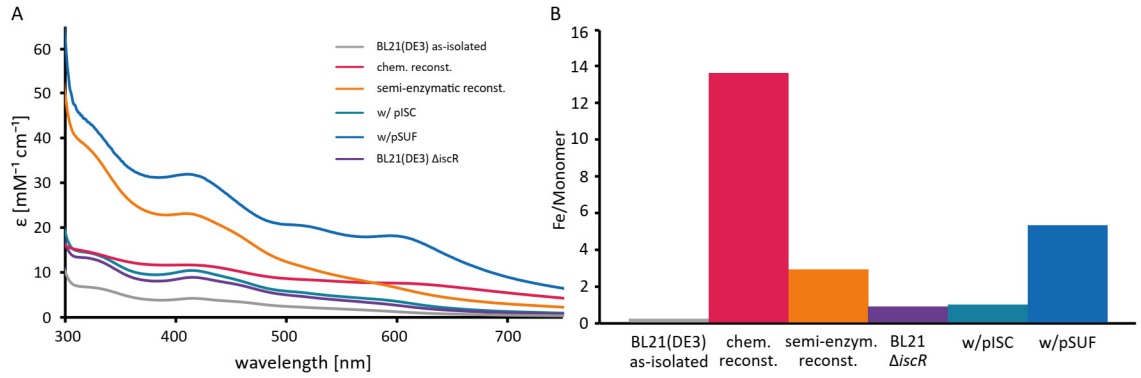


Figure 8.7: Characterization of Fe–S cluster bound to ThnB in different samples. (A) EAS of ThnB samples as-isolated from different cell strains or obtained by different reconstitution methods. Corresponding CD spectra are shown in Figure S8.2D. (B) Iron content per monomer of ThnB proteins shown in (A) determined by ICP-MS measurements. Protein as-isolated from *E. coli* BL21(DE3) is shown in grey, samples obtained by subsequent chemical or semi-enzymatic reconstitution are displayed in pink or orange, respectively. Samples co-expressed with either pISC or pSUF are shown in green or blue, respectively, and protein isolated from BL21(DE3) Δ*iscR* is presented in purple.

Another possible explanation would be that the protein contains one [2Fe–2S] cluster in addition to a [4Fe–4S] cluster. The electronic spectrum of the sample obtained from co-expression contains additional absorption bands at 515 nm and 615 nm. A previous study of the antiviral radical SAM enzyme viperin has assigned the feature at 615 nm to a $[2Fe–2S]^{2+}$ cluster (Duschene & Broderick, 2010). Similar bands have been observed in the electronic spectrum of the active spore photoproduct lyase, a member of the radical SAM superfamily from *Bacillus* (Buis *et al.*, 2006). However, these features have not been analyzed in detail. The Fe content (Figure 8.7B) would be consistent with one [4Fe–4S] and one [2Fe–2S] cluster in ThnB. The CD spectra of ThnB protein obtained by co-expression of the *thnB* gene with the pSUF plasmid and semi-enzymatic reconstitution show a maximum at 450 nm, which indicates that the two [4Fe–4S] clusters are isolated in the reduced form (Figure S8.2D). However, since a CD signal at 450 nm is also characteristic for an oxidized [2Fe–2S] cluster, we cannot exclude the presence of a [2Fe–2S] cluster in ThnB.

Our results show that the presence of the SUF machinery is required for efficient maturation of the radical SAM protein ThnB. This raises the question whether the SUF pathway is the biologically relevant biosynthesis machinery? To our knowledge, the Fe-S biosynthesis in *B. thuringiensis* has not yet been investigated. However, *B. subtilis* does not harbor any protein of the ISC machinery (Albrecht *et al.*, 2010; Yokoyama *et al.*, 2018). We used BLAST (Altschul *et al.*, 1990) to perform sequence alignments in order to determine whether *B. thuringiensis* contains the *isc* operon and we were not able to find any member of the ISC pathway, suggesting that *B. thuringiensis* relies on the SUF pathway for Fe–S cluster biosynthesis.

Table 8.4: Overview on different expression strategies and reconstitution methods for the Fe–S cluster protein ThnB.

Expression strategy	Significant band [nm]	Extinction coefficient [$mM^{-1}cm^{-1}$]	Fe/monomer	Yield [nmol] per l cell culture*	Purity [%]
BL21(DE3)	413	4.2	0.2	31	60
BL21(DE3) chemic. reconst.	413	11.7	13.6	27	–
BL21(DE3) semi-enzym. reconst.	413	24.1	2.5	29	–
BL21 Δ*iscR*	413	9.9	0.9	63	70
BL21(DE3) w/pISC	413	10.5	1.0	61	60
BL21(DE3) w/pSUF	413	31.9	5.3	63	80

* The yield refers to the total amount of protein within the sample not to the exclusive amount of target protein.

8.6 Conclusion

In this study we investigate the maturation of recombinantly produced Fe–S proteins in great detail. We compare six different strategies for protein biosynthesis, including the common expression strain BL21(DE3), three specific cell strains tailored for the maturation of high levels of Fe–S proteins, and two *in vitro* reconstitution methods. The biosynthesis of Fe–S proteins in nature has been studied extensively, however, the maturation of recombinantly produced Fe–S proteins has only been investigated for highly complex systems, such as hydrogenases or nitrogenases. We have, for the first time, systematically evaluated the most common maturation protocols for three well-characterized [4Fe–4S] proteins. The pitfalls of chemical reconstitution pointed out here, such as the formation of Fe–S aggregates, unspecifically bound iron, and different oxidation states of the cluster, have previously led to conflicting results in different studies. These strategies are not limited to proteins containing one [4Fe–4S] cluster. We have previously shown that the same strategies can be applied for the maturation of a [2Fe–2S] protein. In addition, we have demonstrated that this approach can be also used for a radical SAM enzyme. Therefore, we consider the presented methods to be applicable for a large variety of Fe–S cluster-containing proteins. As a consequence, our results will contribute to improve the maturation of Fe-S proteins and obtain samples with correctly assembled cofactors, allowing for understanding the function of the proteins. Maturation of the proteins *in vivo* also leads to more homogeneous samples, which are more amenable for structure determination.

8.7 Supporting data

Supplementary Tables

Table S8.1: Plasmids used in this study.

Plasmid	Reference
pACYC*iscS-fdx* (pISC)	Gräwert *et al.*
pACYC-Duet-1-*suf* (pSUF)	Hänzelmann *et al.*
pET16bTEV-*acnB*	This work
pET16bTEV-*ispH*	This work
pQE30-*ispH*	Gräwert *et al.*
pET46-*nadA*	This work
pET16bTEV-*nadA*	This work
pET16bTEV-*thnB*	This work

Supplementary Figures

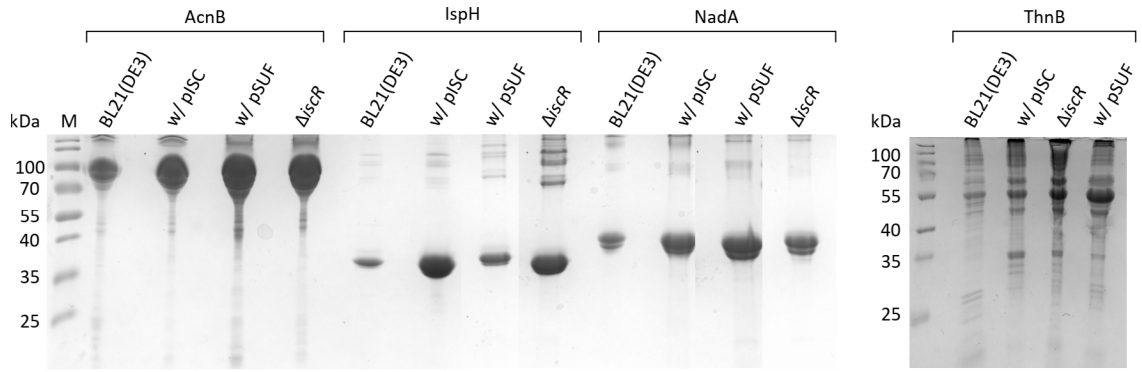


Figure S8.1: SDS PAGE (15%) to analyze the purity of the different proteins after Ni-NTA-column. The fractions containing the target protein were pooled and the volume was reduced to 500 µl, before 5 µl were applied to the gel.

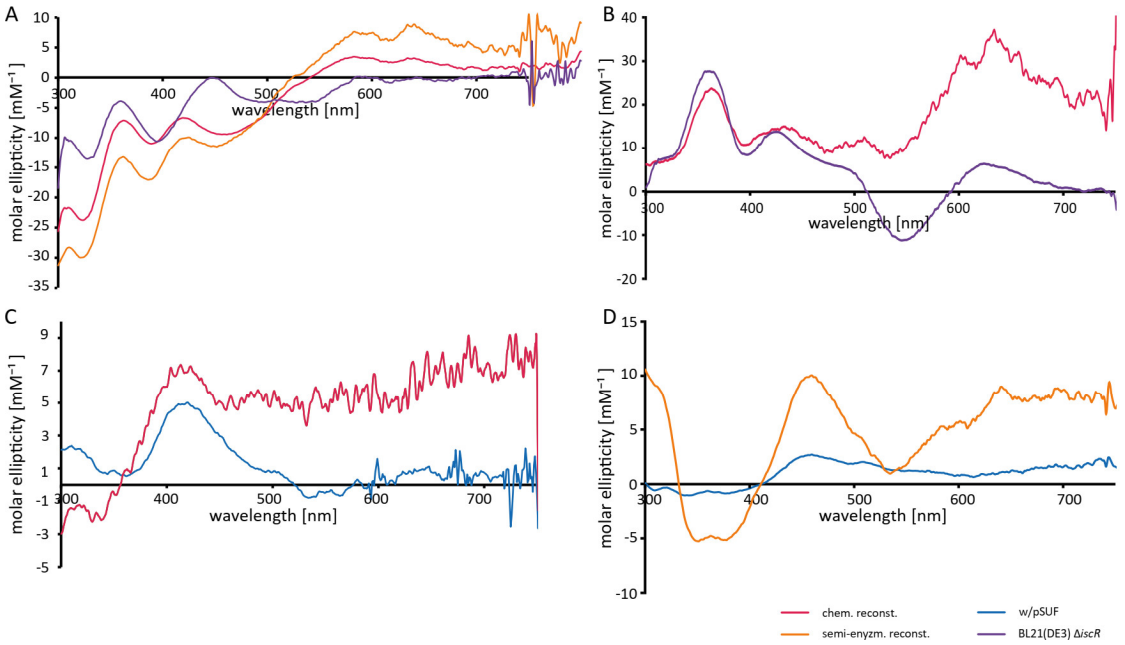


Figure S8.2: CD spectra of the Fe–S cluster containing proteins AcnB (A), IspH (B), NadA (C), and ThnB (D). The concentrations of the protein samples were in the range of 150 µM and 300 µM.

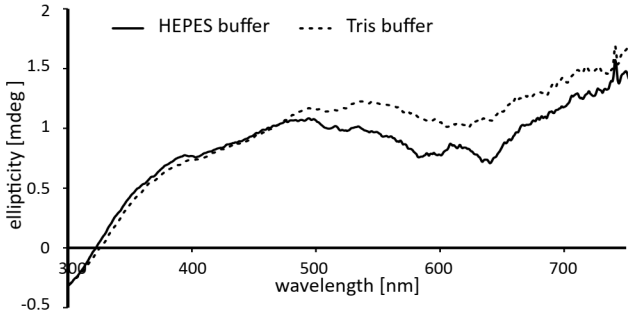


Figure S8.3: CD spectra of HEPES and Tris buffer that have been subtracted from the protein spectra prior to the calculation of the molar ellipticity and plotting the data.

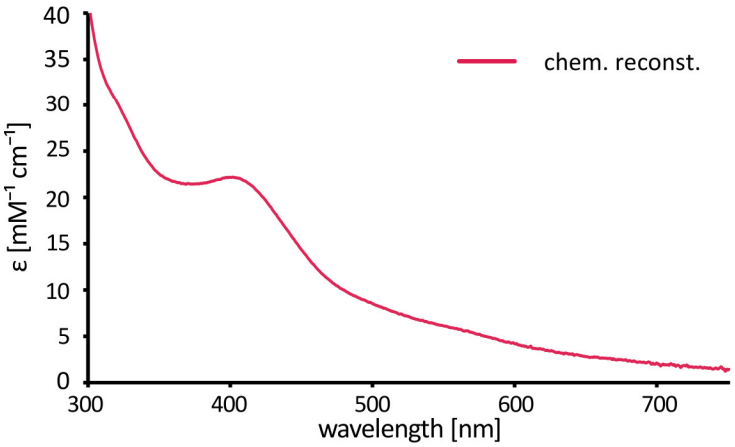


Figure S8.4: EAS of chemically reconstituted IspH after SEC.

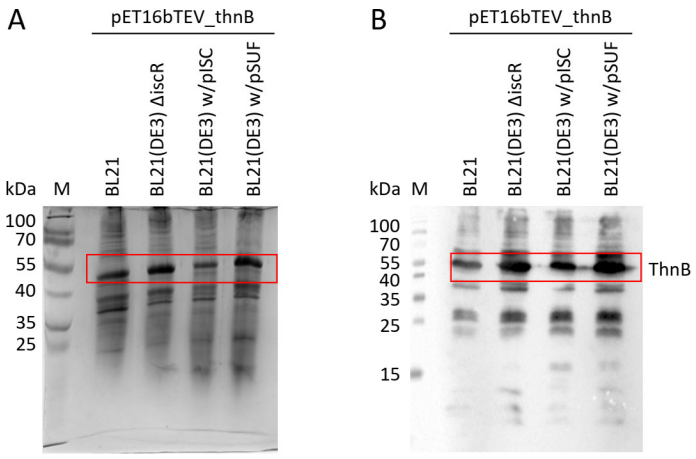


Figure S8.5: Analyzation of the expression of *thnB* and the production of the corresponding protein. (A) Sodium dodecyl sulfate (SDS) polyacrylamide gel (PAGE) (15%) and (B) Western blot using a conjugated anti-His antibody of cell lysates after gene expression.

9 The Asp1 pyrophosphatase from *S. pombe* hosts a $[2Fe\text{-}2S]^{2+}$ cluster *in vivo*

This chapter reflects content of the following publication.

9.1 Publication information

Hannah Rosenbach, Eva K. Walla, George E. Cutsail III, James Birrell, Marina Pascual-Ortiz, Serena DeBeer, Ursula Fleig, and Ingrid Span

To be submitted to: **Journal of Biological Inorganic Chemistry**

9.2 Abstract

The *Schizosaccharomyces pombe* (*S.*) Asp1 protein is a bifunctional kinase/pyrophosphatase that belongs to the highly conserved eukaryotic diphosphoinositol pentakisphosphate kinase (PPIP5K/Vip1) family. The N-terminal Asp1 kinase domain generates specific high-energy inositol pyrophosphate (IPP) molecules, which are hydrolyzed by the C-terminal Asp1 pyrophosphatase domain ($Asp1^{365\text{-}920}$). Thus, Asp1 enzymatic activities regulate the intracellular level of a specific class of IPP molecules, which control a wide number of biological processes ranging from cell morphogenesis to chromosome transmission in the fission yeast *S. pombe*. Recently, it was shown that chemical reconstitution of $Asp1^{365\text{-}920}$ results in a [2Fe–2S] cluster containing protein, however, the biological relevance of the cofactor remained under debate. In this study, we compare the expression in different cell strains and demonstrate that we can isolate the Fe–S-containing $Asp1^{365\text{-}920}$ protein under anaerobic conditions. Expression in *S. pombe* and isolation under anaerobic conditions also results in the Fe–S bound protein. Characterization of the as-isolated protein by X-ray absorption and Mössbauer spectroscopies is in agreement with the presence of a $[2Fe\text{–}2S]^{2+}$ cluster in the enzyme. Furthermore, we have pinpointed the four cysteines involved in coordinating the cofactor by site-directed mutagenesis. We also showed that the Fe–S cluster is sensitive to oxygen and is most likely not involved in redox processes. Our *in vitro* enzymatic activity analysis provides evidence that the $[2Fe\text{–}2S]^{2+}$ cluster does not modulate Asp1 pyrophosphatase activity. Asp1 contains an Fe–S cluster *in vivo* but the absence of the cluster has no measurable effect on the *in vivo* pyrophosphatase function.

9.3 Introduction

Iron-sulfur (Fe–S) clusters are ancient and versatile cofactors that are ubiquitously found in all organisms. Fe–S clusters associated to proteins are essential for numerous biological processes, including electron transfer, substrate binding and activation, redox catalysis, sensing of iron and oxygen, DNA replication and repair, regulation of gene expression, tRNA modification, and genome instability (Beinert *et al.*, 1997; Johnson *et al.*, 2005; Stirling *et al.*, 2011; White & Dillingham, 2012). Fe–S clusters occur in nature in different shapes and nuclearities, the most common types are the rhombic [2Fe–2S] cluster, the cuboidal [3Fe–4S], and the cubic [4Fe–4S] cluster (Johnson *et al.*, 2005). The vast majority of Fe–S clusters are bound to the protein backbone by cysteine residues. In many Fe–S proteins the Cys ligands that bind the cofactor are arranged in characteristic sequence patterns. Despite the constant discovery of novel ligation patterns, it remains challenging to identify Fe–S-containing proteins purely based on sequence analysis. Another challenge in identifying protein-bound Fe–S cofactors is their sensitivity to oxygen. Most Fe–S clusters degrade rapidly in the presence of atmospheric oxygen, thus, numerous native Fe–S proteins are isolated in the absence of the cofactor.

The lack of a well-known characteristic Cys pattern and the sensitivity to oxygen are the reasons why the Fe–S cluster bound to the C-terminal pyrophosphatase domain of Asp1 from *Schizosaccharomyces pombe* (*S. pombe*) has not been discovered until recently (Wang *et al.*, 2015b). Asp1 belongs to the highly conserved Vip1/PPIP5Ks family, which generate a unique subclass of the soluble inositol phosphates (IPs) namely the inositol pyrophosphates (Figure 9.1A) (Menniti *et al.*, 1993; Stephens *et al.*, 1993). Vip1/PPIP5Ks family members are bifunctional enzymes with a N-terminal kinase domain and a C-terminal domain with specific inositol pyrophosphate 1-phosphatase activity (Choi *et al.*, 2007; Dollins *et al.*, 2020; Fridy *et al.*, 2007; Mulugu *et al.*, 2007; Pöhlmann *et al.*, 2014; Wang *et al.*, 2015b). Extensive research has defined numerous biological processes controlled by inositol pyrophosphates and in the fission yeast *S. pombe*, these signaling molecules control cell morphogenesis, microtubule stability, chromosome transmission fidelity, modulation of the actin cytoskeleton and vacuole integrity (Dollins *et al.*, 2020; Feoktistova *et al.*, 1999; Pascual-Ortiz *et al.*, 2018; Pöhlmann & Fleig, 2010; Pöhlmann *et al.*, 2014; Topolski *et al.*, 2016). Specifically, as Asp1 generated inositol pyrophosphates regulate microtubule stability, yeast cell sensitivity or resistance to microtubule poisons such as thiabendazole (TBZ) is directly correlated by lower or higher inositol pyrophosphate levels, respectively (Figure 9.1B, top panel) (Pascual-Ortiz *et al.*, 2018; Pöhlmann *et al.*, 2014; Topolski *et al.*, 2016). The intracellular level of the inositol pyrophosphate 1,5-IP_8 is regulated by the enzymatic activity of the Asp1 pyrophosphatase domain (Pascual-Ortiz *et al.*, 2018). Thus, increased sensitivity to TBZ is a read-out for Asp1 pyrophosphatase activity, while resistance to TBZ is a read-out for an inactive Asp1 pyrophosphatase. Similarly, the transition from the normal surface single-celled growth form to the invasive hyphal growth form is controlled by 1,5-IP_8 and a direct correlation exits between the number of invasively growing colonies and intracellular 1,5-IP_8 levels (Figure 9.1B, bottom panel) (Pöhlmann & Fleig, 2010).

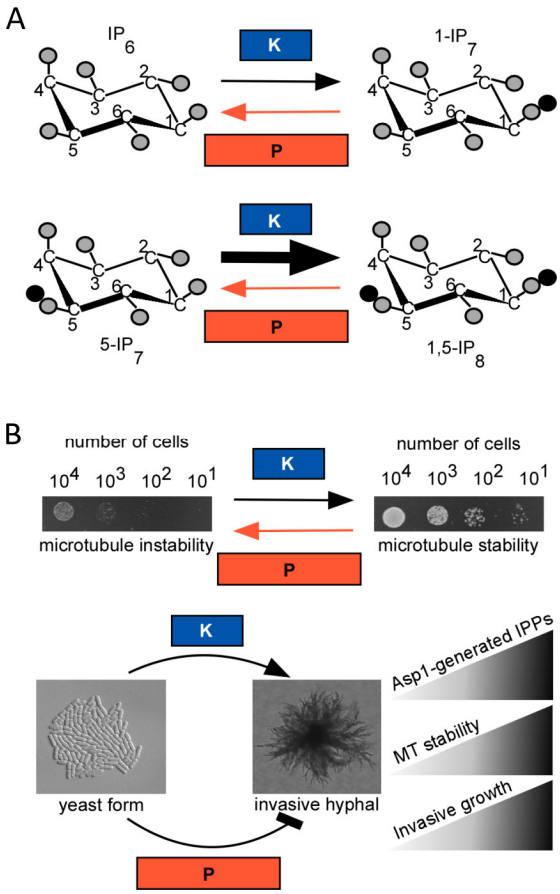


Figure 9.1: Enzymatic activity and biological function of the kinase (K) and pyrophosphatase (P) domain of Asp1. (A) Schematic representation of the reactions catalyzed by the N-terminal kinase and the C-terminal pyrophosphatase domains of *S. pombe* Asp1 (Pascual-Ortiz *et al.*, 2018). Top panel: substrates used for *in vitro* assay; bottom panel: major *in vivo* activities. (B) Biological function of Asp1 in *S. pombe*: Top panels show serial dilution patch tests of yeast cells growing with (left) or without (right) the microtubule destabilizing drug TBZ. The existence of a direct correlation between intracellular 1,5-IP_8 levels and microtubule (MT) stability results in a resistance of the cells to TBZ. Bottom panels: Asp1 kinase activity is required for the switch from surface, single-celled yeast growth (yeast form) to invasive pseudohyphal growth. Asp1 pyrophosphatase activity negatively regulates this switch (Pöhlmann *et al.*, 2014).

A recent study with the Asp1 pyrophosphatase variant $Asp1^{371\text{-}920}$ showed that the recombinantly produced and aerobically purified protein contains substoichiometric amounts of iron as well as acid-labile sulfate, which were increased by chemical reconstitution. The authors proposed that $Asp1^{371\text{-}920}$ binds a $[2Fe\text{–}2S]^{2+}$ cluster, which inhibits its pyrophosphatase activity. Surprisingly, removal of the Fe–S cluster from the reconstituted protein using dithionite (DTH) and ethylenediaminetetraacetic acid (EDTA) did not restore the pyrophosphatase activity. Despite the evidence that Asp1 can bind a cluster *in vitro*, whether the Fe–S cluster is bound to Asp1 protein *in vivo* and its biological relevance remained unclear.

In this study, we demonstrate the presence of the $[2Fe\text{-}2S]^{2+}$ cluster in the Asp1 pyrophosphatase domain *in vivo*. The electronic absorption spectra of the C-terminal pyrophosphatase domain of *S. pombe* Asp1 ($Asp1^{365\text{-}920}$) isolated from *E. coli* BL21(DE3) Δ*iscR* or *S. pombe*

cells under strictly anaerobic conditions show that the protein contains an Fe–S cluster. We have characterized the structure and function of the Fe–S cluster in Asp1$^{365\text{-}920}$ assembled *in vivo* using electronic absorption spectroscopy, X-ray absorption spectroscopy and electron paramagnetic resonance (EPR) spectroscopy. Mössbauer spectra of reconstituted protein support that a [2Fe-2S]$^{2+}$ cluster is the cofactor of Asp1$^{365\text{-}920}$. However, our data show that chemical reconstitution yields an inhomogeneous sample. To identify potential ligands of the cofactor, we have exchanged each of the four cysteines 607, 663, 864 and 879 to serine by site-directed mutagenesis and, in addition, generated a quadruple mutant (QM) with these four cysteines replaced by serine. Furthermore, we have investigated the function of Asp1$^{365\text{-}920}$ in the presence or absence of the [2Fe-2S]$^{2+}$ cluster using *in vivo* and *in vitro* activity assays.

9.4 Material and Methods

9.4.1 Chemicals and buffers

All chemicals were of analytical grade or better. Buffers were prepared using Milli-Q-water.

9.4.2 Strains and plasmids

All strains used are listed in Table S9.1 and all plasmids used are listed in Table S9.2. Generation of the pKM36-*asp1*$^{365\text{-}920}$ construct was performed as previously described (Pöhlmann *et al.*, 2014).

Plasmids harboring *asp1*$^{365\text{-}920,C607S}$, *asp1*$^{365\text{-}920,C663S}$, *asp1*$^{365\text{-}920,C864S}$, *asp1*$^{365\text{-}920,C879S}$, *asp1*$^{365\text{-}920,C607S,C663S,C864S,C879S}$ are derivatives of pKM36. The variants were generated by directed mutagenesis using a QuikChange II site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA, USA) and cloned into the *E. coli* expression vector pKM36 (a gift from Dr. K. Mölleken, Heinrich-Heine-Universität, Düsseldorf, Germany) to generate GST-tagged proteins via homologous recombination in *Schizosaccharomyces cerevisiae* (*S. cerevisiae*) (Jakopec *et al.*, 2011). The expression constructs pJR2-3XL-*asp1*, pJR2-3XL-*asp1*^{D333A}, and pJR2-3XL-*asp1*$^{365\text{-}920}$ for expression in *S. pombe* were generated as previously described (Pöhlmann & Fleig, 2010). The plasmids pJR2-3XL-*asp1*^{C607S}, pJR2-3XL-*asp1*^{C663S}, pJR2-3XL-*asp1*^{C864S}, and pJR2-3XL-*asp1*^{C879S} were generated according to the same protocol. In brief, PCR fragments were generated by directed mutagenesis using a QuikChange II site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA, USA) and cloned into pJR2-3XL18 via homologous recombination in *S. cerevisiae* (Belén Moreno *et al.*, 2000). The quadruple mutant pJR2-3XL-*asp1*$^{365\text{-}920,C607S,\ C663S,\ C864S,\ C879S}$ was obtained by digestion of pJR2-3XL with the restriction enzymes NotI and PstI (New England Biolabs, Ipswich, MA, USA) and homologous recombination in *S. cerevisiae* (Jakopec *et al.*, 2011), followed by leucine selection. For constructing the pJR2-3XL-*asp1*$^{365\text{-}920\text{-}GST}$ plasmid, the GST fragment was generated by PCR and cloned into the vector pJR2-3XL-*asp1*$^{365\text{-}920aa\text{-}GFP}$ (Pöhlmann & Fleig, 2010), which was previ-

ously digested using the restriction enzymes SmaI and NcoI (New England Biolabs) via homologous recombination in *S. cerevisiae* (Jakopec *et al.*, 2011), followed by leucine selection.

9.4.3 Anaerobic gene expression in *E. coli*, cell harvest and lysis

Overnight starting cultures of *E. coli* BL21(DE3), *E. coli* BL21(DE3) pISC, *E. coli* BL21(DE3) pSUF containing the plasmid pKM36-*asp1*$^{365\text{-}920}$, and *E. coli* BL21(DE3) Δ*iscR* containing either the plasmid pKM36-*asp1*$^{365\text{-}920}$ or one of the corresponding plasmids harboring the gene encoding for one of the Asp1$^{365\text{-}920}$ variants were used to inoculate Terrific Broth (TB) medium at 1% (v/v). TB medium was supplemented with kanamycin (50 µg/ml), ampicillin (100 µg/ml) and ferric ammonium citrate (2 mM final concentration). Cells were cultivated aerobically at 37 °C and 160 rpm until the optical density measured at 600 nm (OD_{600}) reached 2. For anaerobic cell growth cultures were then moved to an anaerobic glove box (Coy Laboratory Products, Grass Lake, MI, USA) containing 98% N_2 and 2% H_2. Gene expression was induced by adding 0.5 mM Isopropyl β-d-1-thiogalactopyranoside (IPTG). To facilitate Fe–S cluster assembly and anaerobic metabolism 2 mM L-cysteine as well as 25 mM sodium fumarate were added. Cultures were stirred on a magnetic stirrer at room temperature (RT) for 20 h following induction. Cells were harvested for 10 min at 6,000 × *g* and 4 °C, under an argon atmosphere to maintain anaerobic conditions. For cell lysis, cells were resuspended in 25 mM Tris pH 8.0, 150 mM NaCl. EDTA-free cOmpleteTM Protease Inhibitor Cocktail Tablets (Roche, Basel, Switzerland) were added as needed. After stirring at RT for 20 min under anaerobic conditions the suspension was sonicated (Bandelin electronic, Berlin, Germany) for 20 min with an amplitude of 60% and a pulse of 1 s every 3 s using VS70/T sonotrode, under an argon atmosphere. Lysates were clarified by centrifugation at 40,000 × *g* under an argon atmosphere.

9.4.4 Aerobic gene expression of Asp1$^{365\text{-}920}$ in *E. coli*

In order to obtain a higher protein yield necessary for Mössbauer spectroscopy expression of *asp1*$^{365\text{-}920}$ was carried out under aerobic conditions resulting in a higher cell density and expression level. Aerobic expression was essentially the same as anaerobic expression, but the cultures were not transferred to an anaerobic chamber prior to induction and were incubated at 25 °C and 160 rpm for 20 h. Cell harvest and lysis were performed as previously described, but without exchanging the atmosphere with argon. Aerobically isolated protein was also used for activity assays with the apo protein as well as for reconstitution experiments with the different Asp1$^{365\text{-}920}$ variants.

9.4.5 Aerobic gene expression in *S. pombe*, cell harvest and lysis

S. pombe transformants expressing *asp1*$^{365\text{-}920}$ on plasmid via the *nmt1*$^{+}$ promoter, were grown in minimal medium with supplements for 18 h at 30 °C. Then, 2 × 10^8 cells of a logarithmic *S. pombe* culture were harvested be spinning for 5 min at 3500 rpm. The pellet was washed with 5 ml STOP buffer (0.9% NaCl, 1 mM NaN_3, 10 mM EDTA and 50 mM NaF).

Cells were resuspended under anaerobic conditions in 500 µl HB15 buffer (25 mM MOPS, 60 mM β-glycerophosphate, 15 mM p-nitrophenylphosphate, 15 mM $MgCl_2$, 15 mM Ethylenebis(oxyethylenenitrilo)tetraacetic acid (EGTA), 1 mM DTT, 0.1 mM sodium orthovanadate, 1% Triton X100, 1 mM PMSF and cOmpleteTM Protease Inhibitor Cocktail Tablet, Roche, Basel, Switzerland) and lysed using glass beads. Then, 500 µl of HB15 buffer were added, samples were purged with argon gas, and centrifuged for 30 min at 4 °C and 13,000 rpm. The pelleted cells were then washed with HB15 buffer and centrifuged for 30 min.

9.4.6 Anaerobic protein isolation

Asp1 purifications were carried out under strictly anaerobic conditions at RT. Lysate was applied to a GSTrap HP column (GE Healthcare, Little Chalfont, UK) with a bed volume of 5 ml equilibrated with 25 mM Tris pH 8.0, 150 mM NaCl using an ÄKTAprime plus system (GE Healthcare, Little Chalfont, UK). The column was then washed with 10 column volumes (CV) buffer A before the target protein was eluted with buffer containing 10 mM L-Glutathione. Fractions containing the brownish target protein were pooled. The protein was transferred into 50 mM HEPES, pH 7.5; 150 mM KCl using a HiTrap Desalting column (GE Healthcare, Little Chalfont, UK). Protein concentrations were determined using the Bradford method (Bradford, 1976).

9.4.7 Chemical reconstitution of Asp1$^{365\text{-}920}$ variants

100 µM of aerobically isolated Asp1$^{365\text{-}920}$ was incubated on ice with 50 mM DTT under anaerobic conditions for 1 h and then supplemented with 400 µM ferric ammonium citrate, 400 µM ferrous ammonium sulfate and 800 µM Li_2S. After 1 h the protein was applied to a HiTrap Desalting column (GE Healthcare, Little Chalfont, UK) to separate the protein from excess iron and sulfide. For Mössbauer spectroscopy, aerobically purified Asp1$^{365\text{-}920}$ was reconstituted following the described protocol but with $^{57}FeCl_3$ instead of ferrous ammonium sulfate and ferric ammonium citrate.

9.4.8 *In vitro* activity assay of Asp1$^{365\text{-}920}$

To investigate the activity of aerobically and anaerobically purified as well as reconstituted Asp1$^{365\text{-}920}$ the protein was isolated from *E. coli* BL21(DE3) Δ*iscR* as described in this manuscript. For the kinase reaction, 4 µg of purified Asp1$^{1\text{-}364}$ protein was incubated for 16 h at 37 °C with 300 µM inositol hexakisphosphate (IP_6) (Sigma-Aldrich, St. Louis, Missouri, USA), followed by Asp1$^{1\text{-}364}$ inactivation at 65 °C for 20 min. The Asp1$^{1\text{-}364}$ inactivation was verified by performing a kinase assay with the treated Asp1$^{1\text{-}364}$ protein. For the phosphatase assay, 30 µl of the generated 5-diphosphoinositol(1,2,3,4,6)pentakisphosphate (1-IP_7) were incubated with 2 µg of Asp1$^{365\text{-}920}$ for 3 or 18 h at 37 °C. The assay was carried out under anaerobic conditions. The samples were analyzed on a 35% polyacrylamide gel using electrophoresis, followed by

staining with Toluidine Blue O (Merck KGaA, Darmstadt, Germany) to visualize the inositol compounds.

9.4.9 Invasive-growth assay

Invasive-growth assays were carried out as described (Pöhlmann & Fleig, 2010). *S. pombe* *asp1*^{H397A} transformants were pre-grown in minimal medium without thiamine to allow high expression of the plasmid-encoded *asp1* variants via the thiamine-repressible promoter *nmt1*$^{+}$. 10^4 cells were then patched on plasmid-selective minimal medium agar plates and incubated for 21 days at 25°C. Plates were washed thoroughly to eliminate all surface grown yeast, followed by microscopic determination of the number of invasively growing colonies. Plates were photographed using Axiovert 40FL, Axiocam HR and AxioVisionLE64 programs.

9.4.10 Determination of metal content

10 µM freshly and anaerobically isolated protein were precipitated with 3% trace-metal grade nitric acid before analysis. The protein was pelleted by centrifugation for 20 min at 15,000 $\times$ *g*. The supernatant was then transferred to a metal-free centrifugation tube (VWR, Radnor, PA, USA). The Fe content of the protein was determined by inductively coupled plasma mass spectroscopy (ICP-MS) using an Agilent 7500ce ICP-MS instrument (Agilent Technologies, Ratingen, Germany) in the Central Institute for Engineering, Electronics and Analytics (ZEA-3) at Forschungszentrum Jülich. Samples were measured in triplicates.

9.4.11 Electronic absorption spectroscopy

Electronic absorption spectroscopy (EAS) was used to determine the iron-sulfur cluster content of Asp1$^{365\text{-}920}$, Asp1$^{365\text{-}920,C607S}$, Asp1$^{365\text{-}920,C663S}$, Asp1$^{365\text{-}920,C864S}$, Asp1$^{365\text{-}920,C879S}$ and Asp1$^{365\text{-}920,C607S,C663S,C864S,C879S}$. Electronic absorption spectra were recorded using a Cary-60 spectrophotometer (Agilent Technologies, Ratingen, Germany) with 1 nm bandwidth, a scanning speed of 120 nm/min and a 1-cm-path-length quartz cuvette at RT. For Fe–S cluster stability assays we recorded a spectrum every 30 min and purged the cuvette with oxygen between the measurements.

9.4.12 X-ray absorption spectroscopy

Samples of Asp1 were loaded into custom Delrin X-ray sample cells with a Kapton tape window and frozen and stored in liquid nitrogen until measurement. The total protein concentration of measured sample was 0.35 mM for a total 0.7 mM concentration of Fe. Fe K-edge XAS data were recorded on SSRL beamline 9-3 using a 100-element solid state Ge detector (Canberra) with a SPEAR storage ring current of $\sim$ 500 mA at a power of 3.0 GeV as previously described (Strautmann *et al.*, 2008). The incoming X-rays were selected using a Si(220) double-crystal monochromator and a Rh-coated mirror was utilized for harmonic rejection. Samples

were maintained at $\sim$ 10 K in a liquid helium flow cryostat. Data were calibrated by simultaneously measuring an iron foil, with the first inflection point set to 7111.2 eV. Assessment of short XANES scans ($\sim$ 2-10 min) was used to assess radiation damage and determine dwell time limits. When necessary, the incident beam was attenuated by detuning and/or insertion of aluminum foil into the beam path at varying attenuation lengths. Only scans that showed no evidence of radiation damage were included in the final analysis. Individual PFY-EXAFS scans were evaluated and processed in Matlab 2017a to average selected channels of the multi-element detector and perform normalization of the averaged PFY signal by the incident beam character (I0). Final averaged scans were then further processed within Athena (Ravel & Newville, 2005), where a second order polynomial was fit to the pre-edge region and subtracted throughout the entire EXAFS spectrum. A three-region cubic spline (with the AUTOBK function within Athena) was employed to model the background function to $k = 12$ Å^{-1}. Fourier-transforms were performed over a windowed k range of 2 to 11.75 Å^{-1} and presented without a phase shift correction. Theoretical EXAFS spectra were calculated using Artemis utilizing the multiple scattering FEFF6 code (Ravel & Newville, 2005). The EXAFS amplitude, $\chi(k)$, is given by

$$\chi(k) = \sum_R S_0^2 N \frac{|\, f_{\text{eff}}(k)\,|}{kR^2} \sin(2kR + \phi_k) e^{-2kR/\lambda_k} e^{-2\sigma^2 k^2} \tag{9.1}$$

where S_0^2 is the overall many-body amplitude factor, N is the degeneracy of the paths, $|f_{\text{eff}}(k)|$ is the effective scattering amplitude, and R is the absorber-scatterer distance. A Debye-Waller like factor, exp(-2σ^2k^2) is also included to account for disorder. Lastly, λ_k is the mean free path of the photoelectron and Φ_k is the total photoelectron wave phase shift for the interaction between the absorber and the scatterer. Individual scattering paths were selected and fit within Artemis from paths calculated by FEFF6. The Fourier-transform spectrum (FT) of each were fit over a range of R = 1.0 to 3.0 Å^{-1} (non-phase shift corrected). The FT is the product of a transform of k^3-weighted EXAFS spectrum with a Hann window over the range of k = 2 to 11.75 Å^{-1}. By grouping similar scattering paths of a common coordination shell and increasing its degeneracy, N, the number of variables used for that coordination shell is minimal, 2 variables: σ^2 and ΔR. A single ΔE_0 variable is used for all paths in a given fit. S_0^2 was set to 0.9 for all paths. Goodness of final fits were evaluated by their reduced χ^2-value, defined below and calculated by Artemis,

$$\chi^2 = \frac{N_{\text{idp}}}{\epsilon^2 N_{\text{fit}}} \sum_i^M (d_i - f_i)^2 \tag{9.2}$$

Here d is the raw data, f is the fitted data, ϵ is the estimated noise level, summed over M points of the spectrum. The reduced χ^2 value is normalized for the number of variables used ($N_{\text{idp}}/N_{\text{fit}}$) so that fits of differing number of paths may be statistically compared.

9.4.13 Mössbauer spectroscopy

Mössbauer spectra were recorded on a conventional spectrometer with alternating constant acceleration cooled with an Oxford Instruments Variox cryostat, using a ^{57}Co/Rh (1.8 GBq) γ-

source. Samples were measured at 80 K and with no applied magnetic field. Isomer shifts are relative to iron metal at 300 K. Spectra were simulated and fitted using Lorentzian quadrupole doublets using in house software.

9.5 Results

9.5.1 Expression in BL21(DE3) Δ*iscR* cells leads to Fe–S cluster-containing Asp1$^{365\text{-}920}$

In our previous work we have characterized several recombinant Asp1$^{365\text{-}920}$ variants with electronic absorption spectroscopy and the spectra show no or very low bands at 410 nm, which would indicate the presence of an Fe–S cluster (Pascual-Ortiz *et al.*, 2018). All genes were expressed in *E. coli* Rosetta(DE3) cells and the resulting proteins were isolated in the presence of atmospheric oxygen. High level expression in conventional cell strains, including *E. coli* Rosetta(DE3), leads to partially loaded Fe–S proteins, since the iron-sulfur cluster (ISC) assembly machinery cannot meet the great demand of Fe–S clusters. In addition, protein isolation and purification under aerobic conditions leads to degradation of the oxygen-sensitive Fe–S clusters. Notably, the degree of sensitivity of Fe–S clusters can vary from protein to protein, thus, some Fe–S proteins may not degrade when handled under aerobic conditions. Therefore, we compared different protocols for producing Fe–S proteins to determine if we can isolate Asp1$^{365\text{-}920}$ in the cluster bound form. Hereby, we expressed the truncated version of the *asp1* gene encoding for the residues 365-920 in different cell strains, including *E. coli* BL21(DE3) and BL21(DE3) Δ*iscR* cells, which was genetically engineered to increase the amount of iron-sulfur cluster containing protein significantly (Akhtar *et al.*, 2000). The deletion of the gene encoding the negative regulator protein (IscR) of the *isc* operon leads to an elevated level of the ISC proteins responsible for Fe–S biosynthesis and the cells are capable of producing a higher amount of Fe–S cluster–containing proteins. We also used protein isolated from BL21(DE3) cells and performed chemical and semi-enzymatic reconstitution reactions using the IscS and L-cysteine as sulfur source in the enzymatic reactions. Furthermore, we analyzed protein obtained from BL21(DE3) cells containing an additional plasmid encoding for either the ISC or SUF machinery. The resulting proteins were isolated in an anaerobic chamber with an atmosphere of 96% nitrogen and 4% hydrogen gas and analyzed using EAS as well as ICP-MS.

Our data show that Asp1$^{365\text{-}920}$ contains an [2Fe–2S] cluster, when isolated under anaerobic conditions from cell strains tailored for Fe–S protein production. Electronic spectra of Asp1$^{365\text{-}920}$ (Figure 9.2A) produced in BL21(DE3) cells suggest that the majority of the protein is present in the apo form. Co-expression in cells harboring the pACYC184*iscS-fdx* (pISC) (Gräwert *et al.*, 2004) or pACYC-Duet-1-*suf* (pSUF) (Hänzelmann *et al.*, 2004) plasmid results in protein with a high Fe–S cluster content, indicated by the bands at 325 nm, 410 nm and 470 nm. S-to-Fe charge transfer bands in this range of the spectrum are characteristic for [2Fe–2S]$^{2+}$ clusters (Dailey *et al.*, 1994; Freibert *et al.*, 2018). The features observed at wavelengths higher than 500 nm are not present in every sample and may be a result of Fe–S

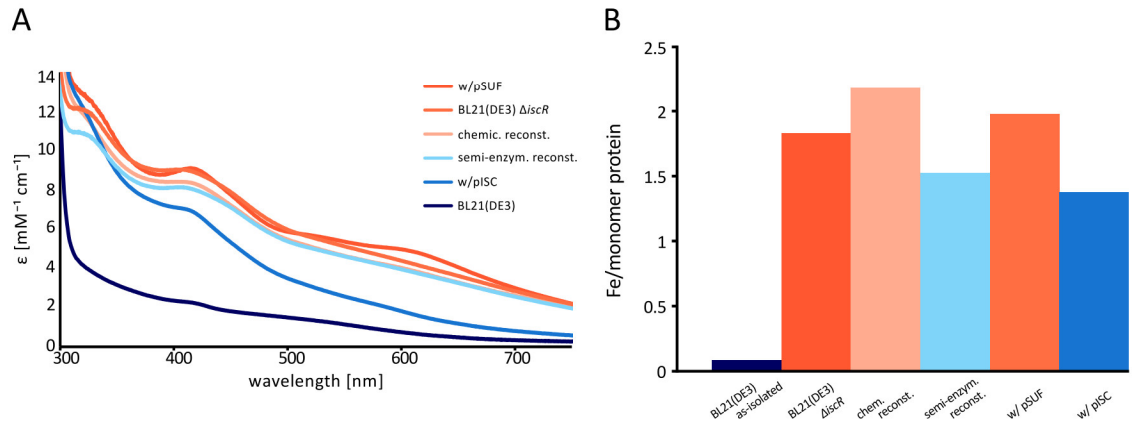


Figure 9.2: EAS (A) and ICP-MS (B) of comparison expression systems.

aggregates that co-purify with the protein. The protein isolated from BL21(DE3) Δ*iscR* cells also shows intense bands in the spectrum, suggesting that the majority of Asp1$^{365-920}$ contains an Fe–S cluster. Protein isolated from BL21(DE3) in the apo form and then reconstituted using either an inorganic or biological source of sulfur show similar levels of Fe–S protein, which are slightly lower than the proteins from the BL21(DE3) Δ*iscR* or BL21(DE3) pSUF cells. The molar extinction coefficients (ε) for all samples except the protein isolated from BL21(DE3) cells are between 7–10 $mM^{-1}cm^{-1}$, which is in good agreement with values found in literature of $\sim$ 1-10 $mM^{-1}cm^{-1}$ (Jacquot, 2017). Interestingly, protein production in BL21(DE3) Δ*iscR* cells results in a higher fraction of Fe–S-containing protein than in BL21(DE3) pISC cells, although the same proteins should be upregulated. A possible explanation could be that the additional plasmid in the cells has a negative effect on protein production or that the levels of ISC proteins varies between the two cell strains. To quantify the amount of iron that is bound to Asp1$^{365-920}$, we analyzed the samples that were used for EAS with ICP-MS. This technique can be used to measure the iron content in the samples and the protein content can be determined using the Bradford method (Bradford, 1976). These two values can then be used to calculate the number of iron ions per monomer. Our results show that protein produced in BL21(DE3) cells and isolated under anaerobic conditions contains almost no Fe–S cluster, since the iron value is close to 0 (Figure 9.2B). The other protocols lead to values close to two irons per monomer, which supports the presence of a [2Fe–2S] cluster in Asp1$^{365-920}$ and suggest that the fraction of holo protein is close to 100%. Chemical reconstitution results in a value that is above two iron per monomer, indicating some unspecific iron binding. In agreement with the spectroscopic results, protein production in BL21(DE3) Δ*iscR* and BL21(DE3) pSUF cell strains results in protein samples with fully assembled [2Fe–2S] clusters. Since the SUF machinery is not present in eukaryotes, including *S. pombe*, we choose the BL21(DE3) Δ*iscR* cells for producing Asp1$^{365-920}$ for all further experiments if not stated otherwise.

9.5.2 Asp1$^{365\text{-}920}$ isolated from the native organism contains an Fe–S cluster

Chemical reconstitution of proteins has previously been reported to result in artefacts. The unspecific incorporation of iron-sulfur clusters inside cells has not been reported previously and is unlikely due to the high specificity of the assembly machinery. However, we aimed at validating that the inorganic cofactor is also incorporated into the Asp1$^{365\text{-}920}$ protein in its native organism, *S. pombe*. The *asp1*$^{365\text{-}920\text{-}GST}$ fusion gene was expressed in *S. pombe* and the resulting protein was isolated and purified under strictly anaerobic conditions. The electronic spectrum (Figure S9.1) confirms the presence of the Fe–S cluster, although with an extinction coefficient that is significantly lower than for Asp1$^{365\text{-}920}$ isolated from the BL21(DE3) strains tailored for Fe–S protein production. The decreased extinction coefficient of the protein isolated from *S. pombe* indicates that only a part of the protein binds the inorganic cofactor, which is most likely a result of the high levels of expression and the inability of the endogenous Fe–S cluster assembly machinery to incorporate the cofactor under these conditions. The presence of the [2Fe–2S]$^{2+}$ cluster in Asp1$^{365\text{-}920}$ isolated from the native organism further supports the biological relevance of the inorganic cofactor for its function.

9.5.3 Characterization of the Fe–S cluster in Asp1 using X-ray absorption spectroscopy

To obtain more information about the electronic structure and the coordinating ligands of the Fe-S cluster we performed X-ray absorption spectroscopy including extended X-ray absorption fine structure (EXAFS) analysis. The Fe K-edge XAS spectrum of the as-isolated Asp1$^{365\text{-}920}$ (Figure 9.3A) exhibits a rising-edge of $\sim$ 7120 eV and a single pre-edge feature at 7112.6 eV of 0.146 units of intensity. These energies are partially consistent with the ferric iron sites in both [2Fe–2S]$^{2+}$ clusters and Fe^{3+} rubredoxin (Figure S9.2) (George *et al.*, 1996). However, the white line of Asp1$^{365\text{-}920}$ appears sharper and more intense than what is typically observed for tetrathiolate coordinated irons. Additionally, the pre-edge has a lower intensity than commonly observed for other Fe–S biological sites. This may indicate light-atom coordination to the ferric iron, or the presence of an additional light-atom coordinated ferric iron species.

The k^3-weighted EXAFS (Figure 9.3B, right) is dominated by an Fe–S scattering interaction as indicated by the intense, partly uniform sinusoidal oscillation in a k-range of 2-7 Å^{-1} and an Fe–Fe at higher k, yielding a deconstructive interference at approximately $k\sim$8 Å^{-1}, characteristic of Fe–S clusters. The non-phase shifted Fourier transform of the EXAFS spectrum (Figure 9.3B, left) yields an intense radial shell at $R \sim$ 1.8 Å, and a second less intense shell at $R \sim$ 2.4 Å. The first and second radial shells are the Fe–S and Fe–Fe scattering interactions, respectively. The intensity ratio of the two shells is typical of [2Fe–2S] cluster, whereas higher nuclearity Fe–S clusters would possess a more intense Fe–Fe scattering interaction. The EXAFS of Asp1$^{365\text{-}920}$ is well fit $N = 4$ Fe–S scattering interaction at a mean scattering distance of 2.25 Å and a Debye-Waller like disorder value of σ^2 = 6.72 $\times$ 10-3 Å^2, Table S9.3. The fitted Fe–S of the four-fold degenerate scattering path is in good agreement with typical Fe–S(cys) and Fe-(μ-S^{2-}) distances found in oxidized ferredoxin clusters. The Fe–Fe scattering interaction

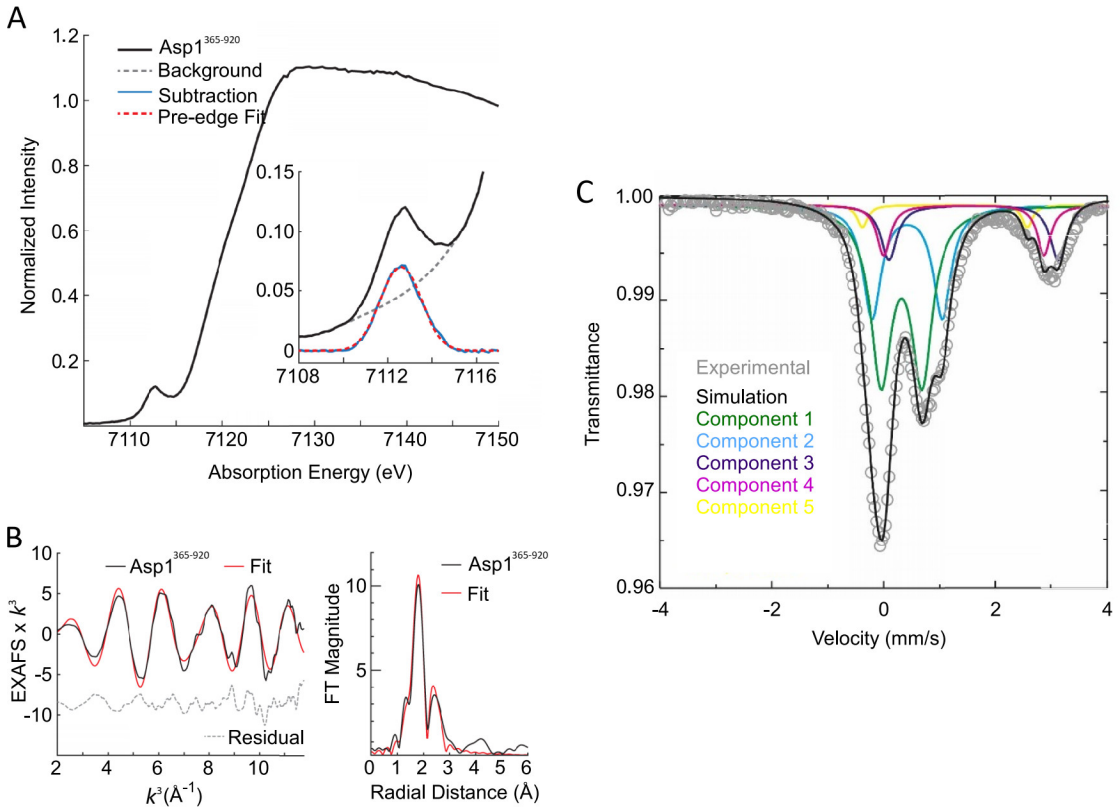


Figure 9.3: Spectroscopic characterization of wild-type Asp1$^{365-920}$. (A) Fe K-edge X-ray absorption spectrum of Asp1$^{365-920}$ with the pre-edge region expanded within the inset. A cubic spline fitting of the edge and background (gray dashed line) is subtracted from the pre-edge feature (blue) and fitted with a single pseudo-Voigt lineshape (red dashed line). The fitted pre-edge feature is centered at 7112.6 eV, a full-width half-max of 1.82 eV and 0.146 units of normalized intensity. (B) The k^3-weighted EXAFS of Asp1$^{365-920}$ (black) with fitted EXAFS (red) and residual (gray dashed) over a k-range of 2-11.75 Å^{-1} (right). The non-phase shifted Fourier Transform of the k^3-weighted EXAFS is displayed (black) with a respected fit (red) (left). (C) Mössbauer spectrum of Asp1$^{365-920}$ enriched with ^{57}Fe measured at 80 K without applied field. The experimental data (gray circles) has been fit (black line) with five components (green, light blue, dark blue, pink and yellow lines). The fitting parameters are presented in Table S9.2.

is fit to distance of 2.73 Å by a well-ordered scattering interaction, reflected by the σ^2-value of 3.57 $\times$ 10-3 Å^2. Attempts to further refine the fitting of the EXAFS data by inclusion of an additional Fe–N/O scattering interaction of shorter distance were not conclusive (Figure S9.2 and Table S9.4). Inclusion of another Fe–N/O scattering interaction yielded fits with better goodness of fit values, however, other parameters of the fit are not physical intuitive, including either negative σ^2-values or no support of Cys ligation in contradiction of mutagenesis result, discussed below. As the technique is a bulk probe and reports the average iron spectrum, influence from minor iron species may contribute. While the assignment of the first coordination shell is not definitive from the EXAFS spectroscopy, a S(cys)4 [2Fe–2S] cluster is supported. The EXAFS does observe the Fe–Fe interaction of the cluster extremely well concluding that a [2Fe–2S]

cluster is present *in vivo* and it is the majority species. Most important finding is that this cluster is formed without the need of iron and sulfur reconstitution.

9.5.4 Characterization of reconstituted wild-type Asp1$^{365\text{-}920}$ by Mössbauer spectroscopy

Mössbauer spectroscopy is a technique that probes transitions between the nuclear ground state and a low-lying nuclear excited state. The nucleus most amenable to Mössbauer spectroscopy is ^{57}Fe, thus, ^{57}Fe Mössbauer spectroscopy is frequently used to obtain detailed information about the chemical environment and electronic structure of iron in Fe–S proteins. One drawback is that the protein samples have to be reconstituted or enriched with the isotope ^{57}Fe. All previous experiments have been performed with as-isolated Asp1$^{365\text{-}920}$ protein and all attempts to fully remove the inorganic cofactor lead to degradation of the protein. Therefore, we used Asp1$^{365\text{-}920}$ protein as isolated from an aerobic purification for chemical reconstitution using $^{57}FeCl_3$ to enrich the ^{57}Fe content. The electronic spectra (Figure S9.3) reveal that the majority of the as isolated protein is in the apo form (extinction coefficient is 2.09 $mM^{-1}cm^{-1}$ at 410 nm). The Mössbauer sample of the reconstituted Asp1 protein has an extinction coefficient of 8.37 $mM^{-1}cm^{-1}$ at 410 nm, which is almost identical to the protein isolated in the Fe–S bound form (extinction coefficient is 8.98 $mM^{-1}cm^{-1}$ at 410 nm). Thus, the protein used for Mössbauer spectroscopy has a similar Fe–S content as the protein produced in BL21(DE3) Δ*iscR* cells. A reasonable fit to the experimental spectrum (Figure 9.3C) was achieved with five components (Table S9.4): two quadrupole doublets with low isomer shifts (δ = 0.32 and 0.42 mm/s) and small quadrupole splittings (ΔE_Q = 0.75 and 1.25 mm/s), typical of iron-sulfur clusters and three quadrupole doublets with high isomer shifts (1.10, 1.44 and 1.61 mm/s) and large quadrupole splittings (2.95, 2.88 and 3.04 mm/s), typical of high spin Fe^{2+}. The latter are likely to be due to some iron aggregates generated during Fe–S cluster reconstitution. The isomer shift and quadrupole splitting for components 1 and 2 are slightly larger than those normally observed for all cysteine ligated [2Fe–2S] clusters ($\sim$ 0.27 mm/s) (Pandelia *et al.*, 2015). However, the isomer shift and the quadrupole splitting can be affected by environment around the cluster (Fee *et al.*, 1984). The unusual Mössbauer properties correlate with the unusual XAS spectral properties and the inability to observe an S = 1/2 signal in the EPR spectrum of the reduced protein.

9.5.5 Cysteine residues 607, 663, 864, and 879 are involved in binding the Fe–S cluster

After characterizing the structure and coordination environment of the $[2Fe\text{–}2S]^{2+}$ cluster in the Asp1$^{365\text{-}920}$ protein, we aimed at identifying the residues involved in linking the inorganic cofactor to the protein backbone. Fe–S clusters are predominantly ligated by cysteine residues, however, they can also be coordinated by histidine, aspartate, arginine, serine tyrosine, or glutamate residues (Freibert *et al.*, 2018). Several consensus sequence motifs are known for

[2Fe–2S] cluster coordination, typically involving four cysteine residues as ligands, yet it remains challenging to identify the coordinating residues based on the sequence. In order to pinpoint the residues involved in cluster coordination, we performed site-directed mutagenesis and characterized the ability to incorporate a [2Fe–2S] cluster. It was previously described that exchange of the cysteine residues 607, 663, 864, 868, 879, and 905 lead to a reduced iron content of Asp1$^{371-920}$ (Wang *et al.*, 2015b). Three of these cysteines, 663, 864, and 879, are highly conserved in some fungal homologs of Asp1. Cysteine 868 in Asp1 aligns with a highly conserved aspartate residue, which could be a potential oxygenic ligand for an Fe–S cluster. As our spectroscopic data suggest that the cluster is exclusively coordinated by cysteines, we concluded that this Cys is not a potential ligand of the cofactor. Cysteine 905 in Asp1 is not conserved in other yeasts and is therefore most likely not involved in cluster ligation. To pinpoint the residues coordinating the cluster, we generated and produced five new Asp1$^{365-920}$ variants: C607S, C663S, C864S, C879S and the quadruple mutant C607S C663S C864S C879S (QM). The position of these cysteine residues in Asp1 protein is shown in Figure 9.4A. All gene constructs were expressed in *E. coli* BL21(DE3) Δ*iscR* and the Asp1$^{365-920}$ variant proteins were isolated under strictly anaerobic conditions followed by analysis of the Fe–S cluster content by electronic absorption spectroscopy and inductively coupled plasma mass spectroscopy (ICP-MS) (Figure 9.4B and C). Our data show that the Asp1$^{365-920/QM}$ is not capable of binding a significant amount of Fe–S cluster and the metal content is lower than 0.1 iron per monomer. The three Asp1$^{365-920}$ variants C663S, C864S, and C879S contain a very low amount of cluster indicated by the band intensities in the electronic spectra and the low metal content of 0.1–0.4 iron per monomer. While the metal content of the Asp1$^{365-920}$ C607S protein is the same as for the wild-type protein with 2.0 iron per monomer, a comparison of the molar extinction coefficients indicates that the amount of Fe–S cluster is approximately 20% lower in the C607 variant. Taken together, our results demonstrate that the cysteines 663, 864, and 879 are essential for binding of the [2Fe–2S] cluster to the protein backbone. Cysteine 607 seems to have an impact on the cluster, since the extinction coefficient at 410 nm is lower compared to Asp1$^{365-920}$, however, the role of this residue in cluster coordination is not clear.

To ensure that the cluster content of all Asp1$^{365-920}$ variants is not a result of accidental degradation during the isolation and purification of the protein, we performed chemical reconstitution on all Asp1$^{365-920}$ variants. The electronic spectra of the proteins after the reconstitution reaction and buffer exchange show that the values for the molar extinction coefficients of Asp1$^{365-920}$ (11.91 $mM^{-1}cm^{-1}$ at 325 nm and 8.98 $mM^{-1}cm^{-1}$ at 410 nm) and C607S (11.64 $mM^{-1}cm^{-1}$ at 325 nm and 7.76 $mM^{-1}cm^{-1}$ at 410 nm) did not increase. This indicates that the Asp1$^{365-920}$ protein was isolated with a fully assembled [2Fe–2S] cluster. It also shows that the lower amount of Fe–S cluster in the C607S variant is due to the inability of this variant to stabilize the cluster, pointing to a possible function of this Cys as cluster ligand. Furthermore, chemical reconstitution of the Asp1$^{365-920}$ variants C663S, C864S, C879S, and QM did not increase the signals at 325 nm and 410 nm, indicating that these mutants are not capable of binding an Fe–S cluster (Figure S9.4).

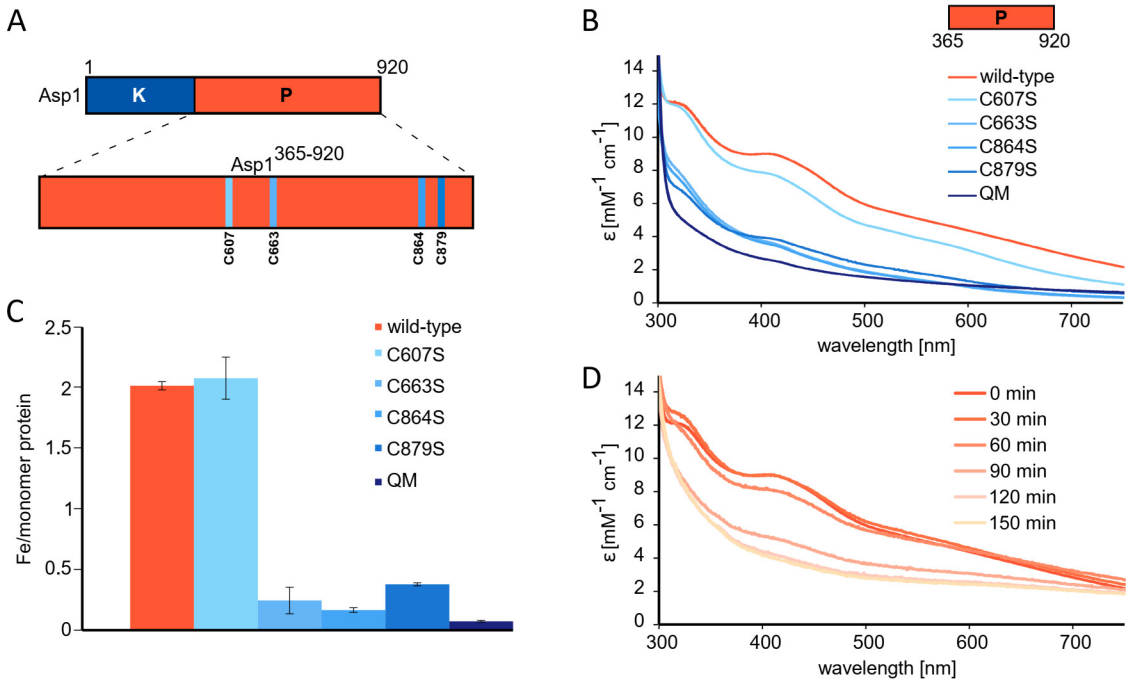


Figure 9.4: Characterization of binding and stability of the Fe–S cluster in the Asp1$^{365\text{-}920}$ variants. (A) Schematic representation of the Asp1$^{365\text{-}920}$ pyrophosphatase domain used for this study in orange with the location of the four cysteines that were replaced utilizing site-directed mutagenesis highlighted in blue. (B) Electronic spectra of Asp1$^{365\text{-}920}$ wt in orange and the variants C607S, C663S, C864S, C879S and QM (different shades of blue) as-isolated from *E. coli* BL21(DE3) Δ*iscR*. (C) Iron content of the proteins shown in (B) determined with ICP-MS. (D) Electronic spectra of Asp1$^{365\text{-}920}$ exposed to molecular oxygen.

9.5.6 The Fe–S cluster degrades when exposed to oxygen

Most iron-sulfur clusters are damaged or degraded when exposed to oxygen, therefore protein isolation and handling is usually performed under anaerobic conditions. [2Fe–2S]$^{2+}$ cluster-containing proteins with four cysteine residues as coordinating ligands are particularly stable and there are examples for highly oxygen-stable Fe–S proteins, such as *E. coli* [2Fe–2S]-ferredoxin (Nakamura *et al.*, 1999). This Fe–S protein can be isolated and even crystallized under aerobic conditions and still contains a fully occupied [2Fe–2S] cluster (Kakuta *et al.*, 2001). The behavior of Fe–S clusters in the presence of oxygen can provide insights into the function of the cofactor. Degradation of the cluster in the presence of oxygen may potentially indicate a role in oxidative stress response. In *E. coli*, the [2Fe–2S] transcriptional factor, SoxR, functions as a sensor of oxidative stress (Ding *et al.*, 1996). Modulation of the Fe–S cluster in SoxR was proposed to control the protein's function in transcription. To analyze the stability of the [2Fe–2S]$^{2+}$ cluster in Asp1$^{365\text{-}920}$ in the presence of oxygen, we exposed the anaerobically isolated protein samples to air and monitored the Fe–S cluster degradation utilizing electronic absorption spectroscopy. Our results show that the cluster in Asp1$^{365\text{-}920}$ is stable in an aerobic environment for approximately 30 min (Figure 9.4D). After 60 min a small fraction of the cluster in Asp1$^{365\text{-}920}$ is degraded and after 120 min most of the cluster is not present anymore. Based on our spectroscopic characterization, we speculate that Cys 607 has an influence on Fe–S

cluster stability, although it is not clear if this residue directly coordinates to one of the iron ions. To gain more insights into the role of residues 607, we also investigated the behavior of the Asp1$^{365\text{-}920}$ C607S variant towards oxygen. In this case, cluster degradation begins within less than 30 min and the cluster is not detectable after 120 min (Figure S9.6A). The stability of the [2Fe–2S] cluster in Asp1$^{365\text{-}920}$ towards degradation by oxygen is, thus, increased by the presence of Cys 607, further supporting its role of in cluster coordination.

9.5.7 The [2Fe–2S]$^{2+}$ cluster in Asp1$^{365\text{-}920}$ is redox inactive

The capability to switch between different oxidation states is a crucial feature for proteins that are involved in electron transfer or redox chemistry. The potential of Fe–S clusters to take part in redox reactions can be estimated by incubating the Fe–S protein with different amounts of reducing agents, such as sodium DTH. Our X-ay absorption spectra suggested the presence of an oxidized [2Fe–2S]$^{2+}$ cluster in Asp1$^{365\text{-}920}$ as isolated from *E. coli* BL21(DE3) Δ*iscR*. This type of cluster is EPR silent. Reduced [2Fe–2S]$^{+}$ clusters with four cysteine ligands typically exhibit average g-values in the range 1.94 to g = 1.96 (Freibert *et al.*, 2018; Orio & Mouesca, 2008). Previous studies using reconstituted and anaerobically reduced Asp1 protein have reported that no EPR signal characteristic of a S = 1/2 [2Fe–2S]$^{1+}$ was observed (Wang *et al.*, 2015b). However, when samples were reduced anaerobically with a twofold excess of DTH in the EPR tube they detected weak rhombic EPR signals with average g-value of 1.96 that accounted for 0.06 spins/Asp1$^{371\text{-}920}$. Our attempts to reduce the *in vivo* prepared samples with 1 and 2 equivalents of DTH were not successful, as shown in the EPR spectra of Asp1$^{371\text{-}920}$ as-isolated and a sample containing 2 equivalents of DTH (Figure S9.7), which do not show evidence of any reduced [2Fe–2S]$^{+}$ cluster as previously observed. Our samples did exhibit a sharp signal at g = 4.3 characteristic of high-spin ferric iron that is estimated to be no more than 10% of the sample. Furthermore, a weak but sharp isotropic signal at g ∼ 2 is also observed with relaxation properties similar to an organic radical. Both the high-spin ferric iron and radical-like signal may arises as a result from cluster decomposition. To further investigate the potential reduction of the cluster, Asp1$^{365\text{-}920}$ and C607S variant, we incubated samples with different amounts of DTH under anaerobic conditions. Reduction of an [2Fe–2S]$^{2+}$ cluster to [2Fe–2S]$^{+}$ may be diagnosed by decreases of the absorbance at 410 nm and 470 nm and corresponding increases at 400 nm and 550 nm (Kim *et al.*, 2013a). The electronic spectra of Asp1$^{365\text{-}920}$ and C607S do not show any changes upon treatment with one equivalent of reducing agent (Figure S9.8), indicating that the oxidation state of the cluster did not change, in agreement with the EPR observations. Further addition of reducing equivalents led to loss of the characteristic absorption bands at 410 nm and 470 nm, suggesting decomposition of the Fe–S cluster.

9.5.8 The Fe–S cluster does not influence phosphatase activity *in vitro*

Our results have shown that Asp1$^{365\text{-}920}$ binds to a [2Fe–2S]$^{2+}$ cluster inside the cells, which is most likely not involved in electron transfer. This raises the questions, whether it influences pyrophosphatase activity or what the role of the inorganic cofactor is. To address the first question, we produced Asp1$^{365\text{-}920}$ in the apo form and compared the activity with the Fe–S-cluster containing protein as isolated. To measure the enzymatic activity of these Asp1$^{365\text{-}920}$ variants, we used an *in vitro* test system that had allowed us previously to show for the first time that a member of the PPIP5K/Vip1 family is a bifunctional enzyme and that the hitherto C-terminal "phosphatase-like domain" of PPIP5K/Vip1 family members is indeed a phosphatase (Pöhlmann *et al.*, 2014). The 1-pyrophosphatase activity of this family is conserved (Dollins *et al.*, 2020; Gu *et al.*, 2017; Wang *et al.*, 2015b; Yousaf *et al.*, 2018). The pyrophosphatase activity measurements were performed in this work were carried out as previously described (Pascual-Ortiz *et al.*, 2018). Briefly, 1-IP$_7$ produced by the Asp1 kinase variant Asp1$^{1\text{-}364}$ was used as a substrate for Asp1$^{365\text{-}920}$ (Figure 9.5A). GST-Asp1$^{365\text{-}920}$ was bacterially produced and isolated under aerobic or anaerobic conditions. Next, 2 µg of anaerobic or aerobically produced Asp1$^{365\text{-}920}$ was added to the mixture and incubated for 3 h or 18 h, followed by the analysis of the resulting inositol polyphosphates via PAGE. Incubation of 1-IP$_7$ with either aerobically or anaerobically produced Asp1$^{365\text{-}920}$ led to a moderate (3 h) or massive (18 h) reduction in the amount of this inositol pyrophosphate (Figure 9.5B and C). Thus, as both forms of Asp1$^{365\text{-}920}$ showed a comparable enzymatic activity, we conclude that the presence of the [2Fe–2S]$^{2+}$ cluster does not inhibit phosphatase activity *in vitro*, in contrast to what has been reported in a previous study by Wang *et al.*.

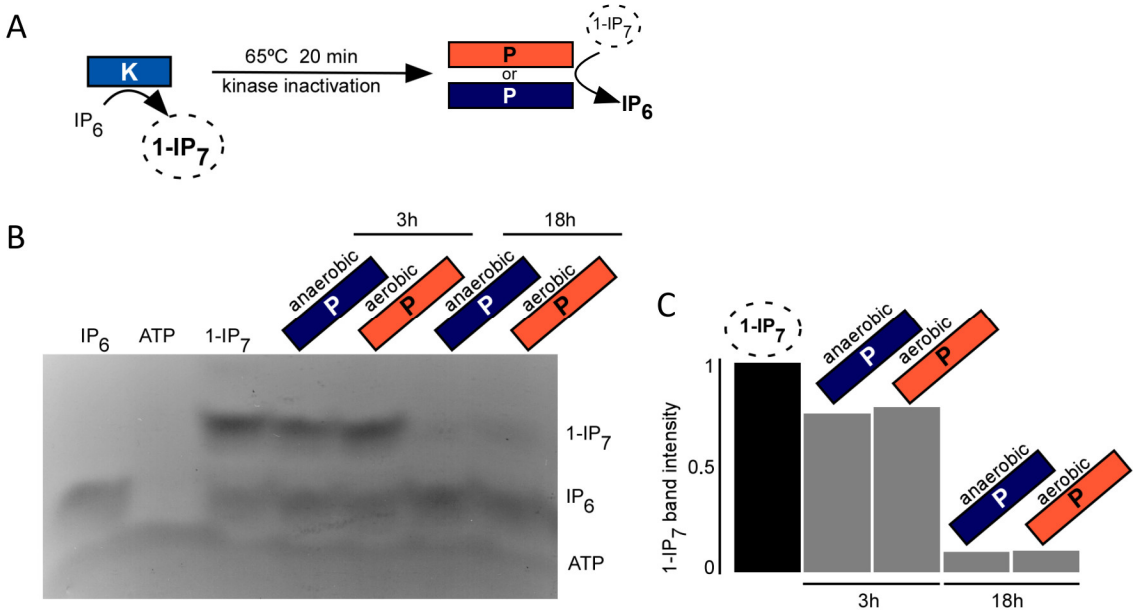


Figure 9.5: Activity of Asp1$^{365-920}$ *in vitro*. (A) Schematic representation of the *in vitro* pyrophosphatase assay. (B) *In vitro* pyrophosphatase assay using the indicated Asp1$^{365-920}$ variants. 2 µg of the indicated proteins were added to Asp1 kinase-generated 1-IP$_7$ (input is shown in the lane labelled 1-IP$_7$) and incubated for 3 h or 18 h. The resulting inositol polyphosphates were resolved on a 35.5% PAGE gel and stained with toluidine blue. The assay was performed twice with reproducible results. (C) Quantification of 1-IP$_7$ bands shown in (B).

9.5.9 Expression of the Asp1$^{365-920/QM}$ variant induces TBZ-hypersensitivity and reduces the ability of yeast cells to grow invasively comparable to Asp1$^{365-920}$

Activity measurements with apo and holo protein have shown that the cofactor does not influence the ability to carry out the pyrophosphatase function *in vitro*. To investigate whether the presence of the Fe–S cluster affects the function of the Asp1 pyrophosphatase *in vivo*, we analyzed two biological processes controlled by 1,5-IP$_8$ in fission yeast. A direct correlation exists between intracellular 1,5-IP$_8$ levels and (i) microtubule stability as well as (ii) the ability to form invasively growing colonies (Figure 9.1B) (Pascual-Ortiz *et al.*, 2018; Pöhlmann & Fleig, 2010; Topolski *et al.*, 2016). As intracellular levels of 1,5-IP$_8$ are controlled by the activity of the Asp1 pyrophosphatase, the *in vivo* activity of Asp1 pyrophosphatase variants can be determined by using the following read-outs: expression of plasmid-borne functional Asp1 pyrophosphatase variants in *S. pombe* transformants will lead to (i) increased sensitivity to the microtubule-destabilizing drug TBZ and (ii) a reduction in the number of cells that can switch to invasive growth. As shown in Figure 9.6B, the plasmid-borne expression of either wild-type Asp1$^{365-920}$ pyrophosphatase or the Asp1$^{365-920/QM}$ variant with cysteines 607, 663, 864 and 879 mutated to serine residues via the thiamine-repressible promoter *nmt1*$^+$ (Figure 9.6A), both resulted in a virtually identical reduction of growth in the presence of the microtubule-destabilizing drug TBZ. Next, we used the *asp1*^{H397A} strain, which expresses an endogenous Asp1 variant with an in-

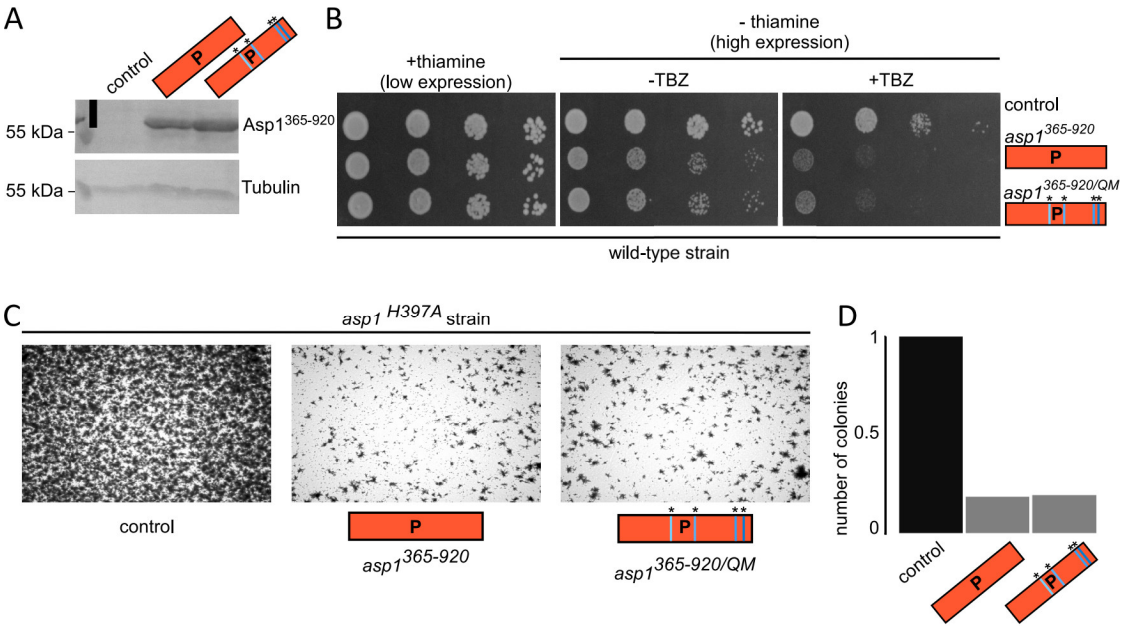


Figure 9.6: Activity of Asp1$^{365\text{-}920}$ *in vivo*. (A) Western blot analysis of a wild-type strain transformed with a control plasmid or plasmids harboring the indicted *asp1* variants. Transformants were grown in thiamine-less, plasmid-selective minimal media to highly express the asp1 variants via the *nmt1*$^{+}$ promoter. Similar amounts of protein extracts were resolved by SDS-PAGE and probed with anti-Asp1 or anti-tubulin antibodies, respectively. (B) Serial dilution patch tests (10^{4}-10^{1} cells) of the transformants shown in (A). Transformants were grown either in the presence or absence of thiamine, which will lead to low or high expression from the *nmt1*$^{+}$ promoter and with or without the microtubule destabilizing drug TBZ. (C) Invasive growth test of strain *asp1*^{H397A} transformed with a control plasmids or plasmids harboring the indicated *asp1* variants. Transformants were grown in the absence of thiamine to allow high expression of *asp1*$^{365\text{-}920}$ and *asp1*$^{365\text{-}920/QM}$. Transformant cells were incubated for 21 d under plasmid-selective conditions, plates washed to eliminate all surface grown cells and the number of invasive colonies quantified via microscopy.

active pyrophosphatase (Pascual-Ortiz *et al.*, 2018). Thus, this strain has significantly higher than wild-type 1,5-IP_8 levels and as a consequence shows an increased ability to switch to invasive growth (Figure 9.6C, left most panel) (Pöhlmann & Fleig, 2010). Expression of either the wild-type Asp1$^{365\text{-}920}$ or the mutant Asp1$^{365\text{-}920/QM}$ in this strain resulted in a decrease of invasively growing colonies (Figure 9.6, middle and right panels and quantification of the number of invasively growing colonies). Again, as shown for the previous *in vivo* assay, the read-out of Asp1$^{365\text{-}920/QM}$ and Asp1$^{365\text{-}920}$ expression is very similar. We conclude that the enzymatic activities of Asp1$^{365\text{-}920/QM}$ and Asp1$^{365\text{-}920}$ are comparable and that the absence of the inorganic cofactor does not affect Asp1$^{365\text{-}920/QM}$ function *in vivo* in the biological processes tested.

9.6 Discussion

The maturation of Fe–S proteins is a strictly regulated process. In bacteria two dedicated machineries are responsible for the correct assembly and transfer of Fe–S clusters: the ISC (Gräwert *et al.*, 2004; Hänzelmann *et al.*, 2004) and SUF pathways are responsible for the

biogenesis of Fe–S proteins under normal and oxidative-stress conditions, respectively. During evolution both machineries were transferred by endosymbiosis to eukaryotes containing Fe–S proteins in mitochondria, cytosol and the nucleus. Biogenesis of Fe–S proteins in the eukaryotic cytosol and nuclei requires the assistance of the mitochondrial ISC assembly machinery, a mitochondrial ISC export system, as well as the cytosolic Fe–S-protein assembly (CIA) machinery (Kispal *et al.*, 1999; Lill, 2009). The first insights into the complex biosynthesis pathways in eukaryotes were established in the yeast *Saccharomyces cerevisiae*, however, subsequent studies have shown that this process is highly conserved from yeast to human. During recombinant gene expression proteins are produced to an abnormally and excessively high level, which frequently leads to incomplete Fe–S cluster incorporation. Several approaches have been used to increase the Fe–S cluster content in proteins, including the co-expression with plasmids containing the *isc* or *suf* operons or the deletion of the regulating *iscR* gene. Expression of the *asp1* gene variant in these cell strains yields high amounts of Fe–S cluster-containing Asp1$^{365-920}$ protein, supporting the biological relevance of the inorganic cofactor. Despite the absence of the SUF machinery in all yeast strains, including *S. pombe*, the co-expression with the *suf* operon results in high protein yields. Since the ISC and SUF machinery are almost completely interchangeable in bacteria, both machineries are supposedly capable to interact with the same proteins. Therefore, it is not surprising that the SUF machinery is efficiently incorporating Fe–S clusters into a eukaryotic protein. Qualitative and quantitative analysis of Asp1$^{365-920}$ in its Fe–S bound form using electronic absorption spectroscopy and ICP-MS, respectively, have shown that expression in BL21(DE3) Δ*iscR* cells results in a high yield of Fe–S protein (Figure 9.2). Chemical or semi-enzymatical reconstitution seem to result in protein with similar spectral properties and iron content. However, the iron/monomer ratio of the samples obtained by chemical reconstitution are above 2, revealing unspecific iron binding. *In vivo* incorporation of Fe–S clusters result in samples with higher homogeneity and purity, since (i) the protein is folded correctly when the Fe–S cofactor is incorporated during or directly after translation, (ii) metal-binding sites may be occupied by non-native metals such as Zn or Ni during expression or purification, leading to inaccessible metal-binding sites (iii) excess Fe–S aggregates can be formed during reconstitution that are quite difficult to separate from the protein, (iv) reconstitution may result in a cluster form that is not identical to the native cluster form, and (v) adventitious iron may bind to the protein in an unspecific manner, interfering with spectroscopic characterization. Spectral features in electronic spectra are not as pronounced as we would expect for a [2Fe–2S] cluster. For rhombic dinuclear clusters, we expect two well resolved bands at 410 and 470 nm. The degree of which structures can be resolved decreases with increasing numbers of iron atom, thus, we only observe one broad feature in the range of 400–450 nm for [4Fe–4S] proteins. The electronic spectra of the Asp1$^{365-920}$ protein as-isolated as well as chemically reconstituted protein reveals a broad feature that is more characteristic for [4Fe–4S] proteins, however, the ICP-MS data show that two iron ions are bound to the protein suggesting a [2Fe–2S] cluster. EXAFS analysis also support the presence of the [2Fe–2S] form, since the intensity ratio of the first and second radial shells, derived from the Fe–S and Fe–Fe scattering interactions, respectively, is typical of [2Fe–2S] cluster. Fe–S clusters with

higher nuclearities would possess a more intense Fe–Fe scattering interaction. The isomer shift and quadrupole splitting in the Mössbauer spectrum are consistent with an all cysteine ligated [2Fe–2S] cluster, even though they are slightly larger. In addition, a majority high spin Fe^{2+} can be fitted into the experimental spectrum, deriving from impurities probably deriving from chemical reconstitution. While all samples for this study were produced *in vivo*, the samples for Mössbauer spectroscopy had to be prepared by chemical reconstitution, since this protocol is by far more cost-effective than *in vivo* incorporation of ^{57}Fe. Our mutagenesis study demonstrates that exchange of cysteines 663, 864, and 879 by serines have a dramatic impact on cluster stability. Replacement of Cys by Ser was chosen, because serines are more likely to maintain the hydrogen bonding network in the structure compared to Ala. Mutation of Cys 607 to Ser also leads to a decreased Fe–S cluster content (Figure 9.4C), however, while the extinction coefficients for the C663S, C864S and C879S variants decrease from 8.98 $mM^{-1}cm^{-1}$ to values ranging between 3.51 and 3.98, the value for the C607S variant only decreases to 7.75. The number of irons per monomer for the C607S variant is also close to the value for $Asp1^{365\text{-}920}$. The hydroxyl group of S607 may be capable of stabilizing the cluster to a certain extent. In this case, the question arises why the hydroxyl group is not able to stabilize the cluster when introduced at another position. A possible explanation may be that this cysteine is more flexible than the others or that the exchange of Cys 663, 864, and 879 leads to local changes in the hydrogen network in a fashion that the hydroxyl group is not available for cluster stabilization. As the $Asp1^{365\text{-}920/QM}$ variant reveals a virtually identical level of phosphatase activity, it is unlikely that amino acid exchange leads to major rearrangements. Another possibility is that there is another Cys located in close proximity of residue 607 that is capable of rearranging to stabilize the Fe-S cluster. It has previously been reported that when cysteine ligands to an Fe-S cluster are changed to alanine, nearby cysteines can substitute and form ligands to the cluster (Golinelli *et al.*, 1996). In absence of structural information, it is difficult to predict which residue would be a potential candidate. Finally, the cluster could be coordinated by three Cys residues, as seen in the scaffold protein ISU1 involved in Fe–S cluster biogenesis (Wu *et al.*, 2002). This would lead to an open coordination site that allows potential binding partners to bind to $Asp1^{365\text{-}920}$ though a Cys residue that occupies the free coordination site, as observed in the IscU:IscS complex (Marinoni *et al.*, 2012). Our EXAFS analysis shows Fe–S scattering interaction that are typical of [2Fe–2S] clusters, which could be modelled satisfactory by majority S coordinated iron-sulfur cluster, in agreement with the electronic and the Mössbauer spectra. Our data are therefore consistent with the [2Fe–2S] cluster coordinated to four cysteines, we therefore propose that Cys 607 is a cluster ligand. Our activity measurements *in vitro* and *in vivo* reveal that the pyrophosphatase function is not influenced by the absence of the [2Fe–2S] cluster. For the *in vitro* analysis, the level of activity of $Asp1^{365\text{-}920}$ isolated under anaerobic conditions in the [2Fe–2S] cluster-containing form or under aerobic condition in the apo form are similar, demonstrating that the cluster is not involved in this reaction. Previously, the FeS cluster was proposed to inhibit the pyrophosphatase activity (Wang *et al.*, 2015b). In this study, activity could not be recovered after removal of the Fe–S cluster, which could indicate that the loss of activity derives from the presence of excess reagents from the reconstitution reaction

or Fe–S aggregates that are formed during chemical reconstitution. Another possibility is that the treatment of the protein to remove the cluster affected the structural integrity leading to an inactive pyrophosphatase domain. The authors did not provide evidence for structural integrity of the samples. This raises the question of the function of the Fe–S cluster in Asp1. A function in redox catalysis or electron transfer are highly unlikely, since all attempts to reduce the cluster have failed so far. A structural role is one possibility, however, since the pyrophosphatase domain is active in absence of the cluster, we conclude that the Asp1 is folded correctly in absence of the Fe–S cluster. Our present *in vivo* analysis tested two biological processes that require 1,5-IP_8: microtubule stability and morphogenesis. In these assays the absence of the [2Fe–2S] cluster in Asp1 did not lead to phenotypic consequences of the yeast strains expressing such Asp1 variants indicting that these two processes are not affected by the absence of the [2Fe–2S] cluster. What then might be the functional role of the Asp1 [2Fe–2S] cluster? We have recently found that Asp1 interacts physically with the mitochondrially localized Met10 protein and that Met10 inhibits the pyrophosphatase activity of Asp1$^{365-920}$ *in vitro* (Pascual-Ortiz *et al.*, 2018). The budding yeast homologue of Met10 protein, *Sc*Met10, is a Fe–S cluster containing protein that physically interacts with Met18/Mms19, which is part of the cytosolic iron-sulfur protein assembly (CIA) machinery (Stehling *et al.*, 2012). Our unpublished observations indicate that an Asp1 subspecies exists in fission yeast mitochondria and Asp1 mutants give rise to altered ATP levels. Altered ATP levels were also observed in human cell lines which did not express PPIP5K (Gu *et al.*, 2017). We therefore speculate that the role of the [2Fe–2S] cluster in Asp1$^{365-920}$ may be related to a possible interaction with the protein Met10.

9.7 Supporting data

Supplementary tables

Table S9.1: *E. coli* and *S. pombe* strains used in this study.

Name	Genotype	Reference
BL21(DE3) Δ*iscR*	F - *ompT hsdS*$_B$ (r_{B-} m_{B-}) *gal dcm iscR*::kan^R (DE3)	Akhtar and Jones
UFY	his3-D1, ade6-M210, leu1-32, ura4-D18, h$^-$	K. Gould
UFY 1579	asp1^{H397A}::kan^R , his3-D1, ade6-M210, leu1-32, ura4-D18, h$^+$	U. Fleig

Table S9.2: Plasmids used in this study.

Name	
For expression in *E. coli*	
pKM36-*asp1*$^{365\text{-}920}$	*amp*^{R}, *tac* Promotor, *lac* Promotor, *GST Factor X site, TRP*, based on pGEX-3X (GE Gealthcare, Chicago, IL, USA)
pKM36-*asp1*$^{365\text{-}920,C607S}$	
pKM36-*asp1*$^{365\text{-}920,C663S}$	
pKM36-*asp1*$^{365\text{-}920,C864S}$	
pKM36-*asp1*$^{365\text{-}920,C879S}$	
pKM36-*asp1*$^{365\text{-}920,C607S,\ C663S,\ C864S,\ C879S}$	
For expression in *S. pombe*	
pJR2-3XL	*LEU2*, *amp*^{R}, pREP3x, thiamin-repressible *nmt1*$^{+}$ promotor
pJR2-3XL-*asp1*$^{365\text{-}920aa}$	
pJR2-3XL-*asp1*$^{365\text{-}920aa\text{-}GST}$	
pJR2-3XL-*asp1*^{D333A}	
pJR2-3XL-*asp1*$^{365\text{-}920,C607S}$	
pJR2-3XL-*asp1*$^{365\text{-}920,C663S}$	
pJR2-3XL-*asp1*$^{365\text{-}920,C864S}$	
pJR2-3XL-*asp1*$^{365\text{-}920,C879S}$	
pJR2-3XL-*asp1*$^{365\text{-}920C607S,C663S,C864S,C879S}$	
For expression in *E. coli*	
pACYC-Duet-*suf* (pSUF)	*Cha*^{R}
pACYC*isc-fdx* (pISC)	*Cha*^{R}

Table S9.3: Comparison of different fits for the EXAFS data collected from Asp1$^{365\text{-}920}$ as-isolated from *E. coli* BL21(DE3) Δ*iscR*.

	N	Path	R(Å)	$\sigma^2 \times 10^3$ (Å^2)	ΔE_0	*F*-value	χ^2
Fit 1	4	Fe–S	2.254	6.72	-0.883	1.437	281.63
	1	Fe–Fe	2.726	3.57			
Fit 2	3	Fe–S	2.258	4.73	0.989	0.876	191.16
	1	Fe–N	2.127	14.92			
	1	Fe–Fe	2.729	4.23			
Fit 3	3.5	Fe–S	2.238	5.79	-1.171	0.814	196.94
	0.5	Fe–N	2.163	-4.06			
	1	Fe–Fe	2.716	4.74			
Fit 4	2	Fe–S	2.268	2.70	3.633	0.549	150.44
	2	Fe–N	2.078	9.22			
	1	Fe–Fe	2.738	4.68			

Table S9.4: Fitting parameters for the Mössbauer spectrum in Figure 9.3C.

Component	1	2	3	4	5
Isomer shift (δ)	0.32	0.42	1.61	1.44	1.1
Quadrupole splitting (ΔE_Q)	0.73	1.25	3.04	2.88	2.95
FWHM	0.45	0.39	0.34	0.29	0.22
Depth	1.78	1.17	0.57	0.52	0.23
% Intensity	48	28	12	9	3
RMSD			0.872		

Supplementary figures

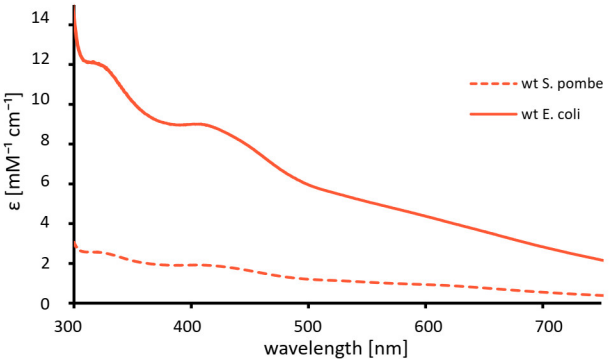


Figure S9.1: Electronic absorption spectra of Asp1$^{365-920}$ as-isolated from *E. coli* BL21(DE3) Δ*iscR* and *S. pombe*.

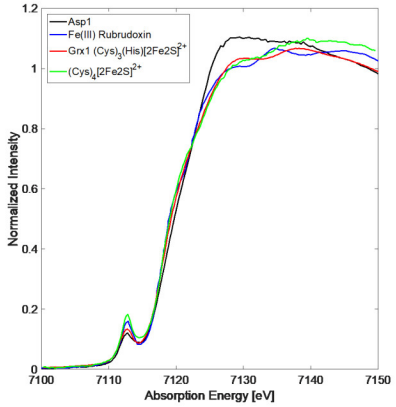


Figure S9.2: Comparison of XAS data collected from Asp1$^{365-920}$ (Asp1) as isolated from *E. coli* BL21(DE3) Δ*iscR* with published data.

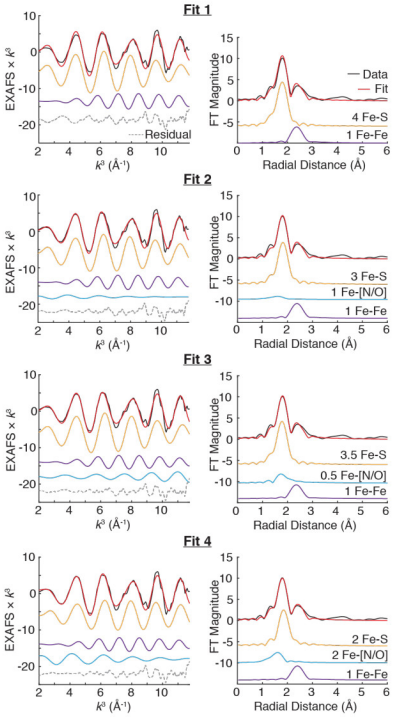


Figure S9.3: EXAFS fits for Asp1$^{365\text{-}920}$.

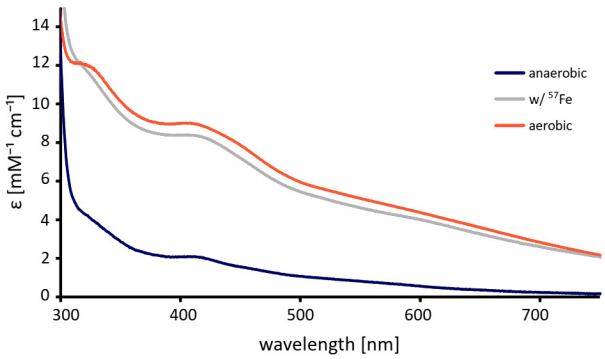


Figure S9.4: Electronic absorption spectra of Asp1$^{365\text{-}920}$ when anaerobically (orange) and aerobicall (blue) isolated from *E. coli* BL21(DE3) Δ*iscR* as well as reconstituted with ^{57}Fe for Mössbauer spectroscopy (grey).

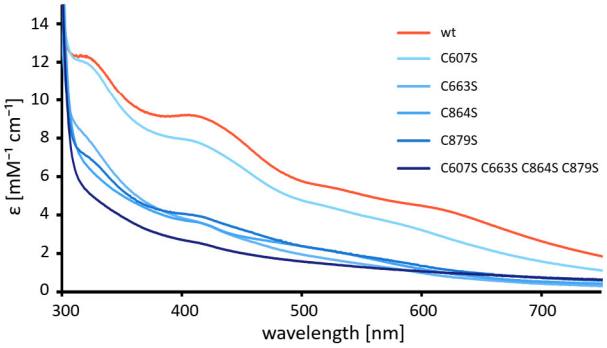


Figure S9.5: Electronic absorption spectra of reconstituted Asp1$^{365-920}$ variants.

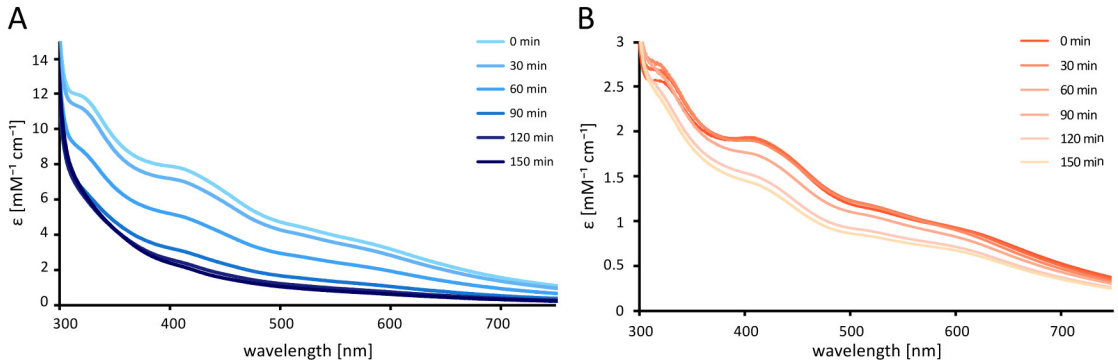


Figure S9.6: O_2-sensitivity assays monitored by EAS. (A) *E. coli* Asp1$^{365-920}$ C607 and (B) *S. pombe* Asp1$^{365-920}$.

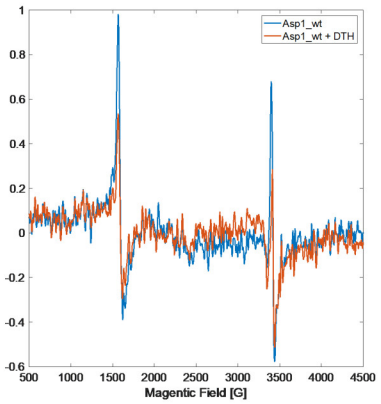


Figure S9.7: Electron paramagnetic resonance spectra of wild type Asp1$^{365-920}$ as-isolated shown in blue and reduced with 2 equivalents DTH shown in orange.

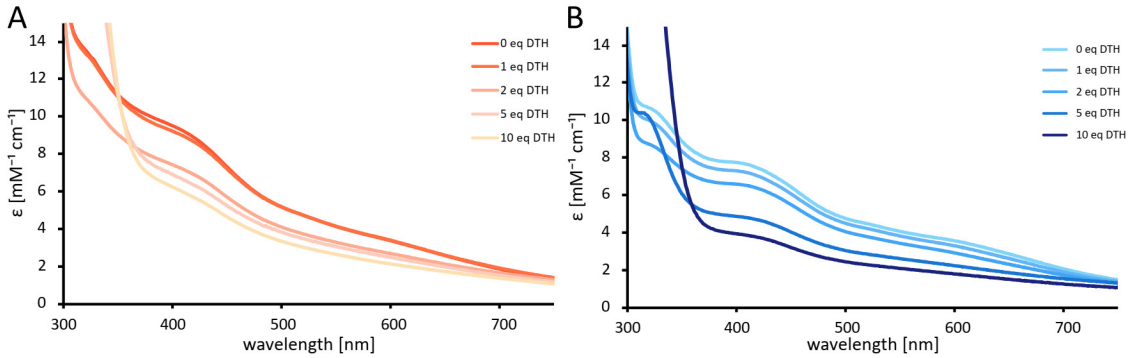


Figure S9.8: Electronic absorption spectra of (A) Asp1$^{365\text{-}920}$ and (B) Asp1$^{365\text{-}920}$ C607S after reduction with different equivalents of sodium DTH.

10 General discussion on iron-sulfur clusters in pro- and eukaryotes

The objective of this thesis was to explore the influence of differnt expression systems on the Fe–S cluster assembly in different target proteins that are known to harbor either one or two [4Fe–4S] clusters or a [2Fe–2S] cluster. While the proteins AcnB, IspH, and NadA, and their Fe–S cofactors were already extensively studied in the past (Bradbury *et al.*, 1996; Gräwert *et al.*, 2004; Jordan *et al.*, 1999; Rousset *et al.*, 2008), only little is known about the proteins ThnB and Asp1. To study the incorporation of [4Fe–4S] clusters we used different expression systems that have been previously successfully tested for the recombinant production of Fe–S cluster containing proteins. These strategies include (i) the co-expression of the target gene with an additional plasmid that harbors the genes encoding for the ISC machinery, (ii) the co-expression of the target gene with an additional plasmid that harbors the genes encoding for SUF machinery, and (iii) the expression of the target gene in an *E. coli* strain that lacks the gene of the transcriptional regulator of the ISC machinery, IscR (Chapter 8). Although many studies focus on the biosynthesis of Fe–S proteins, the maturation of recombinantly produced Fe–S proteins has not been investigated in great detail. We used established expression constructs with a His_6 affinity tag in addition to newly generated constructs with a His_{10} affinity tag in this study. We could detect that the expression levels of the proteins highly depend on the expression construct that is used. However, once a suitable expression construct was found, the expression levels do not show significant variation for the same protein using different expression systems. Only in case of IspH, the expression is weaker in *E. coli* BL21(DE3) in the presence of the additional plasmid pSUF as well as in BL21(DE3) without any additional plasmid (Figure 8.3). A similar trend was observed for the yields after the first purification step. However, we could observe dramatic differences in the cluster incorporation for the different expression systems: For AcnB and IspH isolation from *E. coli* BL21(DE3)Δ*iscR* resulted in highly pure protein with spectroscopic features that are in very good agreement with data reported for the holo proteins (Gräwert *et al.*, 2004; Jordan *et al.*, 1999). ICP-MS data additionally varify the presence of four iron atoms per protein monomer in these samples (Figure 8.4B and 8.5B). The convincing results obtained from this strain may be traced back to the fact that in the absence of an additional plasmid no further stress causing factor is introduced to the expression system. For the protein NadA, the best result was obtained when the protein was expressed in BL21(DE3) in the presence of the plasmid pSUF (Figure 8.6). However, the interaction of NadA with the SUF proteins has not yet been investigated in detail. Previous studies of *E. coli* NadA led to controversial results. Studies on NadA with an *E. coli* strains lacking the *iscS* gene have shown that NAD biosynthesis is impaired, suggesting that *iscS* is essential for the maturation of NadA (Ollagnier-de Choudens *et al.*, 2005). SufS was not able to replace IscS in these experiments and maturate NadA protein. This hypothesis has not been further investigated on a molecular level. Another study reports a specific connection between NadA and the CSD system (Bolstad *et al.*, 2010; Kurihara *et al.*, 2003; Loiseau *et al.*, 2005). The role of

this system in Fe–S cluster biosynthesis has not been established in detail. The only evidence for the interaction of NadA and a member of the SUF machinery was provided in *Arabidopsis thaliana* (Narayana Murthy *et al.*, 2007). It has been reported that the mature SufE3, a SufE-like protein from *Arabidopsis thaliana*, contains two domains: one SufE-like and one with similarity to the bacterial quinolinate synthase, NadA, that hosts a [4Fe–4S] cluster. This cluster could be reconstituted by the SufE-like domaine of the protein in the presence of the cysteine desulfurase CpNifS, cysteine and ferrous iron (Narayana Murthy *et al.*, 2007). It is therefore likely that also during recombinant production of NadA in *E. coli* an interaction between NadA and the proteins of the SUF machinery is supported. The most interesting results have been obtained for the expression of the radical SAM enzyme ThnB. This enzyme has previously been isolated only in its apo form and studies on the active protein have been only performed after chemical reconstitution (Wieckowski *et al.*, 2015). Again, Western blot analysis revealed that the expression was slightly reduced in *E. coli* BL21(DE3) without the knockdown of the transcriptional regulator IscR or the additional plasmids pSUF and pISC, while the expression levels in the modified strains were quite similar. Also, yield and purity after isolation of ThnB from BL21(DE3) were slightly lower than for the other proteins. The most promising results with regard to the [4Fe–4S] cluster specific signals within the electronic spectra and the iron content measured by ISC-MS were obtained from co-expression of *thnB* with pSUF. For this sample with a purity of approximately 80% (Figure S8.1) approximately 6 iron ions could be detected per monomer (Figure 8.7B). Also, the intensities of the molar extinction coefficient of the sample are in the range of what has been reported for other radical SAM enzymes such as AlbA (Flühe *et al.*, 2012). Our results show that the presence of the SUF machinery is required for efficient maturation of the radical SAM protein ThnB. Furthermore, co-expression of other radical SAM enyzmes such as CteB (Grove *et al.*, 2017), CofH, and CofG (Decamps *et al.*, 2012) with pSUF has previously been successfull.Therefore, we suggest that the SUF machinery is the biologically relevant assembly machinery in *B. thuringiensis*. To our knowledge, the Fe–S biosynthesis in *B. thuringiensis* has not yet been investigated. However, it is known that *B. subtilis* does not harbor any protein of the ISC machinery (Albrecht *et al.*, 2010; Yokoyama *et al.*, 2018). Using BLAST (Altschul *et al.*, 1990) we performed sequence alignments in order to determine whether *B. thuringiensis* contains genes of the *isc* operon. We did not find any member of the ISC pathway, suggesting that *B. thuringiensis* relies on the SUF pathway for Fe–S cluster biosynthesis.

For all proteins which have been under investigation chemical reconstitution of the protein led to the formation of aggregates that could not be removed by running an additional purification step. Here, the results might be improved by repeating the desalting step several times. However, this may significantly reduce the protein yield. Taken together, we systematically investigated the suitability of different expression systems for the recombinant production of Fe–S proteins. Our work reveals that one cannot assume a general-purpose system for the production of Fe–S cluster containing holo proteins, but that the expression system needs to be individually optimized for different target proteins. In addition to the EAS that have been discussed in great detail, we performed CD spectroscopy to further analyze the localization of the

absorption bands and to obtain additional information about the oxidation states of the Fe–S clusters (Figure S8.2). We found that the choice of the expression systems can have an impact on the oxidation state of the cluster, indicating the importance of establishing suitable protocols for a native maturation of the target protein.

In an additional study, we investigated the spectroscopic properties of the pyrophosphatase domain of the bifunctional enzyme Asp1 from *S. pombe* (Chapter 9). Previously, it has been reported that this protein hosts a $[2Fe–2S]^{2+}$ cluster after chemical reconstitution. It was reported, that this cluster can be reduced using 2-fold stoichiometric excess of the strong reducing agent dithionite and that the cluster has an inhibitory effect on the phosphatase activity of the protein. These studies were performed on the phosphatase domain inlcuding the amino acids 371 to 920 of the full-length protein. Addtionally, activity assays have been performed on the full-length protein, that also led to a significant decrease of 95% of the phosphatse activity in the presence of the reconstituted Fe–S cluster compared to the as-isolated protein with approximately 40% cluster loading (Wang *et al.*, 2015b). However, it is remarkable, that it was not possible to recover the phosphatase activty after treating both the as-isolated and the reconstituted protein with dithionite and EDTA to remove the cluster. It was therefore suggested, that the Fe–S cluster is not directly involved in phosphatase activity, but indirectly by stabilizing a catalytically relevant conformation. However, studying the activities of the full-length protein it was also found that in the presence of the Fe–S cluster, the product of the kinase reaction is considerably high and comparable to the level produced by a Asp1 mutant that lacks the pyrophosphatase activity. Based on these findings, the authors suggested that under iron-replete conditions *in vivo* the kinase is the predominant activity of Asp1 (Wang *et al.*, 2015b).

The objective of this thesis was to investigate whether Asp1 binds a Fe–S cluster *in vivo* and to characterize the active site of the pyrophosphatase domain of Asp1 ($Asp1^{365\text{-}920}$) as-isolated. For this purpose, we tested the expression in the same expression systems as previously used for the [4Fe–4S] cluster containing proteins. We could show, that the most promising results based on a combination of electronic absorption spectroscopy and ICP-MS measurements were obtained using either the *E. coli* strain BL21(DE3)Δ*iscR* or the additional plasmid pSUF for the co-expression in *E. coli* BL21(DE3) (Figure 9.2). The molar extinction coefficients for the samples after the first purification step at 410 nm, 420 nm, and 325 nm were in the same range as reported for the reconstituted samples of $Asp1^{371\text{-}920}$ by Wang *et al.* (2015b). Additionally, we were not able to further increase the molar extinction coefficients of these samples by adding inorganic iron and sulfur, indicating a full cluster occupancy (Figure S9.5). In the absence of enhanced SUF or ISC activity, nearly no Fe was found in the sample. For further spectroscopic analysis, $Asp1^{365\text{-}920}$ was isolated under anaerobic conditions from BL21(DE3)Δ*iscR*. The same expression system was used for the recombinant production of $Asp1^{365\text{-}920}$ variants with single cysteine mutations to serine at the positions 607, 663, 864, and 879, as well as quadruple mutant with all four cysteine residues mutated to serine. Significantly decreased molar extinction coefficients at 325 nm and 410 nm observed for all of the different mutants indicate that the cysteine residues at all these positions are involved in the coordination of the Fe–S cluster. However, the cysteine 607 seems to be not as critical for stabilizing the cofac-

tor. We propose, that this residue is flexible and therefore capable of reorganizing to create an unoccupied coordination site for binding of another ligand or that another residue, for example another cysteine or a histidine or aspartate close by is capable of stabilizing the Fe–S cluster in the protein variant C607. Exposure of the wild-type Asp1$^{365\text{-}920}$ and the variant C607S to oxygen revealed that the Fe–S cluster is not stable under aerobic conditions.

We also aimed at validating that the Fe–S cluster is present in Asp1$^{365\text{-}920}$ when isolated from the native organism *S. pombe*. For this purpose, we isolated the protein from *S. pombe* under strictly anaerobic conditions and analyzed the protein sample using EAS (Figure S9.1). Although the extinction coefficient was lower than for the same protein isolated from *E. coli* BL21(DE3)Δ*iscR* we could detect a characteristic absorption band, indicating the presence of an Fe–S cluster for this sample. We assumed that the cluster occupancy in the Asp1$^{365\text{-}920}$ isolated from *S. pombe* is lower since the protein is expressed at an abnormally high level and that therefore the Fe–S assembly machineries in *S. pombe* cannot cover the extended amount of Fe–S cluster that is needed. However, the presence of the Fe–S cluster within the protein isolated from *S. pombe* reveals that the inorganic cofactor plays an important biological role.
To verify our assumption that the amino acid residues at positions 607, 663, 864, and 879 are involved in the coordination of the Fe–S cluster, we performed X-ray absorption spectroscopy (XAS) including extended X-ray absorption fine structure (EXAFS) analysis (Figure 9.3A and B, S9.2 and S9.3). The information that we obtained from the data can be summarized in two conclusions: (i) the EXAFS does observe the Fe–Fe interaction of the cluster strongly supporting the presence of a [2Fe–2S] cluster *in vivo*, and (ii) the Fe–S cluster is most likely coordinated by four cysteine residues. However, the white line of Asp1$^{365\text{-}920}$ in the X-ray absorption spectra appears sharper and more intense than what is typically observed for tetrathiolate coordinated irons George *et al.* (1996). Additionally, the pre-edge has a lower intensity than commonly observed for other Fe–S biological sites. XAS experiments performed on the mutant C607S did not lead to high-quality data and have therefore not been published. However, the data suggest that in the Asp1$^{365\text{-}920}$ C607S is not capable of binding a [2Fe–2S] cluster.
Mössbauer spectroscopy (Figure 9.3C) performed on a reconstituted sample also highly suggests the presence of a [2Fe–2S] cluster within the protein, although a low amount of high spin Fe^{2+} has been detected. However, this species is due to aggregates generated during the reconstitution process. The isomer shift and quadrupole splitting for the [2Fe–2S] cluster components are slightly larger than those normally observed for all cysteine ligated [2Fe–2S] clusters (Pandelia *et al.*, 2015), which is in agreement with the unusual spectroscopic properties observed in XAS.

EPR measurements on the as-isolated Asp1$^{365\text{-}920}$ in the presence and absence of the reducing agent dithionate did not yield a S = 1/2 signal characteristic of the reduced $[2Fe–2S]^{+}$ (Figure S9.7) as it was previously detected on a reduced sample of Asp1$^{371\text{-}920}$ (Wang *et al.*, 2015b). Attempts to observe the reduction of the [2Fe–2S] cluster by EAS have also not been successful. Reduction of an $[2Fe–2S]^{2+}$ cluster to $[2Fe–2S]^{+}$ can be detected by decreases of the absorbances at 410 nm and 470 nm and corresponding increases at 400 nm and 550 nm (Kim *et al.*, 2013a). The EAS of the as-isolated Asp1$^{365\text{-}920}$ wildtype and C607S variants did

not show such a change, indicating that it is not possible to change the oxidation state of the cluster (Figure S9.8).

In addition to the spectroscopic analysis, we performed activity measurements both *in vitro* and *in vivo*. *In vitro* we could not detect significant differences in the pyrophosphatase activity of the anaerobically and aerobically isolated protein. Previously, the Fe–S cluster was proposed to inhibit the pyrophosphatase activity (Wang *et al.*, 2015b). In this study, activity could not be recovered after removal of the Fe–S cluster, which may suggest that the loss of activity derives from the presence of excess inorganic iron and sulfur from the reconstitution process or Fe–S aggregates that are formed during chemical reconstitution. Furthermore, it might be possible that the treatment of the protein to remove the cluster affected its structural integrity leading to an inactive pyrophosphatase domain. Unpublished observations from our group provide evidence that Asp1$^{365\text{-}920}$ tends to aggregate when treated with EDTA or DTH. Our observations raise the question of the function of the Fe–S cluster in Asp1. Since all attempts to reduce the cluster have failed, we consider a function in redox catalysis or electron transfer as highly unlikely. In addition, a structural role of the Fe–S cluster can be excluded, since the pyrophosphatase domain is active in absence of the cluster. Our *in vivo* analysis testing two biological processes that require 1,5-IP$_8$, i. e. microtubule stability and morphogenesis, revealed that the absence of the [2Fe–2S] cluster in Asp1 does not lead to phenotypic consequences of the yeast strains. This observation indicates that these two processes are not affected by the absence of the [2Fe–2S] cluster. So what else may be the function of the Fe–S cluster in Asp1? Our collaboration partners recently found that Asp1 interacts physically with the mitochondrially localized Met10 protein and that Met10 inhibits the pyrophosphatase activity of Asp1$^{365\text{-}920}$ *in vitro* (Pascual-Ortiz *et al.*, 2018). The yeast homologue of the Met10 protein, *Sc*Met10, is a Fe–S cluster containing protein that physically interacts with Met18/Mms19, which is part of the cytosolic iron-sulfur protein assembly (CIA) machinery (Stehling *et al.*, 2012). The unpublished observations from our collaboration partners highly suggest that an Asp1 subspecies exits in fission yeast mitochondria and Asp1 mutants give rise to altered ATP levels. Such altered ATP levels were also observed in human cell lines which did not express PPIP5K (Gu *et al.*, 2017).

We did not discuss our unpublished observation that a significant amount of nucleic acids was detected within the Asp1$^{365\text{-}920}$ sample when we did not add any DNase to the purification buffer. We, therefore, additionallz suggest that a possible role of the [2Fe–2S] cluster in Asp1$^{365\text{-}920}$ might be involved in interacting with nucleic acids as for example reported for the protein DNA2 (Lutz, 2018), a nuclease/helicase from *S. cerevisiae*. Another possible function of the Fe–S cluster might be a role in oxygen sensing in *S. pombe*, as it has been reported for the protein AirSR from *Staphylococcus aureus* that hosts a [2Fe–2S] cluster (Sun *et al.*, 2012). However, at this point, we cannot definetly answer the question of the functional role of the [2Fe–2S] cluster in Asp1. More detailed experiments will be necessary to fully understand its function. We are aiming at solving the crystal structure of either the full-length Asp1 or Asp1$^{365\text{-}920}$ in *E. coli* to fully understand the function of the Fe–S cluster and its coordination geometry as well as to interprete the unusual data that we obtained from XAS, EXAFS and Mössbauer spectroscopy in a more conclusive way.

With this work, we obtained important knowledge about the maturation of recombinantly produced proteins that is of great interest for the bioinorganic community. Using two poorly characterized proteins, namely the radical SAM protein ThnB from *B. subtilis* as well as the pyrophosphatase domain Asp1$^{365\text{-}920}$ of a bifunctional protein from *S. pombe* we could demonstrate that both [2Fe–2S] and [4Fe–4S] cluster proteins can become accessible for spectroscopic characterization in their native form. Our work contributes to efficiently produce Fe–S proteins in their biologically relevant form and to obtain a deeper understanding of the function of these ubiquitously existing proteins.

11 Danksagung

Ganz herzlich danke ich meiner Erstgutachterin Frau Jun.-Prof. Dr. Ingrid Span für die freundliche Aufnahme in ihre Arbeitsgruppe und die Bereitstellung der spannenden und herausfordernden Projekte, an denen ich im Rahmen meiner Promotion arbeiten durfte. Ich habe diese Zeit als sehr lehrreich und stets als eine Mischung aus Fordern und Fördern empfunden. Herrn Prof. Dr. Peter Gilch danke ich sehr herzlich für die Übernahme des Zweitgutachtens und die Möglichkeit, meinen Forschungsstand in regelmäßigen Abständen in seiner Arbeitsgruppe vorzustellen und zu diskutieren. Herrn Prof. Dr. Dieter Willbold habe ich für die Bereitstellung der Laboratorien und Verbrauchsmaterialien zu danken. Ganz besonders herzlich bedanke ich mich bei Herrn apl. Prof. Dr. Ing. Gerhard Steger für seine Unterstützung, Hilfe und Tipps in sämtlichen Bereichen des wissenschaftlichen und nicht-wissenschaftlichen Lebens (Nukleinsäuren, IT, Wandern, Kochen, ...), ganz besonders aber natürlich für die tolle Zusammenarbeit im DNAzym-Projekt und die vielen schnellen Hilfen mit LaTeX. Mein besonderer Dank richtet sich auch an Herrn Prof. Dr. Dr. h.c. Detlev Riesner und seine Frau Prof. Dr. Hannelore Riesner. Sie haben mich in den letzten Jahren in vielerlei Hinsicht unterstützt und sind mir stets hilfsbereite und herzliche Mentoren gewesen, sei es in der Wissenschaft als auch darüber hinaus. Ich danke Herrn Dr. Manuel Etzkorn und seinen Arbeitsgruppenmitgliedern Dr. Julian Victor, Dr. Aldino Viegas und Jan Borggräfe für die tolle und interessante Zusammenarbeit am DNAzym-Projekt. Bedanken möchte ich mich auch bei den „guten Seelen“ des IPB, Astrid Wies, Barbara Schulten und Elke Reinartz. Sie haben nicht nur stets ein offenes Ohr für sämtliche Belange des Lebens, sondern sorgen durch ihre „Arbeit im Hintergrund“ jeden Tag dafür, dass das Institut am Laufen gehalten wird, wir immer ausreichend Materialien im Labor und in der Küche haben und unsere Rechnungen bezahlt und Anträge richtig ausgefüllt werden. Allen aktuellen und ehemaligen Mitgliedern der AG Span danke ich für die Zusammenarbeit in den letzten Jahren. Mein Dank richtet sich auch an die vielen Kooperationspartner außerhalb des IPB, mit denen ich in den letzten Jahren an spannenden Projekten zusammenarbeiten durfte und ohne die meine Arbeit in diesem Umfang nicht möglich gewesen wäre. Besonders erwähnen möchte ich hier Prof. Dr. George E. Cutsail III, Dr. James Birrell, apl. Prof. Dr. Ursula Fleig, Eva K. Walla, Prof. Dr. Olav Schiemann, Christine Wübben und Assistant Prof. Dr. Aaron T. Smith.

Ein ganz besonderer Dank gilt meinen Eltern Renate und Achim, die mich stets kompromiss- und bedingungslos auf meinem Weg unterstützt haben und auf die ich mich zu jeder Zeit voll und ganz verlassen kann. Genauso sehr danke ich meinem Freund Steffen, der seit nun drei Jahren der ruhende Gegenpol zu mir ist und mich durch seine Ausgeglichenheit und Fürsorge immer und überall unterstützt. Ohne ihn hätte ich die letzten Monate nur schwer überstanden. Meinen Freundinnen Lisa und Lena danke ich für die nun schon jahrelange Freundschaft und die vielen unbeschwerten Stunden. Danken möchte ich auch Julian und Michael, die ich im IPB nicht nur als tolle Kollegen, sondern auch als Freunde gewonnen habe und die die Zeit im Institut zu etwas Besonderem gemacht haben. Dankend erwähnen möchte ich auch die (Ex-)IPB'ler Anne, Marcel, Marie, Emil, Jan und Alessia für die schönen Stunden.

Bibliography

Achenbach, J. C., Nutiu, R. & Li, Y. (2005). Structure-switching allosteric deoxyribozymes. *Analytica chimica acta*, 534(1), 41–51.

Adams, P. L., Stahley, M. R., Kosek, A. B., Wang, J. & Strobel, S. A. (2004). Crystal structure of a self-splicing group I intron with both exons. *Nature*, 430, 45–50.

Adinolfi, S., Iannuzzi, C., Prischi, F., Pastore, C., Iametti, S., Martin, S. R., Bonomi, F. & Pastore, A. (2009). Bacterial frataxin CyaY is the gatekeeper of iron-sulfur cluster formation catalyzed by IscS. *Nature structural & molecular biology*, 16(4), 390.

Agar, Jeffrey N, Krebs, Carsten, Frazzon, Jeverson, Huynh, Boi Hanh, Dean, Dennis R & Johnson, Michael K (2000). IscU as a scaffold for iron- sulfur cluster biosynthesis: sequential assembly of [2Fe-2S] and [4Fe-4S] clusters in IscU. *Biochemistry*, 39(27), 7856–7862.

Agar, J. N., Zheng, L., Cash, V. L., Dean, D. R. & Johnson, M. K. (2000). Role of the IscU Protein in Iron- Sulfur Cluster Biosynthesis: IscS-mediated Assembly of a [Fe2S2] Cluster in IscU. *Journal of the american chemical society*, 122(9), 2136–2137.

Aguirre, S. D., Ali, M. M., Salena, B. J. & Li, Y. (2013). A sensitive DNA enzyme-based fluorescent assay for bacterial detection. *Biomolecules*, 3(3), 563–577.

Aguzzi, A. (2006). Prion diseases of humans and farm animals: epidemiology, genetics, and pathogenesis. *Journal of neurochemistry*, 97(6), 1726–1739.

Akhtar, M. K. & Jones, P. R. (2008). Deletion of iscR stimulates recombinant clostridial Fe–Fe hydrogenase activity and H 2-accumulation in Escherichia coli BL21 (DE3). *Applied microbiology and biotechnology*, 78(5), 853–862.

Akhtar, S., Hughes, M. D., Khan, A., Bibby, M., Hussain, M., Nawaz, Q., Double, J. & Sayyed, P. (2000). The delivery of antisense therapeutics. *Advanced drug delivery reviews*, 44(1), 3–21.

Albrecht, A. G., Netz, D. J. A., Miethke, M., Pierik, A. J., Burghaus, O., Peuckert, F., Lill, R. & Marahiel, M. A. (2010). SufU is an essential iron-sulfur cluster scaffold protein in Bacillus subtilis. *Journal of bacteriology*, 192(6), 1643–1651.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410.

Anderson, A. C., Earp, B. E. & Frederick, C. A. (1996). Sequence variation as a strategy for crystallizing RNA motifs. *Journal of molecular biology*, 259(4), 696–703.

Archer, V.E., Breton, J., Sanchez-Garcia, I., Osada, H., Forster, A., Thomson, A.J. & Rabbitts, T.H. (1994). Cysteine-rich LIM domains of LIM-homeodomain and LIM-only proteins contain zinc but not iron. *Proceedings of the National Academy of Sciences*, 91(1), 316–320.

Arnott, S., Chandrasekaran, R., Millane, R. P. & Park, H.-S. (1986). DNA-RNA hybrid secondary structures. *Journal of molecular biology*, 188(4), 631–640.

Arnott, S., Chandrasekaran, R., Millane, R. P. & Park, H.-S. (1986). RNA-RNA, DNA-DNA, and DNA-RNA Polymorphism. *Biophysical journal*, 49(1), 3–5.

Asanuma, H., Hayashi, H., Zhao, J., Liang, X., Yamazawa, A., Kuramochi, T., Matsunaga, D., Aiba, Y., Kashida, H. & Komiyama, M. (2006). Enhancement of RNA cleavage activity of 10–23 DNAzyme by covalently introduced intercalator. *Chemical communications*, S. 5062–5064.

Ayala-Castro, C., Saini, A. & Outten, F. W. (2008). Fe-S cluster assembly pathways in bacteria. *Microbiol. Mol. Biol. Rev.*, 72(1), 110–125.

Bachmayer, H., Piette, L. H., Yasunobu, K. T. & Whiteley, H. R. (1967). The binding sites of iron in rubredoxin from Micrococcus aerogenes. *Proceedings of the National Academy of Sciences of the United States of America*, 57(1), 122.

Bai, Y., Chen, T., Happe, T., Lu, Y. & Sawyer, A. (2018). Iron–sulphur cluster biogenesis via the SUF pathway. *Metallomics*, 10(8), 1038–1052.

Barras, F., Loiseau, L. & Py, B. (2005). How Escherichia coli and Saccharomyces cerevisiae build Fe/S proteins. *Advances in microbial physiology*, 50, 41–101.

Basu, S., Sriram, B., Goila, R. & Banerjea, A. C. (2000). Targeted cleavage of HIV-1 coreceptor-CXCR-4 by RNA-cleaving DNA-enzyme: inhibition of coreceptor function. *Antiviral research*, 46(2), 125–134.

Baum, D. A. & Silverman, S. K. (2008). Deoxyribozymes: useful DNA catalysts in vitro and in vivo. *Cellular and molecular life sciences*, 65(14), 2156–2174.

Begley, T. P., Kinsland, C., Mehl, R. A., Osterman, A. & Dorrestein, P. (2001). The biosynthesis of nicotinamide adenine dinucleotides in bacteria. *Vitamins & Hormones*, 61, 103–119.

Beinert, H. (1990). Recent developments in the field of iron-sulfur proteins. *The FASEB Journal*, 4(8), 2483–2491.

Beinert, H. (2000). Iron-sulfur proteins: ancient structures, still full of surprises. *Journal of biological inorganic chemistry: JBIC: a publication of the Society of Biological Inorganic Chemistry*, 5(1), 2–15.

Beinert, H., Holm, R. H. & Münck, E. (1997). Iron-sulfur clusters: nature's modular, multipurpose structures. *Science*, 277(5326), 653–659.

Beinert, H. & Kennedy, M. C. (1993). Aconitase, a two-faced protein: enzyme and iron regulatory factor12. *The FASEB journal*, 7(15), 1442–1449.

Belén Moreno, M., Durán, A. & Carlos Ribas, J. (2000). A family of multifunctional thiamine-repressible expression vectors for fission yeast. *Yeast*, 16(9), 861–872.

Berman, H. M. (2000). The Protein Data Bank. *Nucleic Acids Research*, 28(1), 235–242.

Bernhart, S. H., Hofacker, I. L. & Stadler, P. F. (2006). Local RNA base pairing probabilities in large sequences. *Bioinformatics*, 22(5), 614–615.

Bernhart, S. H., Tafer, H., Mückstein, U., Flamm, C., Stadler, P. F. & Hofacker, I. L. (2006). Partition function and base pairing probabilities of RNA heterodimers. *Algorithms for Molecular Biology*, 1(1), 3.

Bhindi, R., Fahmy, R. G., Lowe, H. C., Chesterman, C. N., Dass, C. R., Cairns, M. J., Saravolac, E. G., Sun, L.-Q. & Khachigian, L. M. (2007). Brothers in arms: DNA enzymes, short interfering RNA, and the emerging wave of small-molecule nucleic acid-based gene-silencing strategies. *The American journal of pathology*, 171(4), 1079–1088.

Bilder, P. W., Ding, H. & Newcomer, M. E. (2004). Crystal structure of the ancient, Fe- S scaffold IscA reveals a novel protein fold. *Biochemistry*, 43(1), 133–139.

Black, K. A. & Dos Santos, P. C. (2015). Shared-intermediates in the biosynthesis of thio-cofactors: Mechanism and functions of cysteine desulfurases and sulfur acceptors. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1853(6), 1470–1480.

Blanc, B., Gerez, C. & de Choudens, S. O. (2015). Assembly of Fe/S proteins in bacterial systems: biochemistry of the bacterial ISC system. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1853(6), 1436–1447.

Blauenburg, B., Mielcarek, A., Altegoer, F., Fage, C. D., Linne, U., Bange, G. & Marahiel, M. A. (2016). Crystal structure of Bacillus subtilis cysteine desulfurase SufS and its dynamic interaction with frataxin and scaffold protein SufU. *PloS one*, 11(7).

Bobbin, M. L. & Rossi, J. J. (2016). RNA interference (RNAi)-based therapeutics: delivering on the promise? *Annual review of pharmacology and toxicology*, 56, 103–122.

Bolstad, H. M., Botelho, D. J. & Wood, M. J. (2010). Proteomic Analysis of Protein- Protein Interactions within the Cysteine Sulfinate Desulfinase Fe- S Cluster Biogenesis System. *Journal of proteome research*, 9(10), 5358–5369.

Bonaccio, M., Credali, A. & Peracchi, A. (2004). Kinetic and thermodynamic characterization of the RNA-cleaving 8-17 deoxyribozyme. *Nucleic acids research*, 32(3), 916–925.

Bonomi, F., Iametti, S., Morleo, A., Ta, D. & Vickery, L. E. (2011). Facilitated transfer of IscU–[2Fe2S] clusters by chaperone-mediated ligand exchange. *Biochemistry*, 50(44), 9641–9650.

Bonomi, F., Iametti, S., Morleo, A., Ta, D. R. & Vickery, L. E. (2008). Studies on the mechanism of catalysis of iron- sulfur cluster transfer from IscU [2Fe2S] by HscA/HscB chaperones. *Biochemistry*, 47(48), 12795–12801.

Bou-Abdallah, F., Adinolfi, S., Pastore, A., Laue, T. M. & Chasteen, N. D. (2004). Iron binding and oxidation kinetics in frataxin CyaY of Escherichia coli. *Journal of molecular biology*, 341(2), 605–615.

Braasch, D. A. & Corey, D. R. (2001). Locked nucleic acid (LNA): fine-tuning the recognition of DNA and RNA. *Chemistry & biology*, 8(1), 1–7.

Bradbury, A. J., Gruer, M. J., Rudd, K. E. & Guest, J. R. (1996). The second aconitase (AcnB) of Escherichia coli. *Microbiology*, 142(2), 389–400.

Bradford, N.A. (1976). A rapid and sensitive method for the quantitation microgram quantities of a protein isolated from red cell membranes. *Anal Biochem*, 72, 248–254.

Braymer, J. J. & Lill, R. (2017). Iron–sulfur cluster biogenesis and trafficking in mitochondria. *Journal of Biological Chemistry*, 292(31), 12754–12763.

Breaker, R. R. (1997). DNA enzymes. *Nature biotechnology*, 15(5), 427–431.

Breaker, R. R., Emilsson, G. M., Lazarev, D., Nakamura, S., Puskarz, I. J., Roth, A. & Sudarsan, N. (2003). A common speed limit for RNA-cleaving ribozymes and deoxyribozymes. *RNA*.

Breaker, R. R., Emilsson, G. M., Lazarev, D., Nakamura, S., Puskarz, I. J., Roth, A. & Sudarsan, N. (2003). A common speed limit for RNA-cleaving ribozymes and deoxyribozymes. *RNA*, 9(8), 949–957.

Breaker, R. R. & Joyce, G. F. (1994). A DNA enzyme that cleaves RNA. *Chemistry & biology*, 1(4), 223–229.

Breaker, R. R. & Joyce, G. F. (1995). A DNA enzyme with Mg^{2+}-dependent RNA phosphoesterase activity. *Chemistry & biology*, 2(10), 655–660.

Broderick, J. B., Duffus, B. R., Duschene, K. S. & Shepard, E. M. (2014). Radical S-adenosylmethionine enzymes. *Chemical reviews*, 114(8), 4229–4317.

Brown, I. D. (1988). What factors determine cation coordination numbers? *Acta Crystallographica Section B: Structural Science*, 44(6), 545–553.

Büeler, H., Aguzzi, A., Sailer, A., Greiner, R.-A., Autenried, P., Aguet, M. & Weissmann, C. (1993). Mice devoid of PrP are resistant to scrapie. *Cell*, 73(7), 1339–1347.

Bühning, M., Valleriani, A. & Leimkühler, S. (2017). The role of SufS is restricted to Fe–S cluster biosynthesis in Escherichia coli. *Biochemistry*, 56(14), 1987–2000.

Buis, J. M., Cheek, J., Kalliri, E. & Broderick, J. B. (2006). Characterization of an active spore photoproduct lyase, a DNA repair enzyme in the radical S-adenosylmethionine superfamily. *Journal of Biological Chemistry*, 281(36), 25994–26003.

Burke, J. E. & Butcher, S. E. (2012). Nucleic Acid Structure Characterization by Small Angle X-Ray Scattering (SAXS). *Current Protocols in Nucleic Acid Chemistry*, 51(1), 7.18.1–7.18.18.

Burkhardt, A., Pakendorf, T., Reime, B., Meyer, J., Fischer, P., Stübe, N., Panneerselvam, S., Lorbeer, O., Stachnik, K., Warmer, M., Rödig, P, Göries, D. & Meents, A. (2016). Status of the crystallography beamlines at PETRA III. *The European Physical Journal Plus*, 131(3), 56.

Cairns, M. J., Hopkins, T. M., Witherington, C. & Sun, L.-Q. (2000). The influence of arm length asymmetry and base substitution on the activity of the 10-23 DNA enzyme. *Antisense and Nucleic Acid Drug Development*, 10(5), 323–332.

Cairns, M. J., Hopkins, T. M., Witherington, C., Wang, L. & Sun, L.-Q. (1999). Target site selection for an RNA-cleaving catalytic DNA. *Nature biotechnology*, 17(5), 480–486.

Cairns, M. J., King, A. & Sun, L.-Q. (2003). Optimisation of the 10–23 DNAzyme–substrate pairing interactions enhanced RNA cleavage activity at purine–cytosine target sites. *Nucleic acids research*, 31(11), 2883–2889.

Carmi, N., Balkhi, S. R. & Breaker, R. R. (1998). Cleaving DNA with DNA. *Proceedings of the National Academy of Sciences*, 95(5), 2233–2237.

Chahal, H. K., Dai, Y., Saini, A., Ayala-Castro, C. & Outten, F. W. (2009). The SufBCD Fe- S scaffold complex interacts with SufA for Fe- S cluster transfer. *Biochemistry*, 48(44), 10644–10653.

Chandra, M., Sachdeva, A. & Silverman, S. K. (2009). DNA-catalyzed sequence-specific hydrolysis of DNA. *Nature chemical biology*, 5(10), 718–720.

Chandramouli, K. & Johnson, M. K. (2006). HscA and HscB stimulate [2Fe-2S] cluster transfer from IscU to apoferredoxin in an ATP-dependent reaction. *Biochemistry*, 45(37), 11087–11095.

Chandramouli, K., Unciuleac, M.-C., Naik, S., Dean, D. R., Huynh, B. H. & Johnson, M. K. (2007). Formation and properties of [4Fe-4S] clusters on the IscU scaffold protein. *Biochemistry*, 46(23), 6804–6811.

Chandrasekar, J. & Silverman, S. K. (2013). Catalytic DNA with phosphatase activity. *Proceedings of the National Academy of Sciences*, 110(14), 5315–5320.

Chandrasekar, J., Wylder, A. C. & Silverman, S. K. (2015). Phosphoserine lyase deoxyribozymes: DNA-catalyzed formation of dehydroalanine residues in peptides. *Journal of the American Chemical Society*, 137(30), 9575–9578.

Chen, F., Bai, M., Cao, K., Zhao, Y., Cao, X., Wei, J., Wu, N., Li, J., Wang, L., Fan, C. *et al.* (2017). Programming enzyme-initiated autonomous DNAzyme nanodevices in living cells. *ACS nano*, 11(12), 11908–11914.

Chen, F., Bai, M., Cao, K., Zhao, Y., Wei, J. & Zhao, Y. (2017). Fabricating MnO2 Nanozymes as Intracellular Catalytic DNA Circuit Generators for Versatile Imaging of Base-Excision Repair in Living Cells. *Advanced Functional Materials*, 27(45), 1702748.

Chen, F., Li, Z., Wang, R., Liu, B., Zeng, Z., Zhang, H. & Zhang, J. (2004). Inhibition of ampicillin-resistant bacteria by novel mono-DNAzymes and di-DNAzyme targeted to β-lactamase mRNA. *Oligonucleotides*, 14(2), 80–89.

Chen, Y., Chen, L., Ou, Y., Wang, Z., Fu, F. & Guo, L. (2016). DNAzyme-based biosensor for Cu2+ ion by combining hybridization chain reaction with fluorescence resonance energy transfer technique. *Talanta*, 155, 245–249.

Chen, Y., Zhang, J., Liu, H., Gao, Y., Li, X., Zheng, L., Cui, R., Yao, Q., Rong, L., Li, J. *et al.* (2017). Unique 5'-P recognition and basis for dG: dGTP misincorporation of ASFV DNA polymerase X. *PLoS biology*, 15(2).

Cheng, R. P., Fisher, S. L. & Imperiali, B. (1996). Metallopeptide design: tuning the metal cation affinities with unnatural amino acids and peptide secondary structure. *Journal of the American Chemical Society*, 118(46), 11349–11356.

Chi, Y.-I., Martick, M., Lares, M., Kim, R., Scott, W. G. & Kim, S.-H. (2008). Capturing hammerhead ribozyme structures in action by modulating general base catalysis. *PLoS biology*, 6(9).

Chinnapen, D. J.-F. & Sen, D. (2004). A deoxyribozyme that harnesses light to repair thymine dimers in DNA. *Proceedings of the National Academy of Sciences*, 101(1), 65–69.

Chiuman, W. & Li, Y. (2007). Efficient signaling platforms built from a small catalytic DNA and doubly labeled fluorogenic substrates. *Nucleic acids research*, 35(2), 401–405.

Cho, E.-A., Moloney, F. J., Cai, H., Au-Yeung, A., China, C., Scolyer, R. A., Yosufi, B., Raftery, M. J., Deng, J. Z., Morton, S. W. *et al.* (2013). Safety and tolerability of an intratumorally injected DNAzyme, Dz13, in patients with nodular basal-cell carcinoma: a phase 1 first-in-human trial (DISCOVER). *The Lancet*, 381(9880), 1835–1843.

Choe, Y. J., Han, H. J., Lee, J. H., Seo, S. W. & Choe, B. S. (2000). Synthesis and NMR Studies of the RNA-Cleaving DNA Enzyme. *Bulletin of the Korean Chemical Society*, 21(10), 1025–1030.

Choi, J. H., Williams, J., Cho, J., Falck, J.R. & Shears, S. B. (2007). Purification, sequencing, and molecular identification of a mammalian PP-InsP5 kinase that is activated when cells are exposed to hyperosmotic stress. *Journal of Biological Chemistry*, 282(42), 30763–30775.

Chu, C.-C., Wong, O. Y. & Silverman, S. K. (2014). A Generalizable DNA-Catalyzed Approach to Peptide–Nucleic Acid Conjugation. *ChemBioChem*, 15(13), 1905–1910.

Cicchillo, R. M., Tu, L., Stromberg, J. A., Hoffart, L. M., Krebs, C. & Booker, S. J. (2005). Escherichia coli quinolinate synthetase does indeed harbor a [4Fe-4S] cluster. *Journal of the American Chemical Society*, 127(20), 7310–7311.

Cieslak, M., Szymanski, J., Adamiak, R. W. & Cierniewski, C. S. (2003). Structural rearrangements of the 10–23 DNAzyme to β3 integrin subunit mRNA induced by cations and their relations to the catalytic activity. *Journal of Biological Chemistry*, 278(48), 47987–47996.

Cochrane, Jesse C, Lipchock, Sarah V & Strobel, Scott A (2007). Structural investigation of the GlmS ribozyme bound to its catalytic cofactor. *Chemistry & biology*, 14(1), 97–105.

Correll, C. C., Wool, I. G. & Munishkin, A. (1999). The two faces of the Escherichia coli 23 S rRNA sarcin/ricin domain: the structure at 1.11 Å resolution. *Journal of molecular biology*, 292(2), 275–287.

Creutzfeldt, H. G. (1920). Über eine eigenartige herdförmige erkrankung des zentralnervensystems (vorläufige mitteilung). *Zeitschrift für die gesamte Neurologie und Psychiatrie*, 57(1), 1–18.

Cuchillo, C. M., Nogués, M. V. & Raines, R. T. (2011). Bovine pancreatic ribonuclease: fifty years of the first enzymatic reaction mechanism. *Biochemistry*, 50(37), 7835–7841.

Cuenoud, B. & Szostak, J. W. (1995). A DNA metalloenzyme with DNA ligase activity. *Nature*, 375(6532), 611–614.

Cunningham, L., Gruer, M. J. & Guest, J. R. (1997). Transcriptional regulation of the aconitase genes (acnA and acnB) of Escherichia coli. *Microbiology*, 143(12), 3795–3805.

Cupp-Vickery, J. R., Urbina, H. & Vickery, L. E. (2003). Crystal structure of IscS, a cysteine desulfurase from Escherichia coli. *Journal of molecular biology*, 330(5), 1049–1059.

Curtis, E. A. & Bartel, D. P. (2001). The hammerhead cleavage reaction in monovalent cations. *Rna*, 7(4), 546–552.

Dahm, S. C., Derrick, W. B. & Uhlenbeck, O. C. (1993). Evidence for the role of solvated metal hydroxide in the hammerhead cleavage mechanism. *Biochemistry*, 32(48), 13040–13045.

Dailey, H. A., Finnegan, M. G. & Johnson, M. K. (1994). Human ferrochelatase is an iron-sulfur protein. *Biochemistry*, 33(2), 403–407.

Dass, C. R., Saravolac, E. G., Li, Y. & Sun, L.-Q. (2002). Cellular uptake, distribution, and stability of 10-23 deoxyribozymes. *Antisense and Nucleic Acid Drug Development*, 12(5), 289–299.

Dean, D. R., Bolin, J. T. & Zheng, L. (1993). Nitrogenase metalloclusters: structures, organization, and synthesis. *Journal of bacteriology*, 175(21), 6737.

Decamps, L., Philmus, B., Benjdia, A., White, R., Begley, T. P & Berteau, O. (2012). Biosynthesis of F0, precursor of the F420 cofactor, requires a unique two radical-SAM domain enzyme and tyrosine as substrate. *Journal of the American Chemical Society*, 134(44), 18173–18176.

delCardayre, S. B. & Raines, R. T. (1994). Structural determinants of enzymic processivity. *Biochemistry*, 33(20), 6031–6037.

Ding, H. & Clark, R. J. (2004). Characterization of iron binding in IscA, an ancient iron-sulphur cluster assembly protein. *Biochemical Journal*, 379(2), 433–440.

Ding, H., Harrison, K. & Lu, J. (2005). Thioredoxin reductase system mediates iron binding in IscA and iron delivery for the iron-sulfur cluster assembly in IscU. *Journal of Biological Chemistry*, 280(34), 30432–30437.

Ding, H., Hidalgo, E. & Demple, B. (1996). The redox state of the [2Fe-2S] clusters in SoxR protein regulates its activity as a transcription factor. *Journal of Biological Chemistry*, 271(52), 33173–33175.

Ding, H., Yang, J., Coleman, L. C. & Yeung, S. (2007). Distinct Iron Binding Property of Two Putative Iron Donors for the Iron-Sulfur Cluster Assembly IscA and the Bacterial Frataxin Ortholog CyaY Under Physiological and Oxidative Stress Conditions. *Journal of Biological Chemistry*, 282(11), 7997–8004.

Doherty, E. A. & Doudna, J. A. (2001). Ribozyme structures and mechanisms. *Annual review of biophysics and biomolecular structure*, 30(1), 457–475.

Dollins, D. E., Bai, W., Fridy, P. C., Otto, J. C., Neubauer, J. L., Gattis, S. G., Mehta, K. P.M. & York, J. D. (2020). Vip1 is a kinase and pyrophosphatase switch that regulates inositol diphosphate signaling. *Proceedings of the National Academy of Sciences*, 117(17), 9356–9364.

Dolot, R., Sobczak, M., Mikołajczyk, B. & Nawrot, B. (2017). Synthesis, crystallization and preliminary crystallographic analysis of a 52-nucleotide DNA/2'-OMe-RNA oligomer mimicking 10–23 DNAzyme in the complex with a substrate. *Nucleosides, Nucleotides and Nucleic Acids*, 36(4), 292–301.

Donghi, D. & Schnabl, J. (2011). Multiple roles of metal ions in large ribozymes. *Met. Ions Life Sci*, 9, 197–234.

Dos Santos, P. C. (2017). 5 Fe-S assembly in Gram-positive bacteria. *Biochemistry, Biosynthesis and Human Diseases*.

Dos Santos, P. C. (2017). B. subtilis as a model for studying the assembly of Fe–S clusters in Gram-positive bacteria. In *Methods in enzymology*, Band 595, S. 185–212. Elsevier.

Dunham, W.R., Harding, L.J. & Sands, R.H. (1994). Mössbauer spectroscopy of metalloproteins and the use of Fourier transforms. In *Ejb Reviews 1993*, S. 77–84. Springer.

Duschene, K. S. & Broderick, J. B. (2010). The antiviral protein viperin is a radical SAM enzyme. *FEBS letters*, 584(6), 1263–1267.

Eaton, G. R., Eaton, S. S., Barr, D. P. & Weber, R. T. (2010). *Quantitative EPR, a practitioner guide*. Springer.

Ellington, A. D. & Szostak, J. W. (1990). In vitro selection of RNA molecules that bind specific ligands. *nature*, 346(6287), 818–822.

Emilsson, G. M., Nakamura, S., Roth, A. & Breaker, R. R. (2003). Ribozyme speed limits. *Rna*, 9(8), 907–918.

Emsley, P. & Cowtan, K. (2004). Coot : model-building tools for molecular graphics. *Acta Crystallographica Section D Structural Biology*, 60(12), 2126–2132.

Famulok, M. & Jenne, A. (1999). Catalysis based on nucleic acid structures. In *Implementation and Redesign of Catalytic Function in Biopolymers*, S. 101–131. Springer.

Fan, H., Zhao, Z., Yan, G., Zhang, X., Yang, C., Meng, H., Chen, Z., Liu, H. & Tan, W. (2015). A smart DNAzyme–MnO2 nanosystem for efficient gene silencing. *Angewandte Chemie International Edition*, 54(16), 4801–4805.

Fan, X., Li, H., Zhao, J., Lin, F., Zhang, L., Zhang, Y. & Yao, S. (2012). A novel label-free fluorescent sensor for the detection of potassium ion based on DNAzyme. *Talanta*, 89, 57–62.

Faulhammer, D. & Famulok, M. (1996). The Ca2+ ion as a cofactor for a novel RNA-cleaving deoxyribozyme. *Angewandte Chemie International Edition in English*, 35(23-24), 2837–2841.

Fedor, M. J. (2002). The role of metal ions in RNA catalysis. *Current opinion in structural biology*, 12(3), 289–295.

Fedor, M. J. & Uhlenbeck, O. C. (1992). Kinetics of intermolecular cleavage by hammerhead ribozymes. *Biochemistry*, 31(48), 12042–12054.

Fedoruk-Wyszomirska, A., Szymański, M., Wyszko, E., Barciszewska, M. Z & Barciszewski, J. (2009). Highly active low magnesium hammerhead ribozyme. *Journal of biochemistry*, 145(4), 451–459.

Fee, J. A., Findling, K.L., Yoshida, T., Hille, R., Tarr, G.E., Hearshen, D.O., Dunham, W.R., Day, E.P., Kent, T.A. & Münck, E. (1984). Purification and characterization of the Rieske iron-sulfur protein from Thermus thermophilus. Evidence for a [2Fe-2S] cluster having non-cysteine ligands. *Journal of Biological Chemistry*, 259(1), 124–133.

Feldman, A. R. & Sen, D. (2001). A new and efficient DNA enzyme for the sequence-specific cleavage of RNA. *Journal of molecular biology*, 313(2), 283–294.

Feoktistova, A., McCollum, D., Ohi, R. & Gould, K. L. (1999). Identification and characterization of Schizosaccharomyces pombe asp1+, a gene that interacts with mutations in the Arp2/3 complex and actin. *Genetics*, 152(3), 895–908.

Ferré-D'Amaré, A. R. (2010). Use of the spliceosomal protein U1A to facilitate crystallization and structure determination of complex RNAs. *Methods*, 52(2), 159–167.

Ferré-D'Amaré, A. R. (2016). Use of the U1A Protein to Facilitate Crystallization and Structure Determination of Large RNAs. In *Nucleic Acid Crystallography*, S. 67–76. NIH Public Access.

Ferré-D'Amaré, A. R. & Doudna, J. A. (2000). Crystallization and structure determination of a hepatitis delta virus ribozyme: use of the RNA-binding protein U1A as a crystallization module. *Journal of Molecular Biology*, 295(3), 541–556.

Ferré-D'Amaré, A. R. & Doudna, J. A. (2000). Methods to Crystallize RNA. *Current Protocols in Nucleic Acid Chemistry*, 1(1), 7.6.1–7.6.13.

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

Flint, D. H (1996). Escherichia coli contains a protein that is homologous in function and N-terminal sequence to the protein encoded by the nifS gene of Azotobacter vinelandii and that can participate in the synthesis of the Fe-S cluster of dihydroxy-acid dehydratase. *Journal of Biological Chemistry*, 271(27), 16068–16074.

Flühe, L., Burghaus, O., Wieckowski, B. M., Giessen, T. W., Linne, U. & Marahiel, M. A. (2013). Two [4Fe-4S] clusters containing radical SAM enzyme SkfB catalyze thioether bond formation during the maturation of the sporulation killing factor. *Journal of the American Chemical Society*, 135(3), 959–962.

Flühe, L., Knappe, T. A., Gattner, M. J., Schäfer, A., Burghaus, O., Linne, U. & Marahiel, M. A. (2012). The radical SAM enzyme AlbA catalyzes thioether bond formation in subtilosin A. *Nature chemical biology*, 8(4), 350–357.

Flühe, L. & Marahiel, M. A. (2013). Radical S-adenosylmethionine enzyme catalyzed thioether bond formation in sactipeptide biosynthesis. *Current opinion in chemical biology*, 17(4), 605–612.

Flynn-Charlebois, A., Wang, Y., Prior, T. K., Rashid, I., Hoadley, K. A., Coppins, R. L., Wolf, A. C. & Silverman, S. K. (2003). Deoxyribozymes with 2 ‘- 5 ‘RNA ligase activity. *Journal of the American Chemical Society*, 125(9), 2444–2454.

Fokina, A. A., Meschaninova, M. I., Durfort, T., Venyaminova, A. G. & François, Jean-Christophe (2012). Targeting insulin-like growth factor I with 10–23 DNAzymes: 2’-O-methyl modifications in the catalytic core enhance mRNA cleavage. *Biochemistry*, 51(11), 2181–2191.

Fokina, A. A., Stetsenko, D. A. & François, J.-C. (2015). DNA enzymes as potential therapeutics: towards clinical application of 10-23 DNAzymes. *Expert opinion on biological therapy*, 15(5), 689–711.

Fontecave, M. & Ollagnier-de Choudens, S. (2008). Iron–sulfur cluster biosynthesis in bacteria: mechanisms of cluster assembly and transfer. *Archives of biochemistry and biophysics*, 474(2), 226–237.

Fontecave, M., Ollagnier De Choudens, S., Py, B. & Barras, F. (2005). Mechanisms of iron–sulfur cluster assembly: the SUF machinery. *JBIC Journal of Biological Inorganic Chemistry*, 10(7), 713–721.

Forconi, M. & Herschlag, D. (2009). Metal ion-based RNA cleavage as a structural probe. *Meth. Enzymol.*, 468, 91–106.

Forlani, F., Cereda, A., Freuer, A., Nimtz, M., Leimkühler, S. & Pagani, S. (2005). The cysteine-desulfurase IscS promotes the production of the rhodanese RhdA in the persulfurated form. *FEBS letters*, 579(30), 6786–6790.

Forster, A. C. & Symons, R. H. (1987). Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites. *Cell*, 49(2), 211–220.

François, B., Lescoute-Phillips, A., Werner, A. & Masquida, B. (2005). Preparation and handling of RNA crystals. *Handbook of RNA Biochemistry*, S. 438–452.

Franke, D., Petoukhov, M. V., Konarev, P. V., Panjkovich, A., Tuukkanen, A., Mertens, H. D. T., Kikhney, A. G., Hajizadeh, N. R., Franklin, J. M., Jeffries, C. M. & Svergun, D. I. (2017). ATSAS 2.8 : a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *Journal of Applied Crystallography*, 50(4), 1212–1225.

Franke, D. & Svergun, D. I. (2009). DAMMIF , a program for rapid ab-initio shape determination in small-angle scattering. *Journal of Applied Crystallography*, 42(2), 342–346.

Freibert, S.-A., Weiler, B. D., Bill, E., Pierik, A. J., Mühlenhoff, U. & Lill, R. (2018). Biochemical Reconstitution and Spectroscopic Analysis of Iron–Sulfur Proteins. In *Methods in enzymology*, Band 599, S. 197–226. Elsevier.

Freisinger, E. & Sigel, R. K. O. (2007). From nucleotides to ribozymes—a comparison of their metal ion binding properties. *Coordination Chemistry Reviews*, 251(13-14), 1834–1851.

Frey, P. A., Hegeman, A. D. & Ruzicka, F. J. (2008). The radical SAM superfamily. *Critical reviews in biochemistry and molecular biology*, 43(1), 63–88.

Fridy, P. C., Otto, J. C., Dollins, D. E. & York, J. D. (2007). Cloning and characterization of two human VIP1-like inositol hexakisphosphate and diphosphoinositol pentakisphosphate kinases. *Journal of Biological Chemistry*, 282(42), 30754–30762.

Fu, S. & Sun, L.-Q. (2015). DNAzyme-based therapeutics for cancer treatment. *Future medicinal chemistry*, 7(13), 1701–1707.

Fu, W., Jack, R. F., Morgan, T. V., Dean, D. R. & Johnson, M. K. (1994). nifU gene product from Azotobacter vinelandii is a homodimer that contains two identical [2Fe-2S] clusters. *Biochemistry*, 33(45), 13455–13463.

Füzéry, A. K., Tonelli, M., Ta, D. T., Cornilescu, G., Vickery, L. E. & Markley, J. L. (2008). Solution structure of the iron- sulfur cluster cochaperone HscB and its binding surface for the iron- sulfur assembly scaffold protein IscU. *Biochemistry*, 47(36), 9394–9404.

Gajdusek, D. C. & Zigas, V. (1957). Degenerative disease of the central nervous system in New Guinea: the endemic occurrence of kuru in the native population. *New England Journal of Medicine*, 257(20), 974–978.

Galle, L. M., Cutsail III, G. E., Nischwitz, V., DeBeer, S. & Span, I. (2018). Spectroscopic characterization of the Co-substituted C-terminal domain of rubredoxin-2. *Biological chemistry*, 399(7), 787–798.

Gambetti, P., Dong, Z., Yuan, J., Xiao, X., Zheng, M., Alshekhlee, A., Castellani, R., Cohen, M., Barria, M. A., Gonzalez-Romero, D. *et al.* (2008). A novel human disease with abnormal prion protein sensitive to protease. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 63(6), 697–708.

Gao, H., Subramanian, S., Couturier, J., Naik, S. G., Kim, S.-K., Leustek, T., Knaff, D. B., Wu, H.-C., Vignols, F., Huynh, B. H. *et al.* (2013). Arabidopsis thaliana Nfu2 accommodates [2Fe-2S] or [4Fe-4S] clusters and is competent for in vitro maturation of chloroplast [2Fe-2S] and [4Fe-4S] cluster-containing proteins. *Biochemistry*, 52(38), 6633–6645.

George, G. N., Pickering, I. J., Prince, R. C., Zhou, Z. H. & Adams, M. W. W. (1996). X-ray absorption spectroscopy of Pyrococcus furiosus rubredoxin. *JBIC Journal of Biological Inorganic Chemistry*, 1(3), 226–230.

Geyer, C. R. & Sen, D. (1997). Evidence for the metal-cofactor independence of an RNA phosphodiester-cleaving DNA enzyme. *Chemistry & biology*, 4(8), 579–593.

Giel, J. L., Nesbit, A. D., Mettert, E. L., Fleischhacker, A. S., Wanta, B. T. & Kiley, P. J. (2013). Regulation of iron–sulphur cluster homeostasis through transcriptional control of the Isc pathway by [2 Fe–2 S]–IscR in E scherichia coli. *Molecular microbiology*, 87(3), 478–492.

Giel, J. L., Rodionov, D., Liu, M., Blattner, F. R. & Kiley, P. J. (2006). IscR-dependent gene expression links iron-sulphur cluster assembly to the control of O2-regulated genes in Escherichia coli. *Molecular microbiology*, 60(4), 1058–1075.

Goila, R. & Banerjea, A. C. (1998). Sequence specific cleavage of the HIV-1 coreceptor CCR5 gene by a hammer-head ribozyme and a DNA-enzyme: inhibition of the coreceptor function by DNA-enzyme. *FEBS letters*, 436(2), 233–238.

Golden, B. L. & Kundrot, C. E. (2003). RNA crystallization. *Journal of structural biology*, 142(1), 98–107.

Golinelli, M.-P., Akin, L. A., Crouse, B. R., Johnson, M. K. & Meyer, J. (1996). Cysteine ligand swapping on a deletable loop of the [2Fe-2S] ferredoxin from Clostridium pasteurianum. *Biochemistry*, 35(27), 8995–9002.

Gong, L., Zhao, Z., Lv, Y.-F., Huan, S.-Y., Fu, T., Zhang, X.-B., Shen, G.-L. & Yu, R.-Q. (2015). DNAzyme-based biosensors and nanodevices. *Chemical Communications*, 51(6), 979–995.

Goodman, R. P., Schaap, I. A. T., Tardin, C. F., Erben, C. M., Berry, R. M., Schmidt, C. F. & Turberfield, A. J. (2005). Rapid chiral assembly of rigid DNA building blocks for molecular nanofabrication. *Science*, 310(5754), 1661–1665.

Gräwert, T., Kaiser, J., Zepeck, F., Laupitz, R., Hecht, S., Amslinger, S., Schramek, N., Schleicher, E., Weber, S., Haslbeck, M. *et al.* (2004). IspH Protein of Escherichia c oli: Studies on Iron- Sulfur Cluster Implementation and Catalysis. *Journal of the American Chemical Society*, 126(40), 12847–12855.

Gräwert, T., Rohdich, F., Span, I., Bacher, A, Eisenreich, W., Eppinger, J. & Groll, M. (2009). Structure of active IspH enzyme from Escherichia coli provides mechanistic insights into substrate reduction. *Angewandte Chemie International Edition*, 48(31), 5756–5759.

Gräwert, T., Span, I., Eisenreich, W., Rohdich, F., Eppinger, J., Bacher, A. & Groll, M. (2010). Probing the reaction mechanism of IspH protein by x-ray structure analysis. *Proceedings of the National Academy of Sciences*, 107(3), 1077–1081.

Grove, T. L., Himes, P. M., Hwang, S., Yumerefendi, H., Bonanno, J. B., Kuhlman, B., Almo, S. C. & Bowers, A. A. (2017). Structural insights into thioether bond formation in the biosynthesis of sactipeptides. *Journal of the American Chemical Society*, 139(34), 11734–11744.

Gruenewald, B., Nicola, C. U., Lustig, A., Schwarz, G. & Klump, H. (1979). Kinetics of the helix—coil transition of a polypeptide with non-ionic side groups, derived from ultrasonic relaxation measurements. *Biophysical chemistry*, 9(2), 137–147.

Gruer, M. J., Bradbury, A. J. & Guest, J. R. (1997). Construction and properties of aconitase mutants of Escherichia coli. *Microbiology*, 143(6), 1837–1846.

Gu, C., Nguyen, H.-N., Ganini, D., Chen, Z., Jessen, H. J., Gu, Z., Wang, H. & Shears, S. B. (2017). KO of 5-InsP7 kinase activity transforms the HCT116 colon cancer cell line into a hypermetabolic, growth-inhibited phenotype. *Proceedings of the National Academy of Sciences*, 114(45), 11968–11973.

Gu, H., Furukawa, K., Weinberg, Z., Berenson, D. F. & Breaker, R. R. (2013). Small, highly active DNAs that hydrolyze DNA. *Journal of the American Chemical Society*, 135(24), 9121–9129.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, 35(3), 849–857.

Gupta, V., Sendra, M., Naik, S. G., Chahal, H. K., Huynh, B. H., Outten, F. W., Fontecave, M. & Ollagnier de Choudens, S. (2009). Native Escherichia coli SufA, coexpressed with SufBCDSE, purifies as a [2Fe- 2S] protein and acts as an Fe- S transporter to Fe- S target enzymes. *Journal of the American Chemical Society*, 131(17), 6149–6153.

Hadjiolov, A. A., Venkov, P. V. & Tsanev, R. G. (1966). Ribonucleic acids fractionation by density-gradient centrifugation and by agar gel electrophoresis: a comparison. *Analytical biochemistry*, 17(2), 263–267.

Hamilton, A. J. & Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*, 286(5441), 950–952.

Hampel, A. & Tritz, R. (1989). RNA catalytic properties of the minimum (-) sTRSV sequence. *Biochemistry*, 28(12), 4929–4933.

Hanna, R. & Doudna, J. A. (2000). Metal ions in ribozyme folding and catalysis. *Current opinion in chemical biology*, 4(2), 166–170.

Hänzelmann, P., Hernández, H. L., Menzel, C., García-Serres, R., Huynh, B. H., Johnson, M. K., Mendel, R. R. & Schindelin, H. (2004). Characterization of MOCS1A, an oxygen-sensitive iron-sulfur protein involved in human molybdenum cofactor biosynthesis. *Journal of Biological Chemistry*, 279(33), 34721–34732.

Hassell, A. M., An, G., Bledsoe, R. K., Bynum, J. M., Carter, H. L., Deng, S.-J. J., Gampe, R. T., Grisard, T. E., Madauss, K. P., Nolte, R. T., Rocque, W. J., Wang, L., Weaver, K. L., Williams, S. P., Wisely, G. B., Xu, R. & Shewchuk, L. M. (2007). Crystallization of protein–ligand complexes. *Acta Crystallographica Section D Structural Biology*, 63(1), 72–79.

He, D., He, X., Yang, X. & Li, H.-W. (2017). A smart ZnO polydopamine-nucleic acid nanosystem for ultrasensitive live cell mRNA imaging by the target-triggered intracellular self-assembly of active DNAzyme nanostructures. *Chemical science*, 8(4), 2832–2840.

He, J., Zhang, D., Wang, Q., Wei, X., Cheng, M. & Liu, K. (2011). A novel strategy of chemical modification for rate enhancement of 10–23 DNAzyme: a combination of A9 position and 8-aza-7-deaza-2'-deoxyadenosine analogs. *Organic & biomolecular chemistry*, 9(16), 5728–5736.

He, Q.-C., Zhou, J.-M., Zhou, D.-M., Nakamatsu, Y., Baba, T. & Taira, K. (2002). Comparison of metal-ion-dependent cleavages of RNA by a DNA enzyme and a hammerhead ribozyme. *Biomacromolecules*, 3(1), 69–83.

He, S., Qu, L., Shen, Z., Tan, Y., Zeng, M., Liu, F., Jiang, Y. & Li, Y. (2015). Highly specific recognition of breast tumors by an RNA-cleaving fluorogenic DNAzyme probe. *Analytical chemistry*, 87(1), 569–577.

Higgs, P. G. & Lehman, N. (2015). The RNA World: molecular cooperation at the origins of life. *Nature Reviews Genetics*, 16(1), 7–17.

Hirabayashi, K., Yuda, E., Tanaka, N., Katayama, S., Iwasaki, K., Matsumoto, T., Kurisu, G., Outten, F. W., Fukuyama, K., Takahashi, Y. *et al.* (2015). Functional dynamics revealed by the structure of the SufBCD complex, a novel ATP-binding cassette (ABC) protein that serves as a scaffold for iron-sulfur cluster biogenesis. *Journal of Biological Chemistry*, 290(50), 29717–29731.

Hoagland, M. B., Stephenson, M. L., Scott, J. F., Hecht, L. I. & Zamecnik, P. C. (1958). A soluble ribonucleic acid intermediate in protein synthesis. *Journal of Biological Chemistry*, 231(1), 241–257.

Hoff, K. G., Cupp-Vickery, J. R. & Vickery, L. E. (2003). Contributions of the LPPVK motif of the iron-sulfur template protein IscU to interactions with the Hsc66-Hsc20 chaperone system. *Journal of Biological Chemistry*, 278(39), 37582–37589.

Hoffman, B. M., Lukoyanov, D., Dean, D. R. & Seefeldt, L. C. (2013). Nitrogenase: a draft mechanism. *Accounts of chemical research*, 46(2), 587–595.

Holbrook, E. L. & Holbrook, S. R. (2001). *Crystallisation of nucleic acids*. Wiley Online Library.

Holden, H. M., Jacobson, B. L., Hurley, J. K., Tollin, G., Oh, B.-H., Skjeldal, L., Chae, Y. K., Cheng, H., Xia, B. & Markley, J. L. (1994). Structure-function studies of [2Fe-2S] ferredoxins. *Journal of bioenergetics and biomembranes*, 26(1), 67–88.

Holland, I. B. & Blight, M. A. (1999). ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. *Journal of molecular biology*, 293(2), 381–399.

Hollenstein, M. (2015). DNA catalysis: the chemical repertoire of DNAzymes. *Molecules*, 20(11), 20777–20804.

Hollenstein, M., Hipolito, C., Lam, C., Dietrich, D. & Perrin, D. M. (2008). A highly selective DNAzyme sensor for mercuric ions. *Angewandte Chemie International Edition*, 47(23), 4346–4350.

Hollenstein, M., Hipolito, C. J., Lam, C. H. & Perrin, D. M. (2009). A self-cleaving DNA enzyme modified with amines, guanidines and imidazoles operates independently of divalent metal cations (M 2+). *Nucleic acids research*, 37(5), 1638–1649.

Horn, W. T., Convery, M. A., Stonehouse, N. J., Adams, C. J., Liljas, L., Phillips, S. E.V. & Stockley, P. G. (2004). The crystal structure of a high affinity RNA stem-loop complexed with the bacteriophage MS2 capsid: Further challenges in the modeling of ligand-RNA interactions. *RNA*, 10(11), 1776–1782.

Horton, T. E., Clardy, D. R. & DeRose, V. J. (1998). Electron paramagnetic resonance spectroscopic measurement of Mn2+ binding affinities to the hammerhead ribozyme and correlation with cleavage activity. *Biochemistry*, 37(51), 18094–18101.

Hu, Y. & Ribbe, M. W. (2016). Biosynthesis of the metalloclusters of nitrogenases. *Annual review of biochemistry*, 85, 455–483.

Huang, P. J. & Liu, J. (2015). Rational evolution of Cd2+-specific DNAzymes with phosphorothioate modified cleavage junction and Cd2+ sensing. *Nucleic acids research*, 43(12), 6125–6133.

Huang, P.-J. J., Lin, J., Cao, J., Vazin, M. & Liu, J. (2014). Ultrasensitive DNAzyme beacon for lanthanides and metal speciation. *Analytical chemistry*, 86(3), 1816–1821.

Huang, P.-J. J. & Liu, J. (2016). An ultrasensitive light-up Cu2+ biosensor using a new DNAzyme cleaving a phosphorothioate-modified substrate. *Analytical chemistry*, 88(6), 3341–3347.

Huang, P.-J. J., Vazin, M. & Liu, J. (2015). Desulfurization activated phosphorothioate DNAzyme for the detection of thallium. *Analytical chemistry*, 87(20), 10443–10449.

Huang, P.-J. J., Vazin, M. & Liu, J. (2016). In vitro selection of a DNAzyme cooperatively binding two lanthanide ions for RNA cleavage. *Biochemistry*, 55(17), 2518–2525.

Huet, G., Daffé, M. & Saves, I. (2005). Identification of the Mycobacterium tuberculosis SUF machinery as the exclusive mycobacterial system of [Fe-S] cluster assembly: evidence for its implication in the pathogen's survival. *Journal of bacteriology*, 187(17), 6137–6146.

Hunsicker, L. M. & DeRose, V. J. (2000). Activities and relative affinities of divalent metals in unmodified and phosphorothioate-substituted hammerhead ribozymes. *Journal of inorganic biochemistry*, 80(3-4), 271–281.

Hunsicker-Wang, L., Vogt, M. & DeRose, V. J. (2009). EPR methods to study specific metal-ion binding sites in RNA. In *Methods in enzymology*, Band 468, S. 335–367. Elsevier.

Hwang, D. M., Dempsey, A., Tan, K.-T. & Liew, C.-C. (1996). A modular domain of NifU, a nitrogen fixation cluster protein, is highly conserved in evolution. *Journal of molecular evolution*, 43(5), 536–540.

Hwang, K., Wu, P., Kim, T., Lei, L., Tian, S., Wang, Y. & Lu, Y. (2014). Photocaged DNAzymes as a general method for sensing metal ions in living cells. *Angewandte Chemie International Edition*, 53(50), 13798–13802.

Iannuzzi, C., Adinolfi, S., Howes, B. D., Garcia-Serres, R., Clémancey, M., Latour, J.-M., Smulevich, G. & Pastore, A. (2011). The role of CyaY in iron sulfur cluster assembly on the E. coli IscU scaffold protein. *PloS one*, 6(7).

Jacquot, Jean-Pierre (2017). Iron–Sulfur Clusters in Chemistry and Biology. Volume 1: Characterization, Properties and Applications. Second edition. Edited by Tracey Rouault. De Gruyter, 2017. *Acta Crystallographica Section D*, 73(12), 1030.

Jakob, A. (1921). Über eigenartige Erkrankungen des Zentralnervensystems mit bemerkenswertem anatomischen Befunde. *Zeitschrift für die gesamte Neurologie und Psychiatrie*, 64(1), 147–228.

Jakobsen, M. R., Haasnoot, J., Wengel, J., Berkhout, B. & Kjems, J. (2007). Efficient inhibition of HIV-1 expression by LNA modified antisense oligonucleotides and DNAzymes targeted to functionally selected binding sites. *Retrovirology*, 4(1), 29.

Jakopec, V., Walla, E. & Fleig, U. (2011). Versatile use of Schizosaccharomyces pombe plasmids in Saccharomyces cerevisiae. *FEMS yeast research*, 11(8), 653–655.

Järup, L. (2003). Hazards of heavy metal contamination. *British medical bulletin*, 68(1), 167–182.

Jäschke, A. (2001). Artificial ribozymes and deoxyribozymes. *Current opinion in structural biology*, 11(3), 321–326.

Jasniewski, A., Hu, Y. & Ribbe, M. W. (2019). Electron Paramagnetic Resonance Spectroscopy of Metalloproteins. In *Metalloproteins*, S. 197–211. Springer.

Jensen, L. T. & Culotta, V. C. (2000). Role of Saccharomyces cerevisiae ISA1and ISA2 in iron homeostasis. *Molecular and cellular biology*, 20(11), 3918–3927.

Jiang, P. & Guo, Z. (2004). Fluorescent detection of zinc in biological systems: recent development on the design of chemosensors and biosensors. *Coordination Chemistry Reviews*, 248(1-2), 205–229.

Johnson, D. C., Dean, D. R., Smith, A. D. & Johnson, M. K. (2005). Structure, function, and formation of biological iron-sulfur clusters. *Annu. Rev. Biochem.*, 74, 247–281.

Jordan, P. A, Tang, Y., Bradbury, A. J., Thomson, A. J. & Guest, J. R. (1999). Biochemical and spectroscopic characterization of Escherichia coli aconitases (AcnA and AcnB). *Biochemical Journal*, 344(3), 739–746.

Jordan, S. R., Whitcombe, T. V., Berg, J. M. & Pabo, C. O. (1985). Systematic variation in DNA length yields highly ordered repressor-operator cocrystals. *Science*, 230(4732), 1383–1385.

Joyce, G. (2001). RNA cleavage by the 10–23 DNA enzyme. *Methods in Enzymology - METH ENZYMOLOGY*, 341, 503–517.

Joyce, G. F. (2007). Forty years of in vitro evolution. *Angewandte Chemie International Edition*, 46(34), 6420–6436.

Kakuta, Y., Horio, T., Takahashi, Y. & Fukuyama, K. (2001). Crystal Structure of Escherichia coli Fdx, an Adrenodoxin-Type Ferredoxin Involved in the Assembly of Iron- Sulfur Clusters. *Biochemistry*, 40(37), 11007–11012.

Kampinga, H. H. & Craig, E. A. (2010). The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nature reviews Molecular cell biology*, 11(8), 579–592.

Kasprowicz, A., Stokowa-Sołtys, K., Jeżowska-Bojczuk, M., Wrzesiński, J. & Ciesiołka, J. (2017). Characterization of Highly Efficient RNA-Cleaving DNAzymes that Function at Acidic pH with No Divalent Metal-Ion Cofactors. *ChemistryOpen*, 6(1), 46–56.

Keefe, A. D. & Cload, S. T. (2008). SELEX with modified nucleotides. *Current opinion in chemical biology*, 12(4), 448–456.

Keel, A. Y., Rambo, R. P., Batey, R. T. & Kieft, J. S. (2007). A general strategy to solve the phase problem in RNA crystallography. *Structure*, 15(7), 761–772.

Keiper, S. & Vyle, J. S. (2006). Reversible photocontrol of deoxyribozyme-catalyzed RNA cleavage under multiple-turnover conditions. *Angewandte Chemie International Edition*, 45(20), 3306–3309.

Kenward, M. & Dorfman, K. D. (2009). Coarse-grained brownian dynamics simulations of the 10-23 dnazyme. *Biophysical journal*, 97(10), 2785–2793.

Kessler, D. (2006). Enzymatic activation of sulfur for incorporation into biomolecules in prokaryotes. *FEMS microbiology reviews*, 30(6), 825–840.

Khachigian, L. M. (2019). Deoxyribozymes as catalytic nanotherapeutic agents. *Cancer research*, 79(5), 879–888.

Kher, G., Trehan, S. & Misra, A. (2011). Antisense oligonucleotides and rna interference. In *Challenges in Delivery of Therapeutic Genomics and Proteomics*, S. 325–386. Elsevier.

Kim, B., Tang, Q., Biswas, P. S., Xu, J., Schiffelers, R. M., Xie, F. Y., Ansari, A. M., Scaria, P. V., Woodle, M. C., Lu, P. *et al.* (2004). Inhibition of ocular angiogenesis by siRNA targeting vascular endothelial growth factor pathway genes: therapeutic strategy for herpetic stromal keratitis. *The American journal of pathology*, 165(6), 2177–2185.

Kim, H. N., Ren, W. X., Kim, J. S. & Yoon, J. (2012). Fluorescent and colorimetric sensors for detection of lead, cadmium, and mercury ions. *Chemical Society Reviews*, 41(8), 3210–3244.

Kim, J. H., Bothe, J. R., Frederick, R. O., Holder, J. C. & Markley, J. L. (2014). Role of IscX in iron–sulfur cluster biogenesis in Escherichia coli. *Journal of the American Chemical Society*, 136(22), 7933–7942.

Kim, J. H., Frederick, R. O., Reinen, N. M., Troupis, A. T. & Markley, J. L. (2013). [2Fe-2S]-Ferredoxin Binds Directly to Cysteine Desulfurase and Supplies an Electron for Iron–Sulfur Cluster Assembly but Is Displaced by the Scaffold Protein or Bacterial Frataxin. *Journal of the American Chemical Society*, 135(22), 8117–8120.

Kim, J. H., Tonelli, M., Frederick, R. O., Chow, D. C.-F. & Markley, J. L. (2012). Specialized Hsp70 chaperone (HscA) binds preferentially to the disordered form, whereas J-protein (HscB) binds preferentially to the structured form of the iron-sulfur cluster scaffold protein (IscU). *Journal of Biological Chemistry*, 287(37), 31406–31413.

Kim, K.-R., Kim, D.-R., Lee, T., Yhee, J. Y., Kim, B.-S., Kwon, I. C. & Ahn, Dae-R. (2013). Drug delivery by a self-assembled DNA tetrahedron for overcoming drug resistance in breast cancer cells. *Chemical Communications*, 49(20), 2010–2012.

Kim, S. H., Quigley, G., Suddath, F. L. & Rich, A. (1971). High-Resolution X-Ray Diffraction Patterns of Crystalline Transfer RNA That Show Helical Regions. *Proceedings of the National Academy of Sciences*, 68(4), 841–845.

Kispal, G., Csere, P., Prohl, C. & Lill, R. (1999). The mitochondrial proteins Atm1p and Nfs1p are essential for biogenesis of cytosolic Fe/S proteins. *The EMBO journal*, 18(14), 3981–3989.

Kisseleva, N., Khvorova, A., Westhof, E. & Schiemann, O. (2005). Binding of manganese (II) to a tertiary stabilized hammerhead ribozyme as studied by electron paramagnetic resonance spectroscopy. *RNA*, 11(1), 1–6.

Kitaoka, S., Wada, K., Hasegawa, Y., Minami, Y., Fukuyama, K. & Takahashi, Y. (2006). Crystal structure of Escherichia coli SufC, an ABC-type ATPase component of the SUF iron–sulfur cluster assembly machinery. *FEBS letters*, 580(1), 137–143.

Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P. *et al.* (2003). Essential Bacillus subtilis genes. *Proceedings of the National Academy of Sciences*, 100(8), 4678–4683.

Kolb, H. C., Finn, M. G. & Sharpless, K. B. (2001). Click chemistry: diverse chemical function from a few good reactions. *Angewandte Chemie International Edition*, 40(11), 2004–2021.

Kondo, J. & Takénaka, A. (2000). Crystallization of the most active RNA-cleaving deoxyribozyme. *Nucleic acids symposium series*, 44(1), 201–202.

Koshkin, A. A. & Wengel, J. (1998). Synthesis of Novel 2 ‘, 3 ‘-Linked Bicyclic Thymine Ribonucleosides. *The Journal of organic chemistry*, 63(8), 2778–2781.

Krebs, C., Agar, J. N., Smith, A. D., Frazzon, J., Dean, D. R., Huynh, B. H. & Johnson, M. K. (2001). IscA, an alternate scaffold for Fe- S cluster biosynthesis. *Biochemistry*, 40(46), 14069–14080.

Krug, N., Hohlfeld, J. M., Kirsten, A.-M., Kornmann, O., Beeh, K. M., Kappeler, D., Korn, S., Ignatenko, S., Timmer, W., Rogon, C. *et al.* (2015). Allergen-induced asthmatic responses modified by a GATA3-specific DNAzyme. *New England Journal of Medicine*, 372(21), 1987–1995.

Kruger, K., Grabowski, P. J., Zaug, A. J, Sands, J., Gottschling, D. E. & Cech, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *cell*, 31(1), 147–157.

Kuchenreuther, J. M., Grady-Smith, C. S., Bingham, A. S., George, S. J., Cramer, S. P. & Swartz, J. R. (2010). High-yield expression of heterologous [FeFe] hydrogenases in Escherichia coli. *PloS one*, 5(11).

Kumar, S., Jain, S., Dilbaghi, N., Ahluwalia, A. S., Hassan, A. A. & Kim, K.-H. (2019). Advanced selection methodologies for DNAzymes in sensing and healthcare applications. *Trends in biochemical sciences*, 44(3), 190–213.

Kurihara, T., Mihara, H., Kato, S.-i., Yoshimura, T. & Esaki, N. (2003). Assembly of iron–sulfur clusters mediated by cysteine desulfurases, IscS, CsdB and CSD, from Escherichia coli. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1647(1-2), 303–309.

Kurreck, J., Bieber, B., Jahnel, R. & Erdmann, V. A. (2002). Comparative study of DNA enzymes and ribozymes against the same full-length messenger RNA of the vanilloid receptor subtype I. *Journal of Biological Chemistry*, 277(9), 7099–7107.

Kurreck, J., Wyszko, E., Gillen, C. & Erdmann, V. A. (2002). Design of antisense oligonucleotides stabilized by locked nucleic acids. *Nucleic acids research*, 30(9), 1911–1918.

Ladbury, J. E. & Chowdhry, B. Z. (1996). Sensing the heat: the application of isothermal titration calorimetry to thermodynamic studies of biomolecular interactions. *Chemistry & biology*, 3(10), 791–801.

Lange, H., Kaut, A., Kispal, G. & Lill, R. (2000). A mitochondrial ferredoxin is essential for biogenesis of cellular iron-sulfur proteins. *Proceedings of the National Academy of Sciences*, 97(3), 1050–1055.

Law, M. J. (2005). Kinetic analysis of the role of the tyrosine 13, phenylalanine 56 and glutamine 54 network in the U1A/U1 hairpin II interaction. *Nucleic Acids Research*, 33(9), 2917–2928.

Layer, G., Gaddam, S. A., Ayala-Castro, C. N., Ollagnier-de Choudens, S., Lascoux, D., Fontecave, M. & Outten, F. W. (2007). SufE transfers sulfur from SufS to SufB for iron-sulfur cluster assembly. *Journal of biological chemistry*, 282(18), 13342–13350.

Layer, G., Ollagnier-de Choudens, S., Sanakis, Y. & Fontecave, M. (2006). Iron-Sulfur Cluster Biosynthesis Characterization of Escherichia Coli CYaY as an iron donor for the assembly of [2Fe-2S] clusters in the scaffold IscU. *Journal of Biological Chemistry*, 281(24), 16256–16263.

Leartsakulpanich, U., Antonkine, M. L. & Ferry, J. G. (2000). Site-Specific Mutational Analysis of a Novel Cysteine Motif Proposed To Ligate the 4Fe-4S Cluster in the Iron-Sulfur Flavoprotein of the Thermophilic MethanoarchaeonMethanosarcina thermophila. *Journal of bacteriology*, 182(19), 5309–5316.

Lee, R. C., Feinbaum, R. L. & Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *cell*, 75(5), 843–854.

Lermer, L., Roupioz, Y., Ting, R. & Perrin, D. M. (2002). Toward an RNaseA mimic: a DNAzyme with imidazoles and cationic amines. *Journal of the American Chemical Society*, 124(34), 9960–9961.

Leung, E. K. Y., Suslov, N., Tuttle, N., Sengupta, R. & Piccirilli, J. A. (2011). The mechanism of peptidyl transfer catalysis by the ribosome. *Annual review of biochemistry*, 80, 527–555.

Leuschner, P. J. F., Ameres, S. L., Kueng, S. & Martinez, J. (2006). Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO reports*, 7(3), 314–320.

Li, H., Gakh, O., Smith, D. Y. & Isaya, G. (2009). Oligomeric yeast frataxin drives assembly of core machinery for mitochondrial iron-sulfur cluster synthesis. *Journal of Biological Chemistry*, 284(33), 21971–21980.

Li, H., Wang, M., Shi, T., Yang, S., Zhang, J., Wang, H.-H. & Nie, Z. (2018). A DNA-Mediated Chemically Induced Dimerization (D-CID) Nanodevice for Nongenetic Receptor Engineering To Control Cell Behavior. *Angewandte Chemie International Edition*, 57(32), 10226–10230.

Li, J. & Lu, Y. (2000). A highly sensitive and selective catalytic DNA biosensor for lead ions. *Journal of the American Chemical Society*, 122(42), 10466–10467.

Li, J., Zheng, W., Kwon, A. H. & Lu, Y. (2000). In vitro selection and characterization of a highly efficient Zn (II)-dependent RNA-cleaving deoxyribozyme. *Nucleic acids research*, 28(2), 481–488.

Li, J., Zhu, D., Yi, Z., He, Y., Chun, Y., Liu, Y. & Li, N. (2005). DNAzymes targeting the icl gene inhibit ICL expression and decrease Mycobacterium tuberculosis survival in macrophages. *Oligonucleotides*, 15(3), 215–222.

Li, L., Feng, J., Fan, Y. & Tang, B. (2015). Simultaneous imaging of Zn2+ and Cu2+ in living cells based on DNAzyme modified gold nanoparticle. *Analytical chemistry*, 87(9), 4829–4835.

Li, P., Du, S., Li, Y. & He, J. (2017). Studies on the Two Thymine Residues in the Catalytic Core of 10-23 DNAzyme: The Impact on the Catalysis of Their 5-Substituted Functional Groups. *Molecules*, 22(7), 1011.

Li, S., Wang, L., Hao, Y., Zhang, L., Zhou, B., Deng, L. & Liu, Y.-N. (2014). An ultrasensitive colorimetric aptasensor for ATP based on peptide/Au nanocomposites and hemin–G-quadruplex DNAzyme. *RSC Advances*, 4(44), 23185–23190.

Li, X., Zhang, H., Tang, Y., Wu, P., Xu, S. & Zhang, X. (2017). A both-end blocked peroxidase-mimicking DNAzyme for low-background chemiluminescent sensing of miRNA. *ACS sensors*, 2(6), 810–816.

Li, Y. & Breaker, R. R. (1999). Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2 ‘-hydroxyl group. *Journal of the American Chemical society*, 121(23), 5364–5372.

Li, Y. & Breaker, R. R. (1999). Phosphorylating DNA with DNA. *Proceedings of the National Academy of Sciences*, 96(6), 2746–2751.

Li, Z., Liu, Y., Liu, G., Zhu, J., Zheng, Z., Zhou, Y. & He, J. (2014). Position-specific modification with imidazolyl group on10–23 DNAzyme realized catalytic activity enhancement. *Bioorganic & medicinal chemistry*, 22(15), 4010–4017.

Li, Z., Zhu, J. & He, J. (2016). Conformational studies of 10–23 DNAzyme in solution through pyrenyl-labeled 2’-deoxyadenosine derivatives. *Organic & biomolecular chemistry*, 14(41), 9846–9858.

Liang, L. & Astruc, D. (2011). The copper (I)-catalyzed alkyne-azide cycloaddition (CuAAC)“click” reaction and its applications. An overview. *Coordination Chemistry Reviews*, 255(23-24), 2933–2945.

Lill, R. (2009). Function and biogenesis of iron–sulphur proteins. *Nature*, 460(7257), 831.

Lilley, D. M. J. (2005). Structure, folding and mechanisms of ribozymes. *Current opinion in structural biology*, 15(3), 313–323.

Lippert, B. (2000). Multiplicity of metal ion binding patterns to nucleobases. *Coordination Chemistry Reviews*, 200, 487–516.

Littlefield, J. W., Keller, E. B., Gross, J. & Zamecnik, P. C. (1955). Studies on cytoplasmic ribonucleoprotein particles from the liver of the rat. *Journal of biological chemistry*, 217(1), 111–124.

Liu, G., Li, Z., Chiang, Y., Acton, T., Montelione, G. T., Murray, D. & Szyperski, T. (2005). High-quality homology models derived from NMR and X-ray structures of E. coli proteins YgdK and Suf E suggest that all members of the YgdK/Suf E protein family are enhancers of cysteine desulfurases. *Protein science*, 14(6), 1597–1608.

Liu, H., Yu, X., Chen, Y., Zhang, J., Wu, B., Zheng, L., Haruehanroengra, P., Wang, R., Li, S., Lin, J. *et al.* (2017). Crystal structure of an RNA-cleaving DNAzyme. *Nature communications*, 8(1), 1–10.

Liu, J., Brown, A. K., Meng, X., Cropek, D. M., Istok, J. D., Watson, D. B. & Lu, Y. (2007). A catalytic beacon sensor for uranium with parts-per-trillion sensitivity and millionfold selectivity. *Proceedings of the National Academy of Sciences*, 104(7), 2056–2061.

Liu, J. & Lu, Y. (2003). A colorimetric lead biosensor using DNAzyme-directed assembly of gold nanoparticles. *Journal of the American Chemical Society*, 125(22), 6642–6643.

Liu, J. & Lu, Y. (2003). Improving fluorescent DNAzyme biosensors by combining inter-and intramolecular quenchers. *Analytical chemistry*, 75(23), 6666–6672.

Liu, J. & Lu, Y. (2006). Fluorescent DNAzyme biosensors for metal ions based on catalytic molecular beacons. In *Fluorescent Energy Transfer Nucleic Acid Probes*, S. 275–288. Springer.

Lohans, C. T. & Vederas, J. C. (2014). Structural characterization of thioether-bridged bacteriocins. *The Journal of antibiotics*, 67(1), 23–30.

Loiseau, L., Ollagnier-de Choudens, S., Lascoux, D., Forest, E., Fontecave, M. & Barras, F. (2005). Analysis of the heteromeric CsdA-CsdE cysteine desulfurase, assisting Fe-S cluster biogenesis in Escherichia coli. *Journal of Biological Chemistry*, 280(29), 26760–26769.

Loiseau, L., Ollagnier-de Choudens, S., Nachin, L., Fontecave, M. & Barras, F. (2003). Biogenesis of Fe-S Cluster by the Bacterial SUF System: SufS and SufE Form a New Type of Cysteine Desulfurase. *Journal of Biological Chemistry*, 278(40), 38352–38359.

Lorenz, R., Bernhart, S. H., Zu Siederdissen, C. H., Tafer, H., Flamm, C., Stadler, P. F. & Hofacker, I. L. (2011). ViennaRNA Package 2.0. *Algorithms for molecular biology*, 6(1), 26.

Lu, J., Yang, J., Tan, G. & Ding, H. (2008). Complementary roles of SufA and IscA in the biogenesis of iron–sulfur clusters in Escherichia coli. *Biochemical Journal*, 409(2), 535–543.

Lu, Z.-X., Ye, M., Yan, G.-R., Li, Q., Tang, M., Lee, L. M., Sun, L.-Q. & Cao, Y. (2005). Effect of EBV LMP1 targeted DNAzymes on cell proliferation and apoptosis. *Cancer gene therapy*, 12(7), 647–654.

Lusic, H., Young, D. D., Lively, M. O. & Deiters, A. (2007). Photochemical DNA activation. *Organic letters*, 9(10), 1903–1906.

Lutz, R. E. (2018). *Elucidating the role of the iron-sulfur cluster in the nuclease/helicase Dna2*. PhD thesis, University of Zurich.

Ma, L. & Liu, J. (2020). Catalytic nucleic acids: biochemistry, chemical biology, biosensors and nanotechnology. *iScience*, S. 100815.

Maher, M., Cross, M., Wilce, M. C. J., Guss, J. M. & Wedd, A. G. (2004). Metal-substituted derivatives of the rubredoxin from Clostridium pasteurianum. *Acta Crystallographica Section D: Biological Crystallography*, 60(2), 298–303.

Majewska, J., Ciesielski, S. J., Schilke, B., Kominek, J., Blenska, A., Delewski, W., Song, J.-Y., Marszalek, J., Craig, E. A. & Dutkiewicz, R. (2013). Binding of the chaperone Jac1 protein and cysteine desulfurase Nfs1 to the iron-sulfur cluster scaffold Isu protein is mutually exclusive. *Journal of Biological Chemistry*, 288(40), 29134–29142.

Malkin, R. & Rabinowitz, J. C. (1966). The reconstitution of clostridial ferredoxin. *Biochemical and biophysical research communications*, 23(6), 822–827.

Maret, W. (2017). Zinc in cellular regulation: the nature and significance of “zinc signals”. *International journal of molecular sciences*, 18(11), 2285.

Marinoni, E. N., de Oliveira, J. S., Nicolet, Y., Raulfs, E. C., Amara, P., Dean, D. R. & Fontecilla-Camps, J. C. (2012). (IscS-IscU) 2 complex structures provide insights into Fe2S2 biogenesis and transfer. *Angewandte Chemie International Edition*, 51(22), 5439–5442.

Martick, M., Lee, T.-S., York, D. M. & Scott, W. G. (2008). Solvent structure and hammerhead ribozyme catalysis. *Chemistry & biology*, 15(4), 332–342.

Mastrianni, J. A., Nixon, R., Layzer, R., Telling, G. C., Han, D., DeArmond, S. J. & Prusiner, S. B. (1999). Prion protein conformation in a patient with sporadic fatal insomnia. *New England Journal of Medicine*, 340(21), 1630–1638.

Matsubara, H. & Saeki, K. (1992). Structural and functional diversity of ferredoxins and related proteins. In *Advances in inorganic chemistry*, Band 38, S. 223–280. Elsevier.

Mayer, M. P., Brehmer, D., Gässler, C. S. & Bukau, B. (2001). Hsp70 chaperone machines. *Advances in protein chemistry*, 59, 1–44.

McCoy, A. J. (2006). Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallographica Section D Structural Biology*, 63(1), 32–41.

McGhee, C. E., Loh, K. Y. & Lu, Y. (2017). DNAzyme sensors for detection of metal ions in the environment and imaging them in living cells. *Current opinion in biotechnology*, 45, 191–201.

McQuillen, K., Roberts, R. B. & Britten, R. J. (1959). Synthesis of nascent protein by ribosomes in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 45(9), 1437.

Meents, ., Reime, B., Stuebe, N., Fischer, P., Warmer, M., Goeries, D., Roever, J., Meyer, J., Fischer, J., Burkhardt, A., Vartiainen, I., Karvinen, P. & David, C. (2013). Development of an in-vacuum x-ray microscope with cryogenic sample cooling for beamline P11 at PETRA III. In *X-Ray Nanoimaging: Instruments and Methods* (Lai, Barry, ed.), Band 8851, S. 88510K.

Menniti, F.S., Miller, R.N., Putney, J.W. & Shears, S.B. (1993). Turnover of inositol polyphosphate pyrophosphates in pancreatoma cells. *Journal of Biological Chemistry*, 268(6), 3850–3856.

Mettert, E. L. & Kiley, P. J. (2014). Coordinate regulation of the Suf and Isc Fe-S cluster biogenesis pathways by IscR is essential for viability of Escherichia coli. *Journal of bacteriology*, 196(24), 4315–4323.

Meyer, A., Betzel, C. & Pusey, M. (2015). Latest methods of fluorescence-based protein crystal identification. *Acta Crystallographica Section F Structural Biology Communications*, 71(2), 121–131.

Meyer, J., Moulis, J.-M. & Lutz, M. (1986). High-yield chemical assembly of [2Fe-2X](X= S, Se) clusters into spinach apoferredoxin: Product characterization by resonance Raman spectroscopy. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, 871(3), 243–249.

Mihara, H. & Esaki, N. (2002). Bacterial cysteine desulfurases: their function and mechanisms. *Applied microbiology and biotechnology*, 60(1-2), 12–23.

Mills, D. R., Peterson, R. L. & Spiegelman, S. (1967). An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proceedings of the National Academy of Sciences of the United States of America*, 58(1), 217.

Milman, G., Langridge, R. & Chamberlin, M. J. (1967). The structure of a DNA-RNA hybrid. *Proceedings of the National Academy of Sciences of the United States of America*, 57(6), 1804.

Morales, R., Charon, M.-H., Hudry-Clergeon, G., Pétillot, Y., Norager, S., Medina, M. & Frey, M. (1999). Refined X-ray structures of the oxidized, at 1.3 Å, and reduced, at 1.17 Å,[2Fe- 2S] ferredoxin from the cyanobacterium Anabaena PCC7119 show redox-linked conformational changes. *Biochemistry*, 38(48), 15764–15773.

Morrison, D., Rothenbroker, M. & Li, Y. (2018). DNAzymes: selected for applications. *Small Methods*, 2(3), 1700319.

Mühlenhoff, U., Gerber, J., Richhardt, N. & Lill, R. (2003). Components involved in assembly and dislocation of iron–sulfur clusters on the scaffold protein Isu1p. *The EMBO journal*, 22(18), 4815–4825.

Mühlenhoff, U., Richter, N., Pines, O., Pierik, A. J. & Lill, R. (2011). Specialized function of yeast Isa1 and Isa2 proteins in the maturation of mitochondrial [4Fe-4S] proteins. *Journal of Biological Chemistry*, 286(48), 41205–41216.

Müller, C., Kuki, K. N., Pinheiro, D. T., de Souza, L. R., Silva, A. I. S., Loureiro, M. E., Oliva, M. A. & Almeida, A. M. (2015). Differential physiological responses in rice upon exposure to excess distinct iron forms. *Plant and Soil*, 391(1-2), 123–138.

Müller, J. J., Müller, A., Rottmann, M., Bernhardt, R. & Heinemann, U. (1999). Vertebrate-type and plant-type ferredoxins: crystal structure comparison and electron transfer pathway modelling. *Journal of molecular biology*, 294(2), 501–513.

Mullinger, R. N., Cammack, R., Rao, K. K., Hall, D.O., Dickson, D.P., Johnson, C.E., Rush, J.D. & Simopoulos, A. (1975). Physicochemical characterization of the four-iron-four-sulphide ferredoxin from Bacillus stearothermophilus. *Biochemical Journal*, 151(1), 75–83.

Mulugu, S., Bai, W., Fridy, P. C., Bastidas, R. J., Otto, J. C., Dollins, D. E., Haystead, T. A., Ribeiro, A. A. & York, J. D. (2007). A conserved family of enzymes that phosphorylate inositol hexakisphosphate. *Science*, 316(5821), 106–109.

Munday, R. (1989). Toxicity of thiols and disulphides: involvement of free-radical species. *Free Radical Biology and Medicine*, 7(6), 659–673.

Murray, J. B., Seyhan, A. A., Walter, N. G., Burke, J. M. & Scott, W. G. (1998). The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations alone. *Chemistry & biology*, 5(10), 587–595.

Murray, J. B., Szöke, H., Szöke, A. & Scott, W. G. (2000). Capture and visualization of a catalytic RNA enzyme-product complex using crystal lattice trapping and X-ray holographic reconstruction. *Molecular cell*, 5(2), 279–287.

Murshudov, G. N .and Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D Structural Biology*, 67(4), 355–67.

Nachin, L., Loiseau, L., Expert, D. & Barras, F. (2003). SufC: an unorthodox cytoplasmic ABC/ATPase required for [Fe—S] biogenesis under oxidative stress. *The EMBO journal*, 22(3), 427–437.

Nagai, K., Oubridge, C., Jessen, T. H., Li, J. & Evans, P. R. (1990). Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature*, 348, 515–520.

Nakamura, M., Saeki, K. & Takahashi, Y. (1999). Hyperproduction of recombinant ferredoxins in Escherichia coli by coexpression of the ORF1-ORF2-iscS-iscU-iscA-hscB-hscA-fdx-ORF3 gene cluster. *The Journal of Biochemistry*, 126(1), 10–18.

Nakano, S., Miyoshi, D. & Sugimoto, N. (2014). Effects of molecular crowding on the structures, interactions, and functions of nucleic acids. *Chemical reviews*, 114(5), 2733–2758.

Narayana Murthy, U. M., Ollagnier-de Choudens, S., Sanakis, Y., Abdel-Ghany, S. E., Rousset, C., Ye, H., Fontecave, M., Pilon-Smits, E. A. H. & Pilon, M. (2007). Characterization of Arabidopsis thaliana SufE2 and SufE3 functions in chloroplast iron-sulfur cluster assembly and NAD synthesis. *Journal of Biological Chemistry*, 282(25), 18254–18264.

Nawrot, B., Widera, K., Sobczak, M., Wojcik, M. & Stec, W. J. (2008). Effect of RP and SP Phosphorothioate Substitution at the Scissile Site on the Cleavage Activity of Deoxyribozyme 10-23. *Current Organic Chemistry*, 12(12), 1004–1009.

Nawrot, B., Widera, K., Wojcik, M., Rebowska, B., Nowak, G. & Stec, W. J. (2007). Mapping of the functional phosphate groups in the catalytic core of deoxyribozyme 10–23. *The FEBS journal*, 274(4), 1062–1072.

Newville, M., Otten, R., Nelson, A., Ingargiola, A., Stensitzki, T., Allan, D., Fox, A., Carter, F., Michaŭ0142, Pustakhod, D., Ram, Y., Glenn, Deil, C., Stuermer, Beelen, A., Frost, O., Zobrist, N., Pasquevich, G., Hansen, A. L. R., Spillane, T., Caldwell, S., Polloreno, A., Hannum, A., Zimmermann, J., Borreguero, J., Fraine, J., deep 42-thought, Maier, B. F., Gamari, B. & Almarza, A. (2019). lmfit/lmfit-py 1.0.0.

Nikzad, N. & Karami, Z. (2018). Label-free colorimetric sensor for sensitive detection of choline based on DNAzyme-choline oxidase coupling. *International journal of biological macromolecules*, 115, 1241–1248.

Nowakowski, J., Shim, P. J., Joyce, G. F. & Stout, C. D. (1999). Crystallization of the 10-23 DNA enzyme using a combinatorial screen of paired oligonucleotides. *Acta Crystallographica Section D: Biological Crystallography*, 55(11), 1885–1892.

Nowakowski, J., Shim, P. J., Prasad, G. S., Stout, C. D. & Joyce, G. F. (1999). Crystal structure of an 82-nucleotide RNA–DNA complex formed by the 10-23 DNA enzyme. *Nature structural biology*, 6(2), 151–156.

Nowakowski, J., Shim, P. J., Stout, C. D. & Joyce, G. F. (2000). Alternative conformations of a nucleic acid four-way junction. *Journal of Molecular Biology*, 300(1), 93–102.

Okumoto, Y. & Sugimoto, N. (2000). Effects of metal ions and catalytic loop sequences on the complex formation of a deoxyribozyme and its RNA substrate. *Journal of inorganic biochemistry*, 82(1-4), 189–195.

Ollagnier-de Choudens, S., Lascoux, D., Loiseau, L., Barras, F., Forest, E. & Fontecave, M. (2003). Mechanistic studies of the SufS–SufE cysteine desulfurase: evidence for sulfur transfer from SufS to SufE. *FEBS letters*, 555(2), 263–267.

Ollagnier-de Choudens, S., Loiseau, L., Sanakis, Y., Barras, F. & Fontecave, M. (2005). Quinolinate synthetase, an iron–sulfur enzyme in NAD biosynthesis. *FEBS letters*, 579(17), 3737–3743.

Ollagnier-de Choudens, S., Mattioli, T., Takahashi, Y. & Fontecave, M. (2001). Iron-Sulfur Cluster Assembly Characterization Of IscA and Evidence For a Specific and Functional Complex With Ferredoxin. *Journal of Biological Chemistry*, 276(25), 22604–22607.

Ollagnier-de Choudens, S., Nachin, L., Sanakis, Y., Loiseau, L., Barras, F. & Fontecave, M. (2003). SufA from Erwinia chrysanthemi CHARACTERIZATION OF A SCAFFOLD PROTEIN REQUIRED FOR IRON-SULFUR CLUSTER ASSEMBLY. *Journal of Biological Chemistry*, 278(20), 17993–18001.

Orio, M. & Mouesca, J.-M. (2008). Variation of Average g Values and Effective Exchange Coupling Constants among [2Fe- 2S] Clusters: A Density Functional Theory Study of the Impact of Localization (Trapping Forces) versus Delocalization (Double-Exchange) as Competing Factors. *Inorganic chemistry*, 47(12), 5394–5416.

Ortigão, J. F. R., Rösch, H., Selter, H., Fröhlich, A., Lorenz, A., Montenarh, M. & Seliger, H. (1992). Antisense effect of oligodeoxynucleotides with inverted terminal internucleotidic linkages: a minimal modification protecting against nucleolytic degradation. *Antisense research and development*, 2(2), 129–146.

Oubridge, C., Ito, N., Evans, P. R., Teo, C.-H. & Nagai, K. (1994). Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature*, 372, 432–438.

Oubridge, C., Ito, N., Teo, C.-H., Fearnley, I. & Nagai, K. (1995). Crystallisation of RNA-protein complexes II. The application of protein engineering for crystallisation of the U1A protein-RNA complex. *Journal of Molecular Biology*, 249(2), 409–423.

Outten, F. W., Wood, M. J., Muñoz, F. M. & Storz, G. (2003). The SufE protein and the SufBCD complex enhance SufS cysteine desulfurase activity as part of a sulfur transfer pathway for Fe-S cluster assembly in Escherichia coli. *Journal of Biological Chemistry*, 278(46), 45713–45719.

Pandelia, M.-E., Lanz, N. D., Booker, S. J. & Krebs, C. (2015). Mössbauer spectroscopy of Fe/S proteins. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1853(6), 1395–1405.

Parent, A., Elduque, X., Cornu, D., Belot, L., Le Caer, J.-P., Grandas, A., Toledano, M. B. & D'autréaux, B. (2015). Mammalian frataxin directly enhances sulfur transfer of NFS1 persulfide to both ISCU and free thiols. *Nature communications*, 6(1), 1–12.

Pascual-Ortiz, M., Saiardi, A., Walla, E., Jakopec, V., Künzel, N. A., Span, I., Vangala, A. & Fleig, U. (2018). Asp1 bifunctional activity modulates spindle function via controlling cellular inositol pyrophosphate levels in Schizosaccharomyces pombe. *Molecular and cellular biology*, 38(9), e00047–18.

Pastore, C., Adinolfi, S., Huynen, M. A., Rybin, V., Martin, S., Mayer, M., Bukau, B. & Pastore, A. (2006). YfhJ, a molecular adaptor in iron-sulfur cluster formation or a frataxin-like protein? *Structure*, 14(5), 857–867.

Paz-Ares, L., Douillard, J.-Y., Koralewski, P., Manegold, C., Smit, E. F., Reyes, J. M., Chang, G.-C., John, W. J., Peterson, P. M., Obasaju, C. K. *et al.* (2006). Phase III study of gemcitabine and cisplatin with or without aprinocarsen, a protein kinase C-alpha antisense oligonucleotide, in patients with advanced-stage non–small-cell lung cancer. *Journal of clinical oncology*, 24(9), 1428–1434.

Pearson, R. G. (1968). Hard and soft acids and bases, HSAB, part 1: Fundamental principles. *Journal of Chemical Education*, 45(9), 581.

Pechlaner, M. & Sigel, R. K. O. (2012). Characterization of metal ion-nucleic acid interactions in solution. In *Interplay between Metal Ions and Nucleic Acids*, S. 1–42. Springer.

Pelzer, W., Mühlenhoff, U., Diekert, K., Siegmund, K., Kispal, G. & Lill, R. (2000). Mitochondrial Isa2p plays a crucial role in the maturation of cellular iron–sulfur proteins. *FEBS letters*, 476(3), 134–139.

Pérard, J. & de Choudens, S. O. (2018). Iron–sulfur clusters biogenesis by the SUF machinery: close to the molecular mechanism understanding. *JBIC Journal of Biological Inorganic Chemistry*, 23(4), 581–596.

Perrotta, A. T. & Been, M. D. (2006). HDV ribozyme activity in monovalent cations. *Biochemistry*, 45(38), 11357–11365.

Plashkevych, O. & Chattopadhyaya, J. (2011). Structure of 10–23 DNAzyme in Complex with the Target RNA In Silico—A Progress Report on the Mechanism of RNA Cleavage by DNA Enzyme. *Medicinal Chemistry of Nucleic Acids*, S. 272–291.

Pley, H. W., Flaherty, K. M. & McKay, D. B. (1994). Model for an RNA tertiary interaction from the structure of an intermolecular complex between a GAAA tetraloop and an RNA helix. *Nature*, 372(6501), 111–113.

Pley, H. W., Flaherty, K. M. & McKay, D. B. (1994). Three-dimensional structure of a hammerhead ribozyme. *Nature*, 372(6501), 68–74.

Pöhlmann, J. & Fleig, U. (2010). Asp1, a conserved 1/3 inositol polyphosphate kinase, regulates the dimorphic switch in Schizosaccharomyces pombe. *Molecular and cellular biology*, 30(18), 4535–4547.

Pöhlmann, J., Risse, C., Seidel, C., Pohlmann, T., Jakopec, V., Walla, E., Ramrath, P., Takeshita, N., Baumann, S., Feldbrügge, M. *et al.* (2014). The Vip1 inositol polyphosphate kinase family regulates polarized growth and modulates the microtubule cytoskeleton in fungi. *PLoS genetics*, 10(9).

Poland, D. (1978). *Cooperative equlibria in physical biochemistry.* Clarendon Press.

Ponce-Salvatierra, A., Wawrzyniak-Turek, K., Steuerwald, U., Höbartner, C. & Pena, V. (2016). Crystal structure of a DNA catalyst. *Nature*, 529(7585), 231–234.

Pradeepkumar, P. I., Höbartner, C., Baum, D. A. & Silverman, S. K. (2008). DNA-Catalyzed Formation of Nucleopeptide Linkages. *Angewandte Chemie International Edition*, 47(9), 1753–1757.

Price, S. R., Ito, N., Oubridge, C., Avis, J. M. & Nagai, K. (1995). Crystallization of RNA-protein complexes I. Methods for the large-scale preparation of RNA suitable for crystallographic studies. *Journal of Molecular Biology*, 249(2), 398–408.

Prischi, F., Konarev, P. V., Iannuzzi, C., Pastore, C., Adinolfi, S., Martin, S. R., Svergun, D. I. & Pastore, A. (2010). Structural bases for the interaction of frataxin with the central components of iron–sulphur cluster assembly. *Nature communications*, 1(1), 1–10.

Privalov, P. L. (1979). Stability of proteins small globular proteins. In *Advances in protein chemistry*, Band 33, S. 167–241. Elsevier.

Prusiner, S. B., Bolton, D. C., Groth, D. F., Bowman, K. A., Cochran, S. P. & McKinley, M. P. (1982). Further purification and characterization of scrapie prions. *Biochemistry*, 21(26), 6942–6950.

Py, B. & Barras, F. (2015). Genetic approaches of the Fe–S cluster biogenesis process in bacteria: historical account, methodological aspects and future challenges. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1853(6), 1429–1435.

Pyle, A. M., Chu, V. T., Jankowsky, E. & Boudvillain, M. (2000). [10] using dnazymes to cut, process, and map rna molecules for structural studies or modification. In *RNA - Ligand Interactions, Part A*, Band 317 der *Methods in Enzymology*, S. 140 – 146. Academic Press.

Raghunathan, K., Harris, P. T. & Arvidson, D. N. (2010). Trial by fire: are the crystals macromolecules? *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 66(5), 615–620.

Rajagopalan, S., Teter, S. J., Zwart, P. H., Brennan, R. G., Phillips, K. J. & Kiley, P. J. (2013). Studies of IscR reveal a unique mechanism for metal-dependent regulation of DNA binding specificity. *Nature structural & molecular biology*, 20(6), 740.

Raulfs, E. C., O'Carroll, I. P., Dos Santos, P. C., Unciuleac, M.-C. & Dean, D. R. (2008). In vivo iron–sulfur cluster formation. *Proceedings of the National Academy of Sciences*, 105(25), 8591–8596.

Ravel, B. & Newville, M. (2005). ATHENA, ARTEMIS, HEPHAESTUS: data analysis for X-ray absorption spectroscopy using IFEFFIT. *Journal of synchrotron radiation*, 12(4), 537–541.

Räz, M. H. & Hollenstein, M. (2015). Probing the effect of minor groove interactions on the catalytic efficiency of DNAzymes 8–17 and 10–23. *Molecular BioSystems*, 11(5), 1454–1461.

Reardan, D. T., Meares, C. F., Goodwin, D. A., McTigue, M., David, G. S., Stone, M. R., Leung, J. P., Bartholomew, R. M. & Frincke, J. M. (1985). Antibodies against metal chelates. *Nature*, 316(6025), 265–268.

Record Jr., M. T., Lohman, T. M. & De Haseth, P. (1976). Ion effects on ligand-nucleic acid interactions. *Journal of molecular biology*, 107(2), 145–158.

Reich, S. J., Fosnot, J., Kuroki, A., Tang, W., Yang, X., Maguire, A. M., Bennett, J., Tolentino, M. J. *et al.* (2003). Small interfering RNA (siRNA) targeting VEGF effectively inhibits ocular neovascularization in a mouse model. *Mol Vis*, 9(5), 210–216.

Rekittke, I., Wiesner, J., Röhrich, R., Demmer, U., Warkentin, E., Xu, W., Troschke, K., Hintz, M., No, J. H., Duin, E. C. *et al.* (2008). Structure of (E)-4-hydroxy-3-methyl-but-2-enyl diphosphate reductase, the terminal enzyme of the non-mevalonate pathway. *Journal of the American Chemical Society*, 130(51), 17206–17207.

Riboldi, G. P., Verli, H. & Frazzon, J. (2009). Structural studies of the Enterococcus faecalis SufU [Fe-S] cluster protein. *BMC biochemistry*, 10(1), 3.

Roach, P. L. (2011). Radicals from S-adenosylmethionine and their application to biosynthesis. *Current opinion in chemical biology*, 15(2), 267–275.

Robaldo, L., Izzo, F., Dellafiore, M., Proietti, C., Elizalde, P. V., Montserrat, J. M. & Iribarren, A. M. (2012). Influence of conformationally restricted pyrimidines on the activity of 10–23 DNAzymes. *Bioorganic & medicinal chemistry*, 20(8), 2581–2586.

Robaldo, L., Montserrat, J. M. & Iribarren, A. M. (2010). 10-23 DNAzyme modified with (2'R)- and (2'S)-2'-deoxy-2'-C-methyluridine in the catalytic core. *Bioorganic & medicinal chemistry letters*, 20(15), 4367–4370.

Robbins, A. H. & Stout, C. D. (1989). The structure of aconitase. *Proteins: Structure, Function, and Bioinformatics*, 5(4), 289–312.

Roberts, R. B., ed. (1958). *Microsomal particles and protein synthesis. First Symposium of the Biophysical Society.* Pergamon, London.

Robertson, D. L. & Joyce, G. F. (1990). Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature*, 344(6265), 467–468.

Roche, B., Huguenot, A., Barras, F. & Py, B. (2015). The iron-binding CyaY and IscX proteins assist the ISC-catalyzed F e-S biogenesis in E scherichia coli. *Molecular microbiology*, 95(4), 605–623.

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

Rousset, C., Fontecave, M. & de Ollagnier-de Choudens, S. (2008). The [4Fe–4S] cluster of quinolinate synthase from Escherichia coli: Investigation of cluster ligands. *FEBS letters*, 582(19), 2937–2944.

Rupert, P. B. (2002). Transition State Stabilization by a Catalytic RNA. *Science*, 298, 1421–1424.

Rupert, P. B. & Ferré-D'Amaré, A. R. (2001). Crystal structure of a hairpin ribozyme–inhibitor complex with implications for catalysis. *Nature*, 410, 780–786.

Ruvkun, G. (2001). Glimpses of a tiny RNA world.(Perspectives: Molecular Biology). *Science*, 294(5543), 797–800.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 239(4839), 487–491.

Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. & Arnheim, N. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, 230(4732), 1350–1354.

Saini, A., Mapolelo, D. T., Chahal, H. K., Johnson, M. K. & Outten, F. W. (2010). SufD and SufC ATPase activity are required for iron acquisition during in vivo Fe-S cluster formation on SufB. *Biochemistry*, 49(43), 9402–9412.

Santiago, F. S., Atkins, D. G. & Khachigian, L. M. (1999). Vascular Smooth Muscle Cell Proliferation and Regrowth after Mechanical Injury in Vitro Are. Egr-1/NGFI-A-Dependent. *The American journal of pathology*, 155(3), 897–905.

Santiago, F. S., Lowe, H. C., Kavurma, M. M, Chesterman, C. N., Baker, A., Atkins, D. G. & Khachigian, L. M. (1999). New DNA enzyme targeting Egr-1 mRNA inhibits vascular smooth muscle proliferation and regrowth after injury. *Nature medicine*, 5(11), 1264–1269.

Santoro, S. W. & Joyce, G. F. (1997). A general purpose RNA-cleaving DNA enzyme. *Proceedings of the national academy of sciences*, 94(9), 4262–4266.

Santoro, S. W. & Joyce, G. F. (1998). Mechanism and utility of an RNA-cleaving DNA enzyme. *Biochemistry*, 37(38), 13330–13342.

Saran, R. & Liu, J. (2016). A silver DNAzyme. *Analytical chemistry*, 88(7), 4014–4020.

Sawata, S., Shimayama, T., Komiyama, M., Kumar, P. KR, Nishikawa, S. & Taira, K. (1993). Enhancement of the cleavage rates of DNA-armed hammerhead ribozymes by various divalent metal ions. *Nucleic acids research*, 21(24), 5656–5660.

Schiemann, O., Fritscher, J., Kisseleva, N., Sigurdsson, S. T. & Prisner, T. F. (2003). Structural Investigation of a High-Affinity MnII Binding Site in the Hammerhead Ribozyme by EPR Spectroscopy and DFT Calculations. Effects of Neomycin B on Metal-Ion Binding. *ChemBioChem*, 4(10), 1057–1065.

Schnabl, J. & Sigel, R. K. O. (2010). Controlling ribozyme activity by metal ions. *Current opinion in chemical biology*, 14(2), 269–275.

Schneider, T. D. & Stephens, R. M. (1990). Sequence logos: a new way to display consensus sequences. *Nucleic Acids Research*, 18(20), 6097–6100.

Schubert, S., Fürste, J. P., Werk, D.and Grunert, H.-P., Zeichhardt, H., Erdmann, V. A & Kurreck, J. (2004). Gaining target access for deoxyribozymes. *Journal of molecular biology*, 339(2), 355–363.

Schubert, S., Guèl, D. C., Grunert, H.-P., Zeichhardt, H., Erdmann, V. A. & Kurreck, J. (2003). RNA cleaving '10-23'DNAzymes with enhanced stability and activity. *Nucleic acids research*, 31(20), 5982–5992.

Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical journal*, 78(3), 1606–1619.

Schwartz, C. J., Djaman, O., Imlay, J. A. & Kiley, P. J. (2000). The cysteine desulfurase, IscS, has a major role in in vivo Fe-S cluster formation in Escherichia coli. *Proceedings of the National Academy of Sciences*, 97(16), 9009–9014.

Schwartz, C. J., Giel, J. L., Patschkowski, T., Luther, C., Ruzicka, F. J., Beinert, H. & Kiley, P. J. (2001). IscR, an Fe-S cluster-containing transcription factor, represses expression of Escherichia coli genes encoding Fe-S cluster assembly proteins. *Proceedings of the National Academy of Sciences*, 98(26), 14895–14900.

Scott, W. G., Finch, J. T., Grenfell, R., Fogg, J., Smith, T., Gait, M. J. & Klug, A. (1995). Rapid crystallization of chemically synthesized hammerhead RNAs using a double screening procedure. *Journal of molecular biology*, 250(3), 327–332.

Scott, W. G., Murray, J. B., Arnold, J. R. P., Stoddard, B. L. & Klug, A. (1996). Capturing the structure of a catalytic RNA intermediate: the hammerhead ribozyme. *Science*, 274(5295), 2065–2069.

Sel, S., Wegmann, M., Dicke, T., Sel, S., Henke, W., Yildirim, A. Ö., Renz, H. & Garn, H. (2008). Effective prevention and therapy of experimental allergic asthma using a GATA-3–specific DNAzyme. *Journal of Allergy and Clinical Immunology*, 121(4), 910–916.

Selbach, B., Earles, E. & Dos Santos, P. C. (2010). Kinetic analysis of the bisubstrate cysteine desulfurase SufS from Bacillus subtilis. *Biochemistry*, 49(40), 8794–8802.

Selbach, B. P., Chung, A. H., Scott, A. D., George, S. J., Cramer, S. P. & Dos Santos, P. C. (2014). Fe-S cluster biogenesis in Gram-positive bacteria: SufU is a zinc-dependent sulfur transfer protein. *Biochemistry*, 53(1), 152–160.

Shakamuri, P., Zhang, B. & Johnson, M. K. (2012). Monothiol glutaredoxins function in storing and transporting [Fe2S2] clusters assembled on IscU scaffold proteins. *Journal of the American Chemical Society*, 134(37), 15213–15216.

Shan, S., Yoshida, A., Sun, S., Piccirilli, J. A. & Herschlag, D. (1999). Three metal ions at the active site of the Tetrahymena group I ribozyme. *Proceedings of the National Academy of Sciences*, 96(22), 12299–12304.

Sharma, A. K., Pallesen, L. J., Spang, R. J. & Walden, W. E. (2010). Cytosolic iron-sulfur cluster assembly (CIA) system: factors, mechanism, and relevance to cellular iron regulation. *Journal of Biological Chemistry*, 285(35), 26745–26751.

Shechner, D. M., Grant, R. A., Bagby, S. C., Koldobskaya, Y., Piccirilli, J. A. & Bartel, D. P. (2009). Crystal Structure of the Catalytic Core of an RNA-Polymerase Ribozyme. *Science*, 326, 1271–1275.

Sheftel, A. D., Wilbrecht, C., Stehling, O., Niggemeyer, B., Elsässer, H.-P., Mühlenhoff, U. & Lill, R. (2012). The human mitochondrial ISCA1, ISCA2, and IBA57 proteins are required for [4Fe-4S] protein maturation. *Molecular biology of the cell*, 23(7), 1157–1166.

Shen, L., Chen, Z., Li, Y., He, S., Xie, S., Xu, X., Liang, Z., Meng, X., Li, Q., Zhu, Z. *et al.* (2008). Electrochemical DNAzyme sensor for lead based on amplification of DNA- Au Bio-Bar codes. *Analytical chemistry*, 80(16), 6323–6328.

Shi, R., Proteau, A., Villarroya, M., Moukadiri, I., Zhang, L., Trempe, J.-F., Matte, A., Armengod, M. E. & Cygler, M. (2010). Structural basis for Fe–S cluster assembly and tRNA thiolation mediated by IscS protein–protein interactions. *PLoS biology*, 8(4).

Shiels, J. C., Tuite, J. B., Nolan, S. J. & Baranger, A. M. (2002). Investigation of a conserved stacking interaction in target site recognition by the U1A protein. *Nucleic Acids Research*, 30(2), 550–558.

Shimomura, Y., Takahashi, Y., Kakuta, Y. & Fukuyama, K. (2005). Crystal structure of Escherichia coli YfhJ protein, a member of the ISC machinery involved in assembly of iron–sulfur clusters. *Proteins: Structure, Function, and Bioinformatics*, 60(3), 566–569.

Shimomura, Y., Wada, K., Fukuyama, K. & Takahashi, Y. (2008). The asymmetric trimeric architecture of [2Fe–2S] IscU: implications for its scaffolding during iron–sulfur cluster biosynthesis. *Journal of Molecular Biology*, 383(1), 133–143.

Sidorov, A. V., Grasby, J. A. & Williams, D. M. (2004). Sequence-specific cleavage of RNA in the absence of divalent metal ions by a DNAzyme incorporating imidazolyl and amino functionalities. *Nucleic Acids Research*, 32(4), 1591–1601.

Sigel, R. K. O. & Pyle, A. M. (2007). Alternative roles for metal ions in enzyme catalysis and the implications for ribozyme chemistry. *Chemical reviews*, 107(1), 97–113.

Silverman, S. K. (2005). In vitro selection, characterization, and application of deoxyribozymes that cleave RNA. *Nucleic acids research*, 33(19), 6151–6163.

Silverman, S. K. (2009). Deoxyribozymes: selection design and serendipity in the development of DNA catalysts. *Accounts of chemical research*, 42(10), 1521–1531.

Silverman, S. K. (2015). Pursuing DNA catalysts for protein modification. *Accounts of chemical research*, 48(5), 1369–1379.

Silverman, S. K. (2016). Catalytic DNA: scope, applications, and biochemistry of deoxyribozymes. *Trends in biochemical sciences*, 41(7), 595–609.

Smuga, D., Majchrzak, K., Sochacka, E. & Nawrot, B. (2010). RNA-cleaving 10–23 deoxyribozyme with a single amino acid-like functionality operates without metal ion cofactors. *New Journal of Chemistry*, 34(5), 934–948.

Soukup, G. A. & Breaker, R. R. (1999). Relationship between internucleotide linkage geometry and the stability of RNA. *Rna*, 5(10), 1308–1325.

Span, I. (2012). *Strukturelle und funktionelle Charakterisierung des Eisen-Schwefel-Proteins IspH in Komplex mit Liganden*. PhD thesis, Technische Universität München.

Span, I., Wang, K., Wang, W., Zhang, Y., Bacher, A., Eisenreich, W., Li, K., Schulz, C., Oldfield, E. & Groll, M. (2012). Discovery of acetylene hydratase activity of the iron–sulphur protein IspH. *Nature communications*, 3(1), 1–8.

Spiro, T. G. & Czernuszewicz, R. S. (1995). Resonance Raman spectroscopy of metalloproteins. In *Methods in enzymology*, Band 246, S. 416–460. Elsevier.

Steger, G. (1994). Thermal denaturation of double-stranded nucleic acids: prediction of temperatures critical for gradient gel electrophoresis and polymerase chain reaction. *Nucleic acids research*, 22(14), 2760–2768.

Stehling, O., Vashisht, A. A., Mascarenhas, J., Jonsson, Z. O., Sharma, T., Netz, D. J. A., Pierik, A. J., Wohlschlegel, J. A. & Lill, R. (2012). MMS19 assembles iron-sulfur proteins required for DNA metabolism and genomic integrity. *Science*, 337(6091), 195–199.

Stepanov, S., Hilgart, M., Yoder, D. W., Makarov, O., Becker, M., Sanishvili, R., Ogata, C. M., Venugopalan, N., Aragao, D., Caffrey, M. *et al.* (2011). Fast fluorescence techniques for crystallography beamlines. *Journal of applied crystallography*, 44(4), 772–778.

Stephens, L., Radenberg, T., Thiel, U., Vogel, G., Khoo, K.-H., Dell, A., Jackson, T.R., Hawkins, P.T. & Mayr, G.W. (1993). The detection, purification, structural characterization, and metabolism of diphosphoinositol pentakisphosphate (s) and bisdiphosphoinositol tetrakisphosphate (s). *Journal of Biological Chemistry*, 268(6), 4009–4015.

Stephens, P.J., Thomson, A.J., Dunn, J.B.R., Keiderling, T.A., Rawlings, J., Rao, K.K. & Hall, D.O. (1978). Circular dichroism and magnetic circular dichroism of iron-sulfur proteins. *Biochemistry*, 17(22), 4770–4778.

Stirling, P. C., Bloom, M. S., Solanki-Patil, T., Smith, S., Sipahimalani, P., Li, Z., Kofoed, M., Ben-Aroya, S., Myung, K. & Hieter, P. (2011). The complete spectrum of yeast chromosome instability genes identifies candidate CIN cancer genes and functional roles for ASTRA complex components. *PLoS genetics*, 7(4).

Strautmann, J. B.H., George, S. D.B., Bothe, E., Bill, E., Weyhermuüller, T., Stammler, A., Bögge, H. & Glaser, T. (2008). Molecular and electronic structures of mononuclear iron complexes using strongly electron donating ligands and their oxidized forms. *Inorganic chemistry*, 47(15), 6804–6824.

Sugimoto, N., Nakano, S., Katoh, M., Matsumura, A., Nakamuta, H., Ohmichi, T., Yoneyama, M. & Sasaki, M. (1995). Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry*, 34(35), 11211–11216.

Sugimoto, N., Okumoto, Y. & Ohmichi, T. (1999). Effect of metal ions and sequence of deoxyribozymes on their RNA cleavage activity. *Journal of the Chemical Society, Perkin Transactions 2*, S. 1381–1386.

Sumner, J. B. & Dounce, A. L. (1937). CRYSTALLINE CATALASE. *Science*, 85, 366–367.

Sun, F., Ji, Q., Jones, M. B., Deng, X., Liang, H., Frank, B., Telser, J., Peterson, S. N., Bae, T. & He, C. (2012). AirSR, a [2Fe-2S] cluster-containing two-component system, mediates global oxygen sensing and redox signaling in Staphylococcus aureus. *Journal of the American Chemical Society*, 134(1), 305–314.

Tafer, H., Ameres, S. L., Obernosterer, G., Gebeshuber, C. A., Schroeder, R., Martinez, J. & Hofacker, I. L. (2008). The impact of target site accessibility on the design of effective siRNAs. *Nature biotechnology*, 26(5), 578–583.

Takahashi, Y. & Tokumoto, U. (2002). A third bacterial system for the assembly of iron-sulfur clusters with homologs in archaea and plastids. *Journal of Biological Chemistry*, 277(32), 28380–28383.

Tang, Y. & Guest, J. R. (1999). Direct evidence for mRNA binding and post-transcriptional regulation by Escherichia coli aconitases. *Microbiology*, 145(11), 3069–3079.

Taylor, G. L (2010). Introduction to phasing. *Acta Crystallographica Section D: Biological Crystallography*, 66(4), 325–338.

Thai, H. B. D., Levi-Acobas, F., Yum, S.-Y., Jang, G., Hollenstein, M. & Ahn, D.-R. (2018). Tetrahedral DNAzymes for enhanced intracellular gene-silencing activity. *Chemical communications*, 54(68), 9410–9413.

Tian, Y., Wang, Y., Xu, Y., Liu, Y., Li, D. & Fan, C. (2015). A highly sensitive chemiluminescence sensor for detecting mercury (II) ions: a combination of Exonuclease III-aided signal amplification and graphene oxide-assisted background reduction. *Science China Chemistry*, 58(3), 514–518.

Ting, R., Lermer, L. & Perrin, D. M. (2004). Triggering DNAzymes with light: a photoactive C8 thioether-linked adenosine. *Journal of the American Chemical Society*, 126(40), 12720–12721.

Tinoco, I., Jr. & Bustamante, C. (1999). How RNA folds. *Journal of molecular biology*, 293(2), 271–281.

Tirupati, B., Vey, J. L., Drennan, C. L. & Bollinger, J. M. (2004). Kinetic and structural characterization of Slr0077/SufS, the essential cysteine desulfurase from Synechocystis sp. PCC 6803. *Biochemistry*, 43(38), 12210–12219.

Tokumoto, U., Nomura, S., Minami, Y., Mihara, H., Kato, S., Kurihara, T., Esaki, N., Kanazawa, H., Matsubara, H. & Takahashi, Y. (2002). Network of protein-protein interactions among iron-sulfur cluster assembly proteins in Escherichia coli. *The Journal of Biochemistry*, 131(5), 713–719.

Tokumoto, U. & Takahashi, Y. (2001). Genetic analysis of the isc operon in Escherichia coli involved in the biogenesis of cellular iron-sulfur proteins. *The Journal of biochemistry*, 130(1), 63–71.

Tolentino, M. J., Brucker, A. J., Fosnot, J., Ying, G.-S., Wu, I.-H., Malik, G., Wan, S. & Reich, S. J. (2004). Intravitreal injection of vascular endothelial growth factor small interfering RNA inhibits growth and leakage in a nonhuman primate, laser-induced model of choroidal neovascularization. *Retina*, 24(1), 132–138.

Tolle, F., Brändle, G. M., Matzner, D. & Mayer, G. (2015). A versatile approach towards nucleobase-modified aptamers. *Angewandte Chemie International Edition*, 54(37), 10971–10974.

Topolski, B., Jakopec, V., Künzel, N. A. & Fleig, U. (2016). Inositol pyrophosphate kinase Asp1 modulates chromosome segregation fidelity and spindle function in Schizosaccharomyces pombe. *Molecular and cellular biology*, 36(24), 3128–3140.

Torabi, S.-F., Wu, P., McGhee, C. E., Chen, L., Hwang, K., Zheng, N., Cheng, J. & Lu, Y. (2015). In vitro selection of a sodium-specific DNAzyme and its application in intracellular sensing. *Proceedings of the National Academy of Sciences*, 112(19), 5903–5908.

Travascio, P., Li, Y. & Sen, D. (1998). DNA-enhanced peroxidase activity of a DNA aptamer-hemin complex. *Chemistry & biology*, 5(9), 505–517.

Trepanier, J. B., Tanner, J. E. & Alfieri, C. (2008). Reduction in intracellular HCV RNA and virus protein expression in human hepatoma cells following treatment with 2'-O-methyl-modified anti-core deoxyribozyme. *Virology*, 377(2), 339–344.

Tuerk, C. & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, 249(4968), 505–510.

Turner, D.H., Sugimoto, N. & Freier, S.M. (1990). Thermodynamics and kinetics of base-pairing and of DNA and RNA self-assembly and helix coil transition. In *Nucleic Acids, Subvolume c, Physical Data I, Spectroscopic and Kinetic Data.* (Saenger, W., ed.), Landolt-Börnstein, Group VII Biophysics, Vol I, S. 201–212.

Unciuleac, M.-C., Chandramouli, K., Naik, S., Mayer, S., Huynh, B. H., Johnson, M. K. & Dean, D. R. (2007). In Vitro Activation of Apo-Aconitase Using a [4Fe-4S] Cluster-Loaded Form of the IscU [Fe- S] Cluster Scaffolding Protein. *Biochemistry*, 46(23), 6812–6821.

Unwalla, H. & Banerjea, A. C. (2001). Novel mono-and di-DNA-enzymes targeted to cleave TAT or TAT-REV RNA inhibit HIV-1 gene expression. *Antiviral research*, 51(2), 127–139.

Urbina, H. D., Silberg, J. J., Hoff, K. G. & Vickery, L. E. (2001). Transfer of sulfur from IscS to IscU during Fe/S cluster assembly. *Journal of Biological Chemistry*, 276(48), 44521–44526.

Uzarska, M. A., Dutkiewicz, R., Freibert, S.-A., Lill, R. & Mühlenhoff, U. (2013). The mitochondrial Hsp70 chaperone Ssq1 facilitates Fe/S cluster transfer from Isu1 to Grx5 by complex formation. *Molecular biology of the cell*, 24(12), 1830–1841.

Valegård, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J. & Liljas, L. (1994). Crystal structure of an RNA bacteriophage coat protein–operator complex. *Nature*, 371, 623–626.

van Gelder, C. W., Gunderson, S. I., Jansen, E. J., Boelens, W. C., Polycarpou-Schwarz, M., Mattaj, I. W. & van Venrooij, W. J. (1993). A complex secondary structure in U1A pre-mRNA that binds two molecules of U1A protein is required for regulation of polyadenylation. *The EMBO Journal*, 12(13), 5191–200.

Van Vranken, J. G., Jeong, M.-Y., Wei, P., Chen, Y.-C., Gygi, S. P., Winge, D. R. & Rutter, J. (2016). The mitochondrial acyl carrier protein (ACP) coordinates mitochondrial fatty acid synthesis with iron sulfur cluster biogenesis. *eLife*, 5, e17828.

Varani, L., Gunderson, S. I., Mattaj, L. W., Kay, L. E., Neuhaus, D. & Varani, G. (2000). The NMR structure of the 38 kDa U1A protein - PIE RNA complex reveals the basis of cooperativity in regulation of polyadenylation by human U1A protein. *Nature Structural Biology*, 7(4), 329–335.

Vester, B., Lundberg, L. B., Sørensen, M. D., Babu, B. R., Douthwaite, S. & Wengel, J. (2002). LNAzymes: incorporation of LNA-type monomers into DNAzymes markedly increases RNA cleavage. *Journal of the American Chemical Society*, 124(46), 13682–13683.

Vester, B. & Wengel, J. (2004). LNA (locked nucleic acid): high-affinity targeting of complementary RNA and DNA. *Biochemistry*, 43(42), 13233–13241.

Vickery, L. E., Silberg, J. J. & Ta, D. T. (1997). Hsc66 and Hsc20, a new heat shock cognate molecular chaperone system from Escherichia coli. *Protein Science*, 6(5), 1047–1056.

Victor, Julian, Steger, Gerhard & Riesner, Detlev (2018). Inability of DNAzymes to cleave RNA in vivo is due to limited Mg^{2+} concentration in cells. *European Biophysics Journal: EBJ*, 47(4), 333–343.

Viladoms, J. & Parkinson, G. N. (2014). HELIX: a new modular nucleic acid crystallization screen. *Journal of Applied Crystallography*, 47(3), 948–955.

Vinella, D., Brochier-Armanet, C., Loiseau, L., Talla, E. & Barras, F. (2009). Iron-sulfur (Fe/S) protein biogenesis: phylogenomic and genetic studies of A-type carriers. *PLoS genetics*, 5(5).

Virtanen, Pauli, Gommers, Ralf, Oliphant, Travis E., Haberland, Matt, Reddy, Tyler, Cournapeau, David, Burovski, Evgeni, Peterson, Pearu, Weckesser, Warren, Bright, Jonathan, van der Walt, Stéfan J., Brett, Matthew, Wilson, Joshua, Jarrod Millman, K., Mayorov, Nikolay, Nelson, Andrew R. J., Jones, Eric, Kern, Robert, Larson, Eric, Carey, CJ, Polat, İlhan, Feng, Yu, Moore, Eric W., VanderPlas, Jake, Laxalde, Denis, Perktold, Josef, Cimrman, Robert, Henriksen, Ian, Quintero, E.A., Harris, Charles R., Archibald, Anne M., Ribeiro, Antônio H., Pedregosa, Fabian, van Mulbregt, Paul & Contributors, SciPy 1.0 (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17, 261–272.

Volkov, V. V. & Svergun, D. I. (2003). Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography*, 36(3), 860–864.

Wada, K., Hasegawa, Y., Gong, Z., Minami, Y., Fukuyama, K. & Takahashi, Y. (2005). Crystal structure of Escherichia coli SufA involved in biosynthesis of iron–sulfur clusters: implications for a functional dimer. *FEBS letters*, 579(29), 6543–6548.

Wahlestedt, C., Salmi, P., Good, L., Kela, J., Johnsson, T., Hökfelt, T., Broberger, C., Porreca, F., Lai, J., Ren, K. *et al.* (2000). Potent and nontoxic antisense oligonucleotides containing locked nucleic acids. *Proceedings of the National Academy of Sciences*, 97(10), 5633–5638.

Walsh, S. M., Konecki, S. N. & Silverman, S. K. (2015). Identification of sequence-selective tyrosine kinase deoxyribozymes. *Journal of molecular evolution*, 81(5-6), 218–224.

Walsh, S. M., Sachdeva, A. & Silverman, S. K. (2013). DNA catalysts with tyrosine kinase activity. *Journal of the American Chemical Society*, 135(40), 14928–14931.

Walter, T. S., Meier, C., Assenberg, R., Au, K.-F., Ren, J., Verma, A., Nettleship, J. E., Owens, R. J., Stuart, D. I. & Grimes, J. M. (2006). Lysine methylation as a routine rescue strategy for protein crystallization. *Structure*, 14(11), 1617–1622.

Wang, B., Cao, L., Chiuman, W., Li, Y. & Xi, Z. (2010). Probing the function of nucleotides in the catalytic cores of the 8-17 and 10-23 DNAzymes by abasic nucleotide and C3 spacer substitutions. *Biochemistry*, 49(35), 7553–7562.

Wang, F., Saran, R. & Liu, J. (2015). Tandem DNAzymes for mRNA cleavage: Choice of enzyme, metal ions and the antisense effect. *Bioorganic & medicinal chemistry letters*, 25(7), 1460–1463.

Wang, Huanchen, Nair, Vasudha S, Holland, Ashley A, Capolicchio, Samanta, Jessen, Henning J, Johnson, Michael K & Shears, Stephen B (2015). Asp1 from Schizosaccharomyces pombe binds a [2Fe-2S] 2+ cluster which inhibits inositol pyrophosphate 1-phosphatase activity. *Biochemistry*, 54(42), 6462–6474.

Wang, Q., Zhang, D., Liu, Y., Cheng, M., He, J. & Liu, K. (2012). A Structure-Activity Relationship Study for 2'-Deoxyadenosine Analogs at A9 Position in the Catalytic Core of 10-23 DNAzyme for Rate Enhancement. *Nucleic acid therapeutics*, 22(6), 423–427.

Wang, W., Wang, K., Liu, Y.-L., No, J.-H., Li, J., Nilges, M. J. & Oldfield, E. (2010). Bioorganometallic mechanism of action, and inhibition, of IspH. *Proceedings of the National Academy of Sciences*, 107(10), 4522–4527.

Wang, X., Feng, M., Xiao, L., Tong, A. & Xiang, Y. (2016). Postsynthetic modification of DNA phosphodiester backbone for photocaged dnazyme. *ACS chemical biology*, 11(2), 444–451.

Wang, Y. & Irudayaraj, J. (2011). A SERS DNAzyme biosensor for lead ion detection. *Chemical Communications*, 47(15), 4394–4396.

Ward, J., Ollmann, E., Maxey, E. & Finney, L. A. (2014). X-ray absorption spectroscopy of metalloproteins. In *Metalloproteins*, S. 171–187. Springer.

Watenpaugh, K. D., Sieker, L. C., Herriott, J. R. & Jensen, L. H. (1972). The structure of a non-heme iron protein: rubredoxin at 1.5 Å resolution. In *Cold Spring Harbor symposia on quantitative biology*, Band 36, S. 359–367.

Webert, H., Freibert, S.-A., Gallo, A.o, Heidenreich, T., Linne, U., Amlacher, S., Hurt, E., Mühlenhoff, U., Banci, L. & Lill, R. (2014). Functional reconstitution of mitochondrial Fe/S cluster synthesis on Isu1 reveals the involvement of ferredoxin. *Nature communications*, 5(1), 1–12.

Wegner, S. V., Okesli, A., Chen, P. & He, C. (2007). Design of an emission ratiometric biosensor from MerR family proteins: A sensitive and selective sensor for Hg2+. *Journal of the American Chemical Society*, 129(12), 3474–3475.

Wells, G. A., Scott, A. C., Johnson, C. T., Gunning, R. F., Hancock, R. D., Jeffrey, M., Dawson, M. & Bradley, R. (1987). A novel progressive spongiform encephalopathy in cattle. *Veterinary Record*, 121(18), 419–420.

Weng, D. E, Masci, P. A., Radka, S. F., Jackson, T. E., Weiss, P. A., Ganapathi, R., Elson, P. J., Capra, W. B., Parker, V. P., Lockridge, J. A. *et al.* (2005). A phase I clinical trial of a ribozyme-based angiogenesis inhibitor targeting vascular endothelial growth factor receptor-1 for patients with refractory solid tumors. *Molecular cancer therapeutics*, 4(6), 948–955.

Wengel, J. (1999). Synthesis of 3 '-C-and 4 '-C-branched oligodeoxynucleotides and the development of locked nucleic acid (LNA). *Accounts of Chemical Research*, 32(4), 301–310.

White, M. F. & Dillingham, M. S. (2012). Iron–sulphur clusters in nucleic acid processing enzymes. *Current opinion in structural biology*, 22(1), 94–100.

Wieckowski, B. M., Hegemann, J. D., Mielcarek, A., Boss, L., Burghaus, O. & Marahiel, M. A (2015). The PqqD homologous domain of the radical SAM enzyme ThnB is required for thioether bond formation during thurincin H maturation. *FEBS letters*, 589(15), 1802–1806.

Wilkosz, P. A., Chandrasekhar, K. & Rosenberg, J. M. (1995). Preliminary characterization of EcoRI–DNA co-crystals: incomplete factorial design of oligonucleotide sequences. *Acta Crystallographica Section D: Biological Crystallography*, 51(6), 938–945.

Williams, C. H., Stillman, T. J., Barynin, V. V., Sedelnikova, S. E., Tang, Y., Green, J., Guest, J. R. & Artymiuk, P. J. (2002). E. coli aconitase B structure reveals a HEAT-like domain with implications for protein–protein recognition. *Nature Structural Biology*, 9(6), 447–452.

Williams, E. S. & Young, S. (1980). Chronic wasting disease of captive mule deer: a spongiform encephalopathy. *Journal of wildlife diseases*, 16(1), 89–98.

Wilson, D. S. & Szostak, J. W. (1999). In vitro selection of functional nucleic acids. *Annual review of biochemistry*, 68(1), 611–647.

Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A. & Breaker, R. R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. *Nature*, 428(6980), 281–286.

Winter, G., Lobley, C. M.C. & Prince, S. M. (2013). Decision making in xia2. *Acta Crystallographica Section D Structural Biology*, 69(7), 1260–1273.

Wojtas, M. N. & Abrescia, N. G. A. (2012). Soaking of DNA into crystals of archaeal RNA polymerase achieved by desalting in droplets. *Acta Crystallographica Section F Structural Biology Communications*, 68(9), 1134–1138.

Wojtas, M. N., Mogni, M., Millet, O., Bell, S. D. & Abrescia, N. G. A. (2012). Structural and functional analyses of the interaction of archaeal RNA polymerase with DNA. *Nucleic Acids Research*, 40(19), 9941–9952.

Wolff, M., Seemann, M., Tse Sum Bui, B., Frapart, Y., Tritsch, D., Estrabot, A. G., Rodrıguez-Concepción, M., Boronat, A., Marquet, A. & Rohmer, M. (2003). Isoprenoid biosynthesis via the methylerythritol phosphate pathway: the (E)-4-hydroxy-3-methylbut-2-enyl diphosphate reductase (LytB/IspH) from Escherichia coli is a [4Fe–4S] protein. *FEBS letters*, 541(1-3), 115–120.

Wollers, S., Layer, G., Garcia-Serres, R., Signor, L., Clemancey, M., Latour, J.-M., Fontecave, M. & Ollagnier de Choudens, S. (2010). Iron-Sulfur (Fe-S) Cluster Assembly the SufBCD complex is a new type of Fe-S scaffold with a flavin redox cofactor. *Journal of Biological Chemistry*, 285(30), 23331–23341.

Wong, O. Y., Pradeepkumar, P.I. & Silverman, S. K. (2011). DNA-catalyzed covalent modification of amino acid side chains in tethered and free peptide substrates. *Biochemistry*, 50(21), 4741–4749.

Wong-Staal, F., Poeschla, E. M. & Looney, D. J. (1998). A Controlled, Phase 1 Clinical Trial to Evaluate the Safety and Effects in Hiv-1 Infected Humans of Autologous Lymphocytes Transduced with a Ribozyme That Cleaves Hiv-1 Rna. University of California San Diego, La Jolla, California. *Human gene therapy*, 9(16), 2407–2425.

Wu, G., Mansy, S. S., Wu, S., Surerus, K. K., Foster, M. W. & Cowan, J. A. (2002). Characterization of an iron- sulfur cluster assembly protein (ISU1) from Schizosaccharomyces pombe. *Biochemistry*, 41(15), 5024–5032.

Wu, S.-P., Mansy, S. S. & Cowan, J. A. (2005). Iron- Sulfur Cluster Biosynthesis. Molecular Chaperone DnaK Promotes IscU-Bound [2Fe-2S] Cluster Stability and Inhibits Cluster Transfer Activity. *Biochemistry*, 44(11), 4284–4293.

Wu, Y., Yu, L., McMahon, R., Rossi, J. J., Forman, S. J. & Snyder, D. S. (1999). Inhibition of bcr-abl oncogene expression by novel deoxyribozymes (DNAzymes). *Human gene therapy*, 10(17), 2847–2857.

Xiang, G., Seki, T., Schuster, M. D., Witkowski, P., Boyle, A. J., See, F., Martens, T. P., Kocher, A., Sondermeijer, H., Krum, H. *et al.* (2005). Catalytic degradation of vitamin D up-regulated protein 1 mRNA enhances cardiomyocyte survival and prevents left ventricular remodeling after myocardial ischemia. *Journal of Biological Chemistry*, 280(47), 39394–39402.

Xiao, Y., Rowe, A. A. & Plaxco, K. W. (2007). Electrochemical detection of parts-per-billion lead via an electrode-bound DNAzyme assembly. *Journal of the American Chemical Society*, 129(2), 262–263.

Yacyshyn, B. R., Chey, W. Y., Goff, J., Salzberg, B., Baerg, R., Buchman, A. L., Tami, J., Yu, R., Gibiansky, E. & Shanahan, W. R. (2002). Double blind, placebo controlled trial of the remission inducing and steroid sparing properties of an ICAM-1 antisense oligodeoxynucleotide, alicaforsen (ISIS 2302), in active steroid dependent Crohn's disease. *Gut*, 51(1), 30–36.

Yan, R., Adinolfi, S. & Pastore, A. (2015). Ferredoxin, in conjunction with NADPH and ferredoxin-NADP reductase, transfers electrons to the IscS/IscU complex to promote iron–sulfur cluster assembly. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1854(9), 1113–1117.

Yen, L., Strittmatter, S. M. & Kalb, R. G. (1999). Sequence-specific cleavage of Huntingtin mRNA by catalytic DNA. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 46(3), 366–373.

Yokoyama, N., Nonaka, C., Ohashi, Y., Shioda, M., Terahata, T., Chen, W., Sakamoto, K., Maruyama, C., Saito, T., Yuda, E. *et al.* (2018). Distinct roles for U-type proteins in iron–sulfur cluster biosynthesis revealed by genetic analysis of the Bacillus subtilis sufCDSUB operon. *Molecular microbiology*, 107(6), 688–703.

Yoon, T. & Cowan, J. A. (2003). Iron- sulfur cluster biosynthesis. Characterization of frataxin as an iron donor for assembly of [2Fe-2S] clusters in ISU-type proteins. *Journal of the American Chemical Society*, 125(20), 6078–6084.

Young, D. D., Lively, M. O. & Deiters, A. (2010). Activation and deactivation of DNAzyme and antisense function with light for the photochemical regulation of gene expression in mammalian cells. *Journal of the American Chemical Society*, 132(17), 6183–6193.

Yousaf, R., Gu, C., Ahmed, Z. M., Khan, S. N., Friedman, T. B., Riazuddin, S., Shears, S. B. & Riazuddin, S. (2018). Mutations in Diphosphoinositol-Pentakisphosphate Kinase PPIP5K2 are associated with hearing loss in human and mouse. *PLoS genetics*, 14(3), e1007297.

Yuda, E., Tanaka, N., Fujishiro, T., Yokoyama, N., Hirabayashi, K., Fukuyama, K., Wada, K. & Takahashi, Y. (2017). Mapping the key residues of SufB and SufD essential for biosynthesis of iron-sulfur clusters. *Scientific reports*, 7(1), 1–12.

Zaborowska, Z., Fürste, J. P., Erdmann, V. A. & Kurreck, J. (2002). Sequence requirements in the catalytic core of the "10-23" DNA enzyme. *Journal of Biological Chemistry*, 277(43), 40617–40622.

Zaborowska, Z., Schubert, S., Kurreck, J. & Erdmann, V. A. (2005). Deletion analysis in the catalytic region of the 10–23 DNA enzyme. *FEBS letters*, 579(2), 554–558.

Zhang, G., Dass, C. R., Sumithran, E., Di Girolamo, N., Sun, L.-Q. & Khachigian, L. M. (2004). Effect of deoxyribozymes targeting c-Jun on solid tumor growth and angiogenesis in rodents. *Journal of the National Cancer Institute*, 96(9), 683–696.

Zhang, L., Gasper, W. J., Stass, S. A., Ioffe, O. B., Davis, M. A. & Mixson, A. J. (2002). Angiogenic inhibition mediated by a DNAzyme that targets vascular endothelial growth factor receptor 2. *Cancer research*, 62(19), 5463–5469.

Zhang, L., Han, B., Li, T. & Wang, E. (2011). Label-free DNAzyme-based fluorescing molecular switch for sensitive and selective detection of lead ions. *Chemical Communications*, 47(11), 3099–3101.

Zhang, L., Zhang, Y., Wei, M., Yi, Y., Li, H. & Yao, S. (2013). A label-free fluorescent molecular switch for Cu 2+ based on metal ion-triggered DNA-cleaving DNAzyme and DNA intercalator. *New Journal of Chemistry*, 37(4), 1252–1257.

Zhang, W., Urban, A., Mihara, H., Leimkühler, S., Kurihara, T. & Esaki, N. (2010). IscS functions as a primary sulfur-donating enzyme by interacting specifically with MoeB and MoaD in the biosynthesis of molybdopterin in Escherichia coli. *Journal of Biological Chemistry*, 285(4), 2302–2308.

Zhang, X., Xu, Y., Ling, H. & Hattori, T. (1999). Inhibition of infection of incoming HIV-1 virus by RNA-cleaving DNA enzyme. *FEBS letters*, 458(2), 151–156.

Zhang, X.-B., Kong, R.-M. & Lu, Y. (2011). Metal ion sensors based on DNAzymes and related DNA molecules. *Annual review of analytical chemistry*, 4, 105–128.

Zhang, X.-B., Wang, Z., Xing, H., Xiang, Y. & Lu, Y. (2010). Catalytic and molecular beacons for amplified detection of metal ions and organic molecules with high sensitivity. *Analytical chemistry*, 82(12), 5005–5011.

Zheng, L., Cash, V. L., Flint, D. H. & Dean, D. R. (1998). Assembly of iron-sulfur clusters identification of an iscSUA-hscBA-fdx gene cluster from Azotobacter vinelandii. *Journal of Biological Chemistry*, 273(21), 13264–13272.

Zheng, L. & Dean, D. R. (1994). Catalytic formation of a nitrogenase iron-sulfur cluster. *Journal of Biological Chemistry*, 269(29), 18723–18726.

Zheng, L., White, R. H., Cash, V. L. & Dean, D. R. (1994). Mechanism for the desulfurization of L-cysteine catalyzed by the nifS gene product. *Biochemistry*, 33(15), 4714–4720.

Zheng, L., White, R. H., Cash, V. L., Jack, R. F. & Dean, D. R. (1993). Cysteine desulfurase activity indicates a role for NIFS in metallocluster biosynthesis. *Proceedings of the National Academy of Sciences*, 90(7), 2754–2758.

Zhou, D.-M., He, Q.-C., Zhou, J.-M. & Taira, K. (1998). Explanation by a putative triester-like mechanism for the thio effects and Mn2+ rescues in reactions catalyzed by a hammerhead ribozyme. *FEBS letters*, 431(2), 154–160.

Zhou, W., Chen, Q., Huang, P.-J. J., Ding, J. & Liu, J. (2015). DNAzyme hybridization, cleavage, degradation, and sensing in undiluted human blood serum. *Analytical chemistry*, 87(7), 4001–4007.

Zhou, W., Ding, J. & Liu, J. (2017). Theranostic dnazymes. *Theranostics*, 7(4), 1010.

Zhou, W., Saran, R., Huang, P.-J. J., Ding, J. & Liu, J. (2017). An exceptionally selective DNA cooperatively binding two Ca2+ ions. *ChemBioChem*, 18(6), 518–522.

Zhou, W., Saran, R. & Liu, J. (2017). Metal sensing by DNA. *Chemical reviews*, 117(12), 8272–8325.

Zhou, W., Vazin, M., Yu, T., Ding, J. & Liu, J. (2016). In vitro selection of chromium-dependent DNAzymes for sensing chromium (III) and chromium (VI). *Chemistry–A European Journal*, 22(28), 9835–9840.

Zhu, J., Li, Z., Wang, Q., Liu, Y. & He, J. (2016). The contribution of adenines in the catalytic core of 10-23 DNAzyme improved by the 6-amino group modifications. *Bioorganic & medicinal chemistry letters*, 26(18), 4462–4465.

Zhu, J., Li, Z., Yang, Z. & He, J. (2015). Studies on the preferred uracil–adenine base pair at the cleavage site of 10–23 DNAzyme by functional group modifications on adenine. *Bioorganic & medicinal chemistry*, 23(15), 4256–4263.

Declaration of Contributions

Molecular features and metal ions that influence 10-23 DNAzyme activity

Hannah Rosenbach[a], Julian Victor[a], Gerhard Steger[a], Detlev Riesner[a], Manuel Etzkorn[a,b], and Ingrid Span[a,*]

Submitted to: **Molecules**, 31.05.2020
Journal Impact Factor (2018): 3.060
Contribution: 75%

H.R. selected the references and wrote the first draft of the manuscript. H.R. prepared Tables 2.2 and 2.3, and Figures 2.3, 2.4, 2.6, 2.7, and 2.8. All authors commented on the manuscript.

Influence of monovalent metal ions on metal binding and catalytic activity of the 10-23 DNAzyme

Hannah Rosenbach[a], Jan Borggräfe[a,b], Julian Victor[a], Christine Wübben[c], Olav Schiemann[c], Gerhard Steger[a], Manuel Etzkorn[a,b,*], and Ingrid Span[a,*]

Submitted to: **Biological Chemistry**, 03.06.2020
Journal Impact Factor (2017): 3.014
Contribution: 65%

All authors designed the experiments. H.R. performed the kinetic and activity measurements shown in the article. H.R. performed EPR measurements together with C.W. and contributed to analysis of both kinetic and EPR data. H.R. wrote the first draft of the manuscript. All authors commented on the manuscript.

Expanding crystallization tools for nucleic acid complexes using U1A protein variants

Hannah Rosenbach[a], Julian Victor[a], Jan Borggräfe[a,b], Ralf Biehl[d], Gerhard Steger[a], Manuel Etzkorn[a,b], and Ingrid Span[a,*]

Published in: **Journal of Structural Biology**, 2020, Vol. 210(2):107480
DOI: https://doi.org/10.1016/j.jsb.2020.107480
Rosenbach *et al.* (2020)
Journal Impact Factor(2017): 3.489
Contribution: 75%

H.R., J.V., M.E., and I.S. designed the experiments and wrote the manuscript. H.R. isolated all protein variants, performed crystallization and soaking experiments, as well as activity measurements. H.R. collected X-ray diffraction data and solved the crystal structures published in this article. H.R. prepared samples for NMR and SAXS measurements. H.R. wrote the first draft of the manuscript. All authors commented on the manuscript.

Approaches to crystallize the 10-23 DNAzyme in a biologically relevant conformation

Hannah Rosenbach[a], Julian Victor[a], Thomas Pauly[b], Luitgard Nagel-Steger[b], Manuel Etzkorn[a,b], Gerhard Steger[a], and Ingrid Span[a,*]

To be submitted in: **Crystals**
Journal Impact Factor (2017): 2.144
Contribution: 90%

H.R., J.V., and I.S. designed the experiments. H.R. isolated the U1A protein used in this study. H.R. performed all the experiments shown in the article, except the AUC measurements. H.R. prepared the samples analyzed by AUC. H.R. wrote the manuscript.

Comparing the maturation of [4Fe–4S] cluster containing proteins *in vivo* and *in vitro*

Hannah Rosenbach[a,†], Steffen Mielenbrink[a,†], Melissa Jansing[a], Sabine Metzger[e,f], and Ingrid Span[a,*]

To be submitted to: **ChemBioChem**
Journal Impact Factor (2018): 2.641
Contribution: 45%

H.R., S.M., and I.S. designed the experiments. H.R. supervized and assisted M.J. in performing protein isolation and electronic absorption spectroscopy. H.R. analyzed and plotted the data and prepared the figures shown in the article. H.R. wrote the first draft of the manuscript. H.R., S.M. and I.S. finalized the manuscript.

The Asp1 pyrophosphatase from *S. pombe* hosts a $[2Fe–2S]^{2+}$ cluster *in vivo*

Hannah Rosenbach[a], Eva K. Walla[g], George E. Cutsail III[h], James Birrell[h], Marina Pascual-Ortiz[i], Serena DeBeer, Ursula Fleig[g], and Ingrid Span[a,*]

To be submitted to: **Journal of Biological Inorganic Chemistry**
Journal Impact Factor (2018): 3.632
Contribution: 70%

H.R., U.F. and I.S. designed the experiments. H.R. isolated the protein for spectroscopic analysis and activity measurements. H.R. performed electronic absorption spectroscopy and *in vitro* activity measurements shown in this article. H.R. prepared samples for EPR, X-ray absorption and Mössbauer spectroscopy. H.R., U.F., G.E.C., J.B., and I.S. wrote the manuscript. All authors commented on the manuscript.

[a] Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany

[b] Institute for Biological Information Processing: Structural Biochemistry (IBI-7), Research Center Jülich, Wilhelm-Johnen-Str., 52428 Jülich, Germany

[c] Institute of Physical and Theoretical Chemistry, University of Bonn, Wegelerstr. 12, 53115 Bonn, Germany

[d] Jülich Centre for Neutron Science (JCNS-1/ICS-1), Research Center Jülich, Wilhelm-Johnen-Str., 52428 Jülich, Germany

[e] MS-Platform Biocenter, Cluster of Excellence on Plant Science (CEPLAS), University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany.

[f] IUF-Leibniz Research Institute for Environmental Medicine, Heisenberg-Group - Environmentally-Induced Cardiovascular Degeneration, Mass Spectrometry Core Unit, Auf'm Hennekamp 50, 40225 Düsseldorf, Germany.

[g] Institute for Functional Genome Research of Microorganisms, Heinrich-Heine-Universität Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany

[h] Max Planck Institute for Chemical Energy Conversion, Stiftstr. 34-36, 45470 Mülheim an der Ruhr, Germany.

[i] Department of Biomedical Sciences, Faculty of Health Sciences, Universidad Cardenal Herrera, CEU Universities, 46113, Valencia, Spain

[†] equally contributed

* corresponding authors
Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf,
Universitätsstr. 1, 40225 Düsseldorf, Germany
Ingrid Span: ingrid.span@hhu.de; +(49) (0)211 - 8115582
Manuel Etzkorn: manuel.etzkorn@hhu.de; +(49) (0)211 - 8112023

Erklärung

Ich versichere an Eides statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, 21.09.2020
Ort, Datum

Hannah Rosenbach

Danksagung

Ganz herzlich danke ich meiner Erstgutachterin Frau Jun.-Prof. Dr. Ingrid Span für die freundliche Aufnahme in ihre Arbeitsgruppe und die Bereitstellung der spannenden und herausfordernden Projekte, an denen ich im Rahmen meiner Promotion arbeiten durfte. Ich habe diese Zeit als sehr lehrreich und stets als eine Mischung aus Fordern und Fördern empfunden. Herrn Prof. Dr. Peter Gilch danke ich sehr herzlich für die Übernahme des Zweitgutachtens und die Möglichkeit, meinen Forschungsstand in regelmäßigen Abständen in seiner Arbeitsgruppe vorzustellen und zu diskutieren. Herrn Prof. Dr. Dieter Willbold habe ich für die Bereitstellung der Laboratorien und Verbrauchsmaterialien zu danken. Ganz besonders herzlich bedanke ich mich bei Herrn apl. Prof. Dr. Ing. Gerhard Steger für seine Unterstützung, Hilfe und Tipps in sämtlichen Bereichen des wissenschaftlichen und nicht-wissenschaftlichen Lebens (Nukleinsäuren, IT, Wandern, Kochen, ...), ganz besonders aber natürlich für die tolle Zusammenarbeit im DNAzym-Projekt und die vielen schnellen Hilfen mit LaTeX. Mein besonderer Dank richtet sich auch an Herrn Prof. Dr. Dr. h.c. Detlev Riesner und seine Frau Prof. Dr. Hannelore Riesner. Sie haben mich in den letzten Jahren in vielerlei Hinsicht unterstützt und sind mir stets hilfsbereite und herzliche Mentoren gewesen, sei es in der Wissenschaft als auch darüber hinaus. Ich danke Herrn Dr. Manuel Etzkorn und seinen Arbeitsgruppenmitgliedern Dr. Julian Victor, Dr. Aldino Viegas und Jan Borggräfe für die tolle und interessante Zusammenarbeit am DNAzym-Projekt. Bedanken möchte ich mich auch bei den „guten Seelen“ des IPB, Astrid Wies, Barbara Schulten und Elke Reinartz. Sie haben nicht nur stets ein offenes Ohr für sämtliche Belange des Lebens, sondern sorgen durch ihre „Arbeit im Hintergrund“ jeden Tag dafür, dass das Institut am Laufen gehalten wird, wir immer ausreichend Materialien im Labor und in der Küche haben und unsere Rechnungen bezahlt und Anträge richtig ausgefüllt werden. Allen aktuellen und ehemaligen Mitgliedern der AG Span danke ich für die Zusammenarbeit in den letzten Jahren. Mein Dank richtet sich auch an die vielen Kooperationspartner außerhalb des IPB, mit denen ich in den letzten Jahren an spannenden Projekten zusammenarbeiten durfte und ohne die meine Arbeit in diesem Umfang nicht möglich gewesen wäre. Besonders erwähnen möchte ich hier Prof. Dr. George E. Cutsail III, Dr. James Birrell, apl. Prof. Dr. Ursula Fleig, Eva K. Walla, Prof. Dr. Olav Schiemann, Christine Wübben und Assistant Prof. Dr. Aaron T. Smith.

Ein ganz besonderer Dank gilt meinen Eltern Renate und Achim, die mich stets kompromiss- und bedingungslos auf meinem Weg unterstützt haben und auf die ich mich zu jeder Zeit voll und ganz verlassen kann. Genauso sehr danke ich meinem Freund Steffen, der seit nun drei Jahren der ruhende Gegenpol zu mir ist und mich durch seine Ausgeglichenheit und Fürsorge immer und überall unterstützt. Ohne ihn hätte ich die letzten Monate nur schwer überstanden. Meinen Freundinnen Lisa und Lena danke ich für die nun schon jahrelange Freundschaft und die vielen unbeschwerten Stunden. Danken möchte ich auch Julian und Michael, die ich im IPB nicht nur als tolle Kollegen, sondern auch als Freunde gewonnen habe und die die Zeit im Institut zu etwas Besonderem gemacht haben. Dankend erwähnen möchte ich auch die (Ex-)IPB'ler Anne, Marcel, Marie, Emil, Jan und Alessia für die schönen Stunden.