DIPLOMARBEIT

Detecting metal ion binding pockets within the group II intron ai5γ *in vivo*

angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer.nat.)

Verfasserin / Verfasser:	Michael Wildauer (0400880)
Studienrichtung /Studienzweig (lt. Studienblatt):	Molekulare Biologie (A 490)
Betreuerin / Betreuer:	Dr. rer. nat. Christina Waldsich

Wien, im August 2010

Table of Content

1   Introduction

1.1   RNA folding

Throughout the past three decades the role of RNA has undergone a dramatic change in science. Upon discovery RNA was long seen in the perspective of a messenger to pass genetic information from DNA to protein synthesis. Therefore not much thought had been dedicated these days to the structure of RNA molecules. A shift of paradigm occurred with discovery of the first catalytic RNAs in the 1980s by Thomas R. Cech and Sidney Altman (Guerrier-Takada et al., 1983); (Kruger et al., 1982). In their ground breaking work they found that a group I intron is capable of catalyzing its own excision from precursor RNA and that the RNA component of the RNase P complex is the crucial catalytic part. These findings shook the belief in proteins being the only catalytically active players (Guerrier-Takada et al., 1983); (Kruger et al., 1982). From this point on evidence for RNA being involved in a lot of cellular processes grew with overwhelming speed. Several new classes of RNA were discovered, each of it seems to play a role in regulating cellular processes and nowadays it is commonly accepted that RNA molecules are involved in almost every aspect going on in a living cell (Harrison et al., 2009). These new facts strengthened the hypothesis that adopting a certain structure is an important prerequisite for a RNA molecule in order to carry out its task. First thoughts on structure being important came up when investigating the first crystals of tRNAs in the 1970s.

To address the topics of RNA structure and catalytic function one has to keep in mind the chemical properties of RNA. Each RNA molecule is composed of nucleotides and each of the nucleotides contains a ribose sugar, a phosphate and one of the four bases adenine, cytosine, guanine and uracil. The carbon atoms of the ribose molecule are numbered 1' to 5' and the phosphate links two riboses while being bound to the 5' position of one ribose molecule and to the 3' position of the other ribose - the phosphodiester bond. From a chemical point of view the bases can be divided into purines (adenine and guanine) and pyrimidines (cytosine and uracil) and are attached to C1' position of the ribose. This built-up grants each RNA molecule an explicit orientation, namely 5' to 3'.
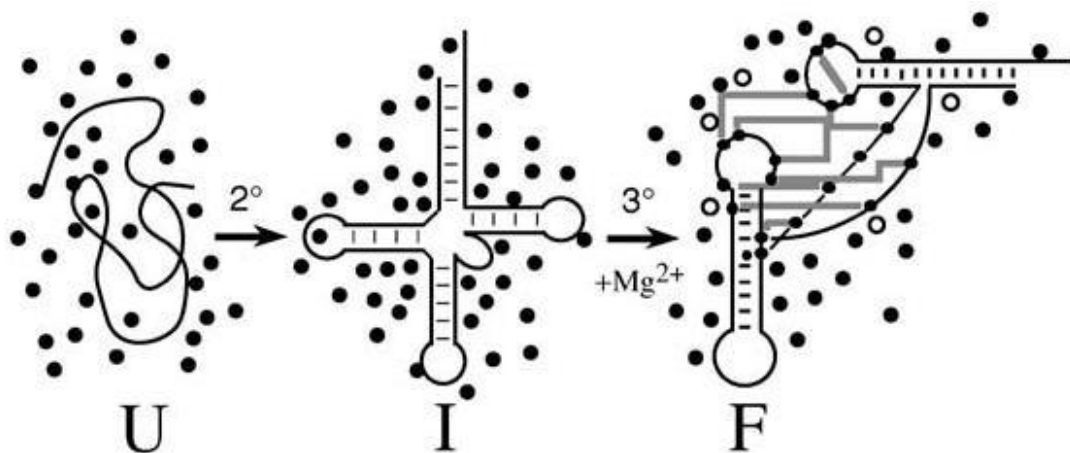

Figure 1. Schematic view of RNA folding. The unfolded state (U) forms stable secondary structures with the help of monovalent ions and the RNA switches into an intermediate form (I). Upon presence of divalent ions like Mg^{2+} the RNA molecule can occupy its folded state with long-range tertiary interactions.

Draper, D. E. 2005

Each RNA must obtain a well-defined structure with exceptions for the very short siRNAs, miRNAs and piRNAs, and the transition from an unfolded, disordered state to the native, folded state is called folding and is shown in Figure 1. It is hierarchical by the means that secondary structure forms first and is then followed by development of the tertiary structure which can bring elements that are far apart in sequence space in close proximity. Secondary structures are composed of two basic elements: **a)** single-stranded regions (like loops and bulges) and **b)** double-stranded regions which are formed through Watson-Crick base pairs. From a thermodynamic point of view the energy required to form secondary structure elements is larger than those required for tertiary interactions, and therefore, most secondary structure is able to stabilize itself while this is not the case for tertiary interactions (Tinoco and Bustamante, 1999).

However RNA molecules encounter a problem upon their path to native conformation which is called "RNA folding problem" and shown in figure 2. This term describes the two fundamental problems: **a)** RNA molecules may fold into non-native, inactive conformations; such non-native interaction can be very stable, thus trapping the RNA in this misfolded conformation incapable of carrying out its original task within the cell; **b)** the native structure is energetically not favored over other structures; in this case the RNA requires assistance to form and maintain the native state (Herschlag, 1995). The high thermodynamic stability of RNA duplexes depends on the number of bases involved and the G/C-richness of the sequence. Accordingly, opening of helix can take up to minutes or even years (Turner et al., 1990). In terms of RNA folding these long time periods make it pretty clear that *in vivo* a mechanism is needed to deal with RNAs trapped in alternative conformations as well as with RNAs with low thermodynamic stability.
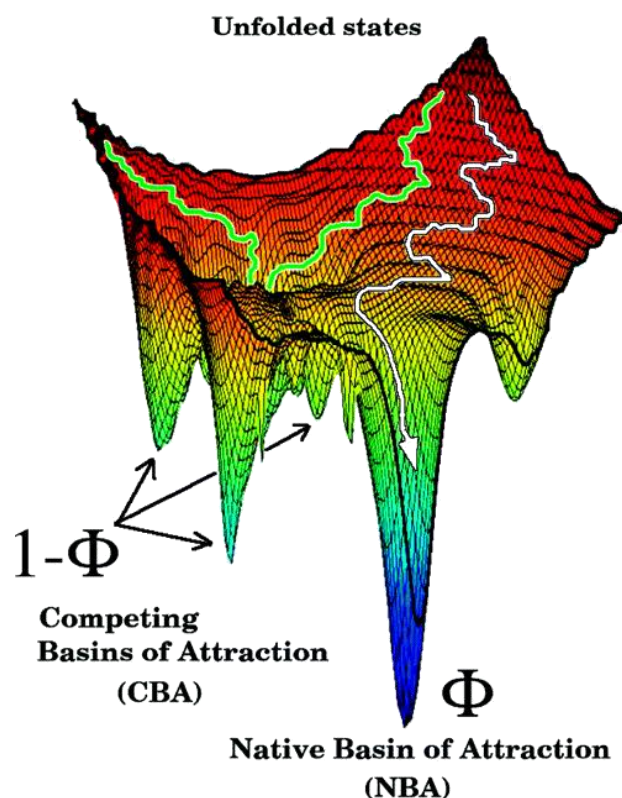
Figure 2. Energetic landscape of a RNA molecule with focus on the folding problem. The native folded state named NBA (native basin of attraction) competes with other energetically favorable folds namely CBA (competing basins of attraction). If the RNA molecule gets trapped along its folding pathway in an alternative basin it may not be able to proceed to the native fold without help. Note that the native state is the energetically most favored, but proteins with chaperone activity may strongly contribute to this final state. Thirumalai, D. & Hyeon, C 2004

Evolutionary, two different players can help a RNA molecule along its folding pathway; **a)** proteins and **b)** metal ions. Proteins can contribute in two ways: first of all non-specific RNA chaperones (e.g. S12, StpA) can resolve misfolded RNAs without requiring an external energy source. After the escape from the kinetic trap the RNA molecule has another chance to reach its native fold (Rajkowitsch et al., 2007; Schroeder et al., 2002; Waldsich et al., 2002). In contrast, specific RNA–binding proteins recognize and bind to a specific sequence or structure of the RNA (e.g. Cyt18, Cbp2, and LtrA), thereby conferring stability to RNA folding intermediates and/or the native conformation (Herschlag, 1995; Schroeder et al., 2004; Waldsich et al., 2002; Weeks, 1997). Upon binding to the RNA these specific cofactor proteins favor a particular structure over non-functional, alternative structures, thereby guiding folding to the native state (Herschlag, 1995; Schroeder et al., 2004; Waldsich et al., 2002; Weeks, 1997).

Aside from proteins metal ions play an important role in RNA folding as well. Especially the neutralization of negative charges of the RNA backbone can be easily achieved by metal ions, thus allowing the RNA molecule to compact and eventually properly folded. A similar function has also been reported for polyamines and it was shown that folding of the *Tetrahymena* ribozyme by polyamines is possible (Koculi et al., 2004). Size and therefore charge of the polyamine molecules can greatly vary and they seem to influence RNA folding in way not completely understood yet. Lastly, it is important to point out that majority of folding studies was carried out *in vitro* under non-physiological conditions meaning increased temperature and high salt conditions.

1.2 Metal ions

With a quick gaze at the periodic table of elements one can see that about two-thirds of all elements can be classified as metals. One of their main characteristics is the low amount of energy needed to ionize them and accordingly metals prevalently exist in nature in their ionic forms. Therefore the multitude of reactions metal ions can participate in have made them an ideal tool to be used by life. This seems to be intrinsically logical due to the high abundance in the primordial soup about 3.8×10^9 years ago and the ability to compensate positive and negative charges. It is generally believed that the concentration of metal ions have been significantly higher in the prebiotic ocean than nowadays (Bengston, 1994). Evidence that metal ions may have played an important role in the RNA world was found in the late 1970s. R. Hach and D. M. Crothers, both working on tRNA, discovered the strong effect of divalent ions like Mg^{2+} on the structural transition from unfolded to a folded state (Römer and Hach, 1975; Stein and Crothers, 1976).

Properties of selected metal ions

Metal	Oxidation state	Hydration number	Ionic radius [Å]
Group 1A			
Li	1+	4	0.59
Na	1+	6	1.02
Ka	1+	6	1.38
Group 2A			
Mg	2+	6	0.72
Ca	2+	8	1.12
Miscellaneous			
Pb	2+	6	1.19
Tb	3+	8	0.92

Figure 3. Table enlisting important characteristics for the most physiologically relevant metal ions. Note the small ionic radius of Mg^{2+} compared to other metal ions important *in vivo.* Pb^{2+} which is often used in folding studies has a much larger radius. This has to be taken into account when performing *in vitro* studies. Adapted from Feig and Uhlenbeck, 1999.

Unlike a polypeptide chain with all side-chains pointing away from the center a RNA duplex has the overwhelming majority of its side-chains pointing towards the molecule's center. A RNA molecule presents itself with a highly negatively charged backbone derived from the phosphate groups and the center region with the bases which can interact with other partners mainly via hydrogen bonds (Herschlag, 1995).

In the absence of ions the RNA molecule in an extended state which changes when ions become available. The process of compaction is known as counter-ion mediated condensation and compensates negative charges from the phosphate groups and therefore prevents unfavorable electrostatic repulsive forces when two or more phosphates come in close spatial proximity. As each phosphate group of the RNA backbone has a formal charge of − 1 it is obvious that each RNA molecule is surrounded by a huge number of metal ions. As mentioned earlier the effectiveness to stabilize tertiary interactions is not the same for mono- and divalent ions. The reasons for this behavior are different ionic radii and charges. Ions with a small ionic radius and therefore a high charge density on its surface have a greater potential for electrostatic interactions and therefore fewer ions are needed to compensate for the same amount of negative charge in contrast to larger ions (Draper, 2004; Woodson, 2005a).
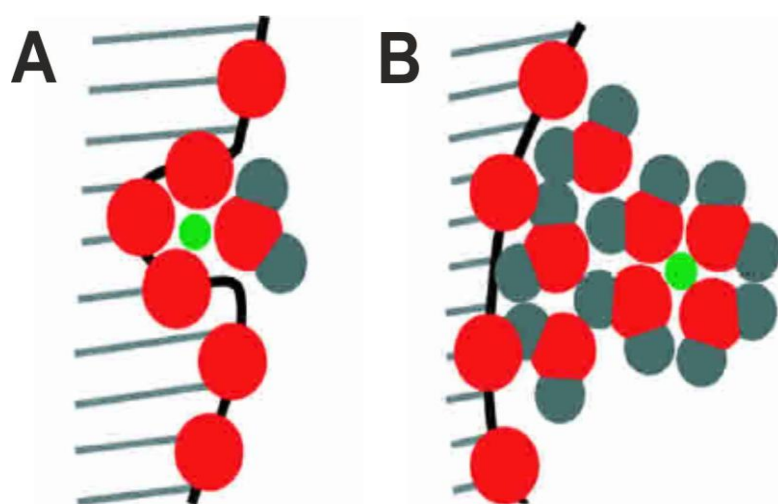

Figure 4. Interactions between metal ions and RNA. Inner-sphere contact (A) versus outer-sphere interaction (B). Oxygen from phosphate groups of the RNA backbone are depicted as red circles, the metal ion is represented by green dot and water molecules are colored red and grey. Adapted from Draper, D. E. 2004

To get a complete picture of how metal ions contribute upon RNA folding it is important to be aware of differences how metal ions can interact with RNA. Defined by their spatial position one can distinguish between delocalized or also termed diffuse ions and site-specific bound ions. Each group differs in their interactions with the RNA backbone or bases and with water molecules. Site-specific bound ions can form both inner-sphere as well as outer-sphere contacts with parts of the RNA molecule, while delocalized ions interact with the RNA exclusively by the formation of outer-sphere contacts. Figure 4 shows the difference between these two interaction types: a metal ion engages in an inner-sphere contact with the RNA backbone, if the Mg^{2+} directly interacts with an RNA atom and which lacks its hydration shell (or part thereof; Figure 4A). In contrast, a metal ion is involved in an outer-sphere contact, if it interacts indirectly with an RNA atom through its hydration shell backbone (Figure 4B).

1.3 Delocalized ions

In the cell ions like Na^+ and K^+ and also Mg^{2+} are surrounded by aqueous solvent and they strongly interact with H_2O molecules thus having a hydration shell. This is true in the presence of single-strand and duplex DNA as well as RNA (Braunlin, 1995). Ions showing such characteristics are often called the "ion atmosphere" of the RNA or in more technical terms "diffuse ions". By definition these ions do not interact directly with the backbone of the RNA molecule, rather they form long-range electrostatic interactions through water molecules and the phosphate groups of the RNA. Looking at a single ion will show that it is not bound to a fixed place but following the electrostatic field of the RNA backbone resulting in areas with different concentrations of ions. The strength of the electrostatic potential and the ion concentration correlate positively and, for example, the highest potential as well as the highest concentration are found deep within the major groove of RNA (Draper, 2004). It has been shown that the sum of non-specific interactions between "diffuse ions" and the RNA contribute much more in stabilizing the global fold of the RNA than site-specifically bound ions within the RNA structure. This is due to the high energetic cost for removing the hydration hull of such chelated ions (Misra et al., 2002). Calculations for such models are usually based on a nonlinear Poisson-Boltzmann equation which describes electrostatic interactions between molecules in an ionic solution.

1.4 Site-specific ions

The opposite of free diffuse ions mentioned above are chelated ions. In this situation the metal-ion is coordinated in a clear defined place in close proximity of the RNA and/or the backbone with which it interacts. In order to contribute to the formation of long-range tertiary interactions within the RNA molecule three different kinds of energetics must be taken into account:

a) The first one is the partial removal of the hydration shell, which encloses the metal ion, referred to as *Hydration (ΔG$_{hyd}$).* This process requires energy is needed making this process energetically costly. As smaller ions tend to have larger hydration shells with two or more layers of water molecules one can understand that more energy is needed to partly dehydrate Mg^{2+} than K^+ or Na^+ due to their differences

in ionic radii. It is also worth mentioning that removal of water molecules is not only needed around the metal ion, of course, this is also the case for the RNA backbone which also interacts with water molecules. Although the energy needed for this purpose has a tendency to be much smaller than for dehydration of the ion itself.

b) The second form of energetics which must be considered are those between the chelated ion and other diffuse ions associated near the RNA backbone, or ($\Delta G_{ion\text{-}diffuse}$). The chelated ion and the ions forming the "ionic atmosphere" around the RNA tend to influence each other strongly via electrostatic interactions. Energetically speaking such interactions are not favorable for two reasons. First, the metal ion in its binding site decreases the number of diffuse ions which can electrostatically interact with the RNA backbone via their hydration shell, therefore resulting in a net unfavorable free energy. This effect is much stronger than the entropically favorable reduction of diffuse ions in sheer terms of numbers. The second unwanted effect which also contributes to the net unfavorable free energy is due the repulsive nature of electrostatic interactions between the chelated ion and the free diffuse ions. The magnitude of energy required for these processes is nearly the same than for dehydration, tens of kcals/mole (Draper, 2004).

c) With the first and second forms of energetics described above being energetically unfavorable it is clear that the third one must shift the balance towards energetically favorable. This is achieved by inner-sphere contact between the chelated metal-ion and the RNA molecule itself, or *($\Delta G_{ion\text{-}RNA}$)*. The overwhelming majority is due to interactions between the ionic charge and the electrostatic field of the RNA.

The experimental setup which is required to identify chelated Mg^{2+}-ions within a certain RNA molecule is often enormous. Crystallization studies are in most cases the method of choice but they require great experience and knowledge. One problem is the similarity of Mg^{2+} electron density peaks to those of hydrated sodium or water (Feig and Uhlenbeck, 1999). Several criteria have been proposed to distinguish between Mg^{2+} ions and its look-alikes, for instance the height and size of the electron peak or the coordination geometry of the investigated species (Holbrook and Kim, 1997). Furthermore it is also possible to soak the crystals for X-ray crystallization with solutions containing cobalt-hexamine or thallium thus replacing Mg^{2+} and K^+ ions, respectively. These ions can be identified much more easier than their predecessors (Gill et al., 2006; Holbrook and Kim, 1997).

Of course the frequency of chelated metal ions is much lower than for diffuse ions due to all negative interactions resulting in a net unfavorable free energy which must be offset by the binding site of the metal ion. As an example a 58mer from the 23S rRNA from *E. coli* has 11 crystallographically observed Mg^{2+}-ions, but only two of them are chelated by the RNA structure, whereas only one shows an overall favorable free energy for binding when all three different types of energetics are taken into account. The ion is situated in a deep pocket where the electrostatic potential is an order of magnitude higher than for any other metal ion found in this structure (Misra et al., 2002).

In order to study the contribution of metal ions to folding of a RNA molecule it is great advantage to work with a very well known model system. This is true for instance for a special class of RNAs called group II introns.

1.5 Group II Introns

In order to sustain life the removal of intronic sequences from precursor RNA is necessary. For the majority of pre-RNA this process is carried out by a large machinery called spliceosome, which itself consists of snRNAs and associated proteins (Jurica and Moore, 2003). The discovery of catalytic RNAs, in particular group II introns in the 1980s (Kruger et al., 1982) revealed a group of RNA molecules, group I and group II introns, whose splicing reaction did not require the spliceosome. These molecules are large, autocatalytic molecules capable of excising themselves from their precursor RNA. Group II introns are very abundant in bacteria, fungi as well as in organelles of plants and lower eukaryotes, but there is no evidence for group II introns in the genome of animals (Fedorova and Zingler, 2007). That group II introns share remarkable similarities in the splicing mechanism with nuclear spliceosomal introns may point to a common evolutionary ancestor (Michel and Ferat, 1995).
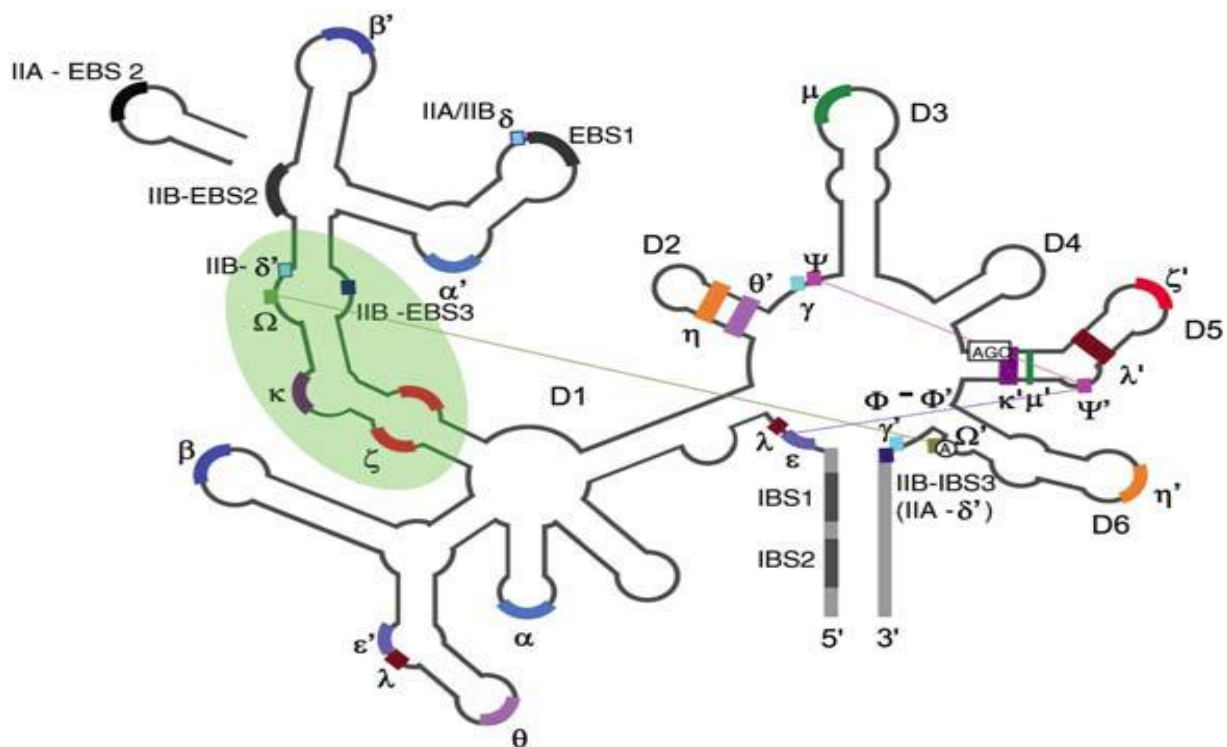
Figure 5. Secondary structure map of a group II intron. The colored regions show long-range tertiary interactions and are labeled by Greek letters. Domains 1 to 6 are labeled D1 to D6 and the boxed letters in D5 and D6 show the catalytic bases and the branch-point adenosine, respectively. The green ellipse highlights the folding control element. Most of the information shown here is based on results obtained from mitochondrial ai5γ intron of *S. cerevisiae.* Fedorova, O. & Zingler, N. 2007

Furthermore, it has also been suggested that group II introns are predecessors of modern non-LTR retrotransposons due to the fact that group II introns are autonomous mobile elements as well (Lambowitz and Zimmerly, 2004; Robart and Zimmerly, 2005). These two facts of close evolutionary relation to elements which account for approximately 45% of the human genome in case of non-LTR retrotransposons make group II introns a very interesting model system to gain new insights in RNA biology and biochemistry.

Additionally, some group II introns contain an open reading frame in domain 4 encoding a maturase protein which is important for splicing of some group II introns under near-physiological conditions *in vitro* but also *in vivo* (Lambowitz and Zimmerly, 2004). Another important function of the maturase is that is promotes intron homing. Domain 4 also harbors the binding site for the maturase protein but due to its location on the surface of the folded intron it may play a more general role as protein-binding element (Fedorova and Zingler, 2007). It is worth mentioning that group II introns are not very conserved in sequence but in contrast their secondary structure organization shows a very high degree of conservation. Based on secondary structure group II introns can be divided in three major groups, the IIA, IIB and IIC introns, and two minor groups, the IID and IIE introns (Michel et al., 1989; Toor et al., 2001). The differences between these families are, for example the location of the EBS2 site between IIA and IIB, and the size and secondary structure of domain 3. For the IIC group, which is believed to be the most ancient and therefore the possible ancestor of group IIA and IIB, some major differences were found; domain 5 which is usually the most conserved among all group II introns lacks two base pairs and there are also substantial differences in positioning of the 5`-splice site at the active site as well (Fedorova and Zingler, 2007). An example for a very well characterized group II intron is ai5γ from *S. cerevisiae.* It is located in the mitochondrial genome, in more detail in the Cox1 gene which is coding for the cytochrome oxidase 1 and therefore its correct splicing is absolutely important for the mitochondrial respiratory chain. The majority of studies investigating group II introns were performed using ai5γ as a model system.

1.6 Splicing reactions

The splicing of group II introns has been well characterized *in vitro*. Depending on metal ions the splicing product forms a lariat structure (Daniels et al., 1996). This lasso-like structure is interestingly observed for spliceosome-mediated splicing of nuclear introns as well (van der Veen et al., 1986). Although the conditions needed *in vitro* are strictly non-physiologically in terms of salt concentration and temperature it has been shown that there is no major difference in mechanics between *in vitro* and *in vivo* once the intron is folded (Podar et al., 1998). Although the excision of the intron from the precursor RNA is catalyzed by the intron itself there is still a need for co-factors to contribute. These additional factors are proteins most of the time and group II introns are known to code often for a maturase protein which is involved in intron splicing (Lambowitz and Zimmerly, 2004). However there is a second class of proteins are essential for intron splicing *in vivo*, but these proteins are encoded in the nuclear

genome (Lambowitz et al., 1999). In contrast to the co-evolved maturase protein the recruitment of host proteins is usually to be thought to be a young evolutionary adaption allowing group II introns to proliferate within new environments (Lambowitz and Zimmerly, 2004). Although the basic mechanisms how certain proteins interact with splicing are not clear yet, there is a well studied example called Mss116p from *Saccharomyces cerevisiae.* It belongs to the family of DEAD-box proteins and is absolutely essential for splicing of all yeast mitochondrial group II introns, including ai5γ (Huang et al., 2005). Another example is MRS2p which is the yeast mitochondrial membrane Mg^{2+} transporter protein. MRS2p is responsible for regulating the mitochondrial Mg^{2+} concentration and was found to decrease group II intron splicing activity when not being present (Gregan et al., 2001). Nevertheless it is important to keep in mind that the splicing reaction *per se* is carried out by the catalytic RNA with protein collaborators only as helpers. A brief summary of splicing reactions catalyzed by group II ribozymes is shown in figure 6.

Forward splicing

The splicing reaction of group II introns is composed of two transesterification reactions and results in the excision of the group II intron from its precursor RNA (Jarrell et al., 1988b; Peebles et al., 1987). Based on the nucleophile which carries out the nucleophilic attack in the first transesterification reaction two alternative splicing pathways can be distinguished; **a)** the branching pathway and **b)** hydrolytic splicing.

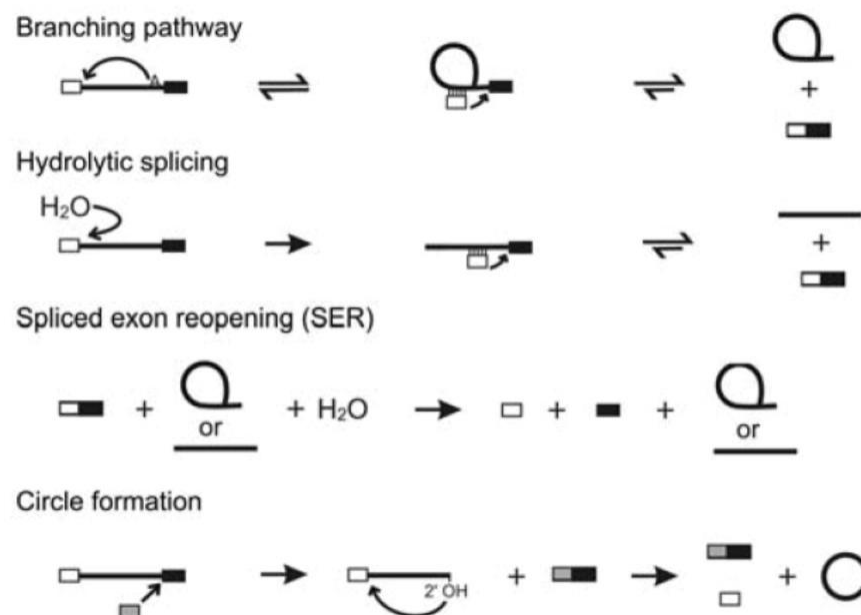**Figure 6**.Overview of splicing reactions of group II introns. The branching pathway generates lariat forms of the intron while hydrolytic splicing leads to the linear form of the intron. Two rather unusual splicing reactions, SER and circle formation are also depicted. Exons are shown as boxes (white, 5' exon; black, 3' exon; grey, free 5'-exon in *trans*), the intron as black line. Adapted from Fedorova and Zingler, 2007.

Branching pathway

The first step of splicing is initiated by the 2´-hydroxyl group from a bulged adenosine located in D6 which is highly conserved. It makes a nucleophilic attack at the phosphate at the 5´-end of the intron resulting in the release of the 5´-exon and the lariat intron – 3' exon intermediate. The chemical mechanism for the reaction described is S_N2 (Padgett et al., 1994) and it is the slower one of the two transesterification reactions thus being the rate-limiting step (Daniels et al., 1996). In the second step of splicing a nucleophilic attack is driven by the 3´-OH of the released 5´-exon which is bound through base pairing to the lariat structure. The target is the 3´-splice site resulting in the release of the intron in its lariat form as well as the ligation of the two exons. The intron in its lariat form is kept together by a 2' 5' linkage.

Hydrolytic splicing

Based on early *in vitro* studies a second splicing pathway was proposed in addition to the branching pathway. The excision of the intron from the precursor RNA is catalyzed by a water or hydroxyl ion which takes over the role of the branch point adenosine (Daniels et al., 1996; Jarrell et al., 1988b; van der Veen et al., 1987). In contrast to the lariat splicing pathway the intermediate is composed of the linear intron with the still attached 3´-exon and the free 5´-exon positioned via base pairing for the second nucleophilic attack. In this case the linear form of the intron and the exons are ligated as well. When first observed it was not clear what parameter can shift the balance between these two splicing reactions. Meanwhile it is known that the choice of monovalent cations in the reaction is key to that question (Daniels et al., 1996). Introns which do not have a branch-point *per se* can maintain splicing efficiency through the hydrolytic splicing pathway which suggests that hydrolytic splicing has biological relevance as well (Vogel and Borner, 2002).

Reverse splicing

It has been shown that both of the reactions in the splicing pathway of group II introns are reversible (Augustin et al., 1990; Mörl and Schmelzer, 1990). However, the rate of reverse splicing is not the same for the lariat form and the linear form, respectively. It has recently been shown that the linear form of the ai5γ intron is capable to catalyze reverse splicing at an astonishing rate (Roitzsch and Pyle, 2009). This contradicts the general belief that linear introns show no or only poor reverse splicing kinetics. Furthermore, it is also interesting that reverse splicing is not only restricted to RNA substrates, it also works with DNA targets. Reverse splicing into DNA is the first step towards intron homing facilitated by the maturase (Griffin et al., 1995; Mörl et al., 1992).

Unusual splicing reactions

Besides the branching and hydrolytic splicing pathways there are some other, rather unusual reactions. Most of them are not very well characterized and also their *in vivo* relevance is not yet known. Among these unusual reactions are; **a)** Spliced exon reopening (SER), which involves the hydrolysis of the pre-

vious ligated exons exactly at the 5´-3´ splice site junction (Jarrell et al., 1988a). This reaction seems to be rather uncommon in IIA introns but is often observed for IIB introns (Hebbar et al., 1992; Schmidt et al., 1990); **b)** Circle formation which was described almost 20 years ago and initially thought to be the product of splicing (Osiewacz and Esser, 1984). Here again the biological function is not known.

1.7   Structural organization of group II introns

As mentioned earlier group II introns are composed of six different domains which protrude from a central wheel. Group II introns are also very large molecules ranging between 500 and 1000 bases placing them among the largest catalytic RNAs known. Each domain contributes to either the efficiency of splicing, conformational rearrangement or catalysis. The majority of the structural information available by today had been obtained from phylogenetic and biochemical analysis of interactions within the group II introns. Recently there was also a high resolution crystal structure at 3.1 angstrom published which reports an intact, self-spliced group IIC intron from *Oceanobacillus iheyensis* (Toor et al., 2008).

Domain 1

Being the largest of all domains D1 is long known to be absolutely essential in order to have a functionally active group II intron. Its importance can be explained by the fact that D1 provides a scaffold for all other domains to dock on. Furthermore recognition of the 5´-exon as well as for the 3'-exon is also performed by D1 mainly via EBS1 and EBS2 sites which ensure correct positioning by base pairing with the corresponding regions IBS1 and IBS2 on the 5´-exon (Qin and Pyle, 1998). Mismatch analysis from EBS1-IBS1 and EBS2-IBS2 revealed that even disruption of a single base pair may lead to drastically decreased cleavage efficiency (Xiang et al., 1998). D1 itself is characterized by the five-way junction from where stems a to i protrude and by several intra-domain tertiary contacts like α-α', β-β' and ε-ε'. Additionally, the θ-θ' interaction between D1 and D2 seems to be important for structural stabilization of the intron as well as for recruitment of D3 and J2/3 into the active site (Costa et al., 1997; Fedorova et al., 2003). Domain 1 also harbors three long-range inter-domain tertiary interactions named ζ-ζ´, κ-κ´ and λ-λ´ and all of them target the catalytic center of the group II intron, domain 5 (Boudvillain and Pyle, 1998; Boudvillain et al., 2000; Costa and Michel, 1995). These interactions are responsible for docking of D5 on to D1 as well as spatial positioning D5 in order to allow catalysis. Furthermore D1 provides structural elements which are critical for the folding process of the entire group II intron and it is known that D1 folding is the rate-limiting step for folding of the ai5γ intron (Su et al., 2005).

Domain 2

Very little is known about the role of this domain however it establishes long-range tertiary interactions to D1 (θ-θ') as well as D6 (η-η') (Chanfreau and Jacquier, 1996; Costa et al., 1997). Domain 2 shows the largest variation in sequence as well as in size and the few data available suggest that the η-η' interaction acts like a switch between different conformation during 1st and 2nd step of splicing (Chanfreau and Jacquier, 1996). Furthermore D2 also plays an important role in positioning the structural element J2/3 located between D2 and D3 in the active site of the ribozyme (Toor et al., 2008).

Domain 3

This domain is not strictly required for catalysis but its presence substantially accelerates the rate of catalysis (Fedorova et al., 2003; Griffin et al., 1995; Koch et al., 1992; Xiang et al., 1998). Like other domains D3 harbors a long range tertiary interaction (μ-μ') which involves two adenosine residues from the D3 pentaloop and the 2'-hydroxyl group of G844 located in D5 (Fedorova and Pyle, 2005). Interestingly, D3 contains one of the most conserved elements of a group II intron, an adenosine-rich internal bulge which is essential for catalytic activity (Michel and Ferat, 1995).

Domain 4

As mentioned previously, in some group II introns D4 harbors the open reading frame for a maturase protein, which is involved in intron splicing as well as reverse splicing (Lambowitz and Zimmerly, 2004). In addition, D4 contains a binding site for the maturase protein but it has also been shown that other intron-specific proteins can dock onto this domain (Matsuura et al., 2001; Watanabe and Lambowitz, 2004). Aside from carrying an open reading frame D4 has no other known function. For full catalytic activity of a group II intron D4 is not required.

Domain 5

Besides D1 this is the second domain absolutely required for catalysis and its absence leads to the complete loss of catalytic activity of every group II intron (Fedorova and Zingler, 2007). Furthermore, D5 is also the smallest domain with only 34 nucleotides as well as the most conserved one (Michel and Ferat, 1995). It harbors two very important structural elements: **a)** the so-called "catalytic triad" AGC at positions 815-817 (ai5γ numbering); and **b)** the dinucleotide bulge at positions 838-839 (ai5γ numbering) of the basal stem. Both of these elements are critical for splicing; the catalytic triad with its invariant guanosine residue is required for both *in vivo* and *in vitro* splicing reactions (Boulanger et al., 1995; Peebles et al., 1995), while the dinucleotide bulge harbors a magnesium ion-binding site which is believed to be important for catalytic activity (Schmidt et al., 1996). The crystal structure (figure 7, panel A) of a group II intron from *Oceanobacillus iheyensis* revealed that those two elements are brought close together in the folded intron resulting in the formation of a negatively charged pocket which can coordinate two magnesium ions (Toor et al., 2008). With the ideal distance of 3.9 Å (figure 7, panel B) between the two magnesium ions they perfectly match the prerequisites for a classic two metal ion

reaction mechanism observed in group I introns as well as for DNA and RNA polymerases (Steitz and Steitz, 1993). Furthermore, the linker region between D2 and D3 called J2/3 is also brought in close spatial proximity of the catalytic triad and the dinucleotide bulge forming a catalytic triplex which contains all the residues absolutely important for catalysis (Toor et al., 2008). Interestingly the AGC catalytic triad is also found in U6 snRNA (Valadkhan and Manley, 2001). Like the other domains D5 also harbors tertiary interactions. These inter-domain long-range tertiary interactions κ-κ' (Boudvillain and Pyle, 1998) and ζ-ζ' (Costa and Michel, 1995) which are both tetraloop-receptor interaction targeting D1 are responsible for firmly anchoring D5 onto D1. The third long-range tertiary interaction λ-λ' which also targets D1 brings D5 in correct position with respect to the 5'-splice site (Boudvillain et al., 2000).
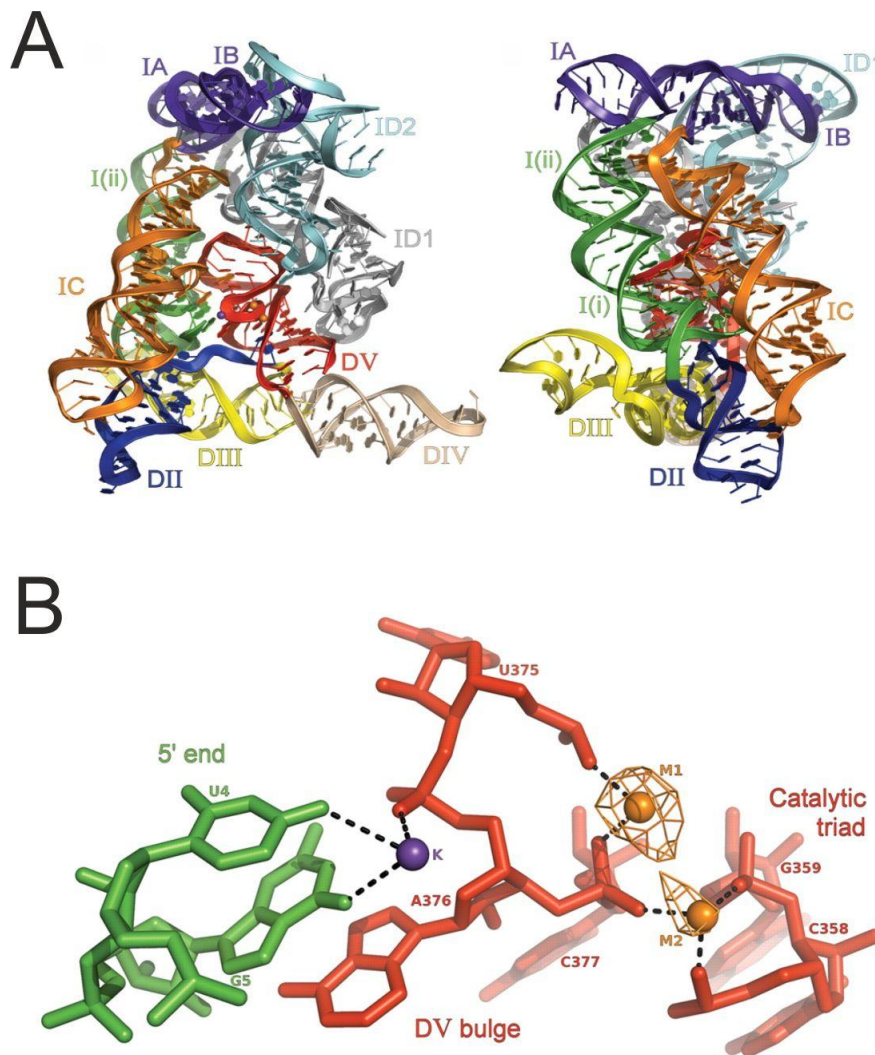
Figure 7. Panel A shows the crystal structure of an intact, self-spliced group IIC intron at 3.1 angstrom resolution from *Oceanobacillus iheyensis*. On the left side the intron is shown in ribbon drawing with color-coded domains; the right side is the image shown on the left side rotated by 90°. Panel B is a representation of the catalytic triad and its interactions with two metal ions (M1 and M2) as well as the interactions between the 5' end and a K^+ ion in the bulge region.   Adapted from Toor, 2008.

Domain VI

Although containing the branch-point adenosine which carries out the nucleophilic attack in the first transesterification reaction of the branching pathway, D6 *per se* is not very conserved among group II introns (Lehmann and Schmidt, 2003). This is a very interesting fact given the importance of the branch-point adenosine for *in vivo* splicing although the residue itself is of course highly conserved throughout the family of group II introns. D6 has only one inter-domain long range tertiary interaction (η-η' to D2).

1.8   Folding pathway of ai5γ in vitro

The majority of what is known about the folding pathway for the group II intron ai5γ comes from studies on the D135 ribozyme *in vitro* (Su et al., 2001). This ribozyme consists of domains 1, 3 and 5, whereby domains 2 and 4 were shortened to small stable hairpins and domain 6 and the exons were deleted. Early findings using hydroxyl radical footprinting showed that almost half of the intron is protected from cleavage (Swisher et al., 2001) and therefore suggested that ai5γ RNA is tightly folded. Surprisingly the regions which are most protected are also the most conserved ones (Swisher et al., 2001) (Pyle et al., 2007). Furthermore all footprints evolve comparably as a function of divalent ion concentration and time (Swisher et al., 2002); in other words, all internalized regions become protected at the same Mg^{2+} concentration and on the same timescale. Furthermore urea denaturation studies performed with the D135 ribozyme show that the folding of D135 is highly reversible, thus allowing the folded RNA to return to the native state after unfolding, and as expected similar thermodynamic parameters between folding and unfolding were observed (Su et al., 2003; Swisher et al., 2002). Thus, it was first suggested that D135 follows a two-state folding pathway. However some characteristics of D135 folding were puzzling; first it was generally believed that with increasing length of a RNA molecule the chances of getting kinetically trapped in a misfolded state increase as well and in turn the escape from such a kinetic trap would be the rate-limiting step along the folding pathway (Treiber and Williamson, 1999). Curiously the D135 ribozyme combines a slow folding rate (1 min^{-1}) with evidence that no kinetic trap is present (Fedorova et al., 2007; Su et al., 2005; Swisher et al., 2001; Swisher et al., 2002). This unusual behavior has recently been explained in part by a chemogenetic study (NAIM) on D135 folding (Su et al., 2005; Waldsich and Pyle, 2007; Waldsich and Pyle, 2008). These further folding studies with other ai5γ-derived ribozymes showed first that the rate-limiting step overall for D135 folding is compaction of D1, revealing that D135 folding involves a two-step mechanism with at least 3 distinct states (Fedorova and Zingler, 2007; Su et al., 2005). In other words, D1 formation constitutes an intermediate state between unfolded and native conformation. Employing NAIM Pyle and coworker then showed that a small substructure, the κ–ζ element, at the core of D1 appears to dictate folding of the entire ribozyme and thus it was termed folding control element (Waldsich and Pyle, 2007; Waldsich and Pyle, 2008). In addition, it has been suggested that the κ–ζ element induces compaction of D1 and the compact state is then locked in by the intra-domain tertiary contacts α-α' and β-β' (Waldsich and Pyle, 2007; Waldsich and Pyle, 2008). Importantly, the folding control element serves as a docking site

for the catalytic center during later stages of D135 folding (Boudvillain et al., 2000; Costa et al., 1998). In brief, proper folding and positioning of this key element seems to trigger compaction and moreover folding of the D135 ribozyme (Waldsich and Pyle, 2007) and suggests that folding is not limited by the formation of long-range interactions. The folding pathway of the ai5γ intron is summarized in figure 8.
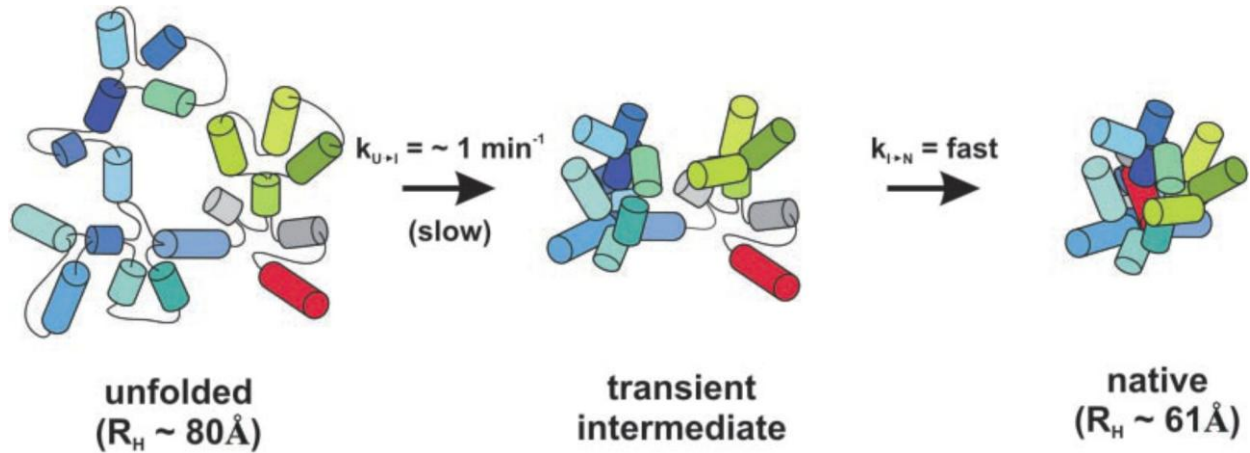

$k_{U \to I} = \sim 1\ min^{-1}$ (slow) $k_{I \to N} = fast$

unfolded ($R_H \sim 80$Å)   transient intermediate   native ($R_H \sim 61$Å)

Figure 8. *In vitro* folding pathway of the ai5γ ribozyme D135. First the intron collapses slowly to adopt a compact intermediate state, which is limited by the formation of the κ–ζ element. Then the other intron domains can dock rapidly onto the D1 scaffold resulting in the adoption of the native state. Adapted from Su et al.,2005.

Both folding and compaction of the ribozyme require high concentrations of Mg^{2+} ions (Su et al., 2005; Swisher et al., 2002) and it is suggested that the formation of the folding control element is mediated by metal ion binding in this area (Waldsich and Pyle, 2008). *In vivo* metal ions are not present in such high concentrations as they are needed for *in vitro* studies. This clearly points out that there must be some differences between the folding pathway of ai5γ *in vitro* and *in vivo.* One of several possible options are proteins which can stand in for metal ions. For example it is well-known that Mss116p is essential to promote efficient splicing of the ai5γ intron in yeast mitochondria (REF: Huang et al 2005).

1.9 Mss116p and its role for intron folding

Mss116p was identified in the 1980s as a gene which affects yeast mitochondrial function (Faye and Simon, 1983) and later on as a member of the DEAD-box protein family (Seraphin et al., 1989). This protein family is large and their members have in common the usage of energy derived from ATP hydrolysis to mediate structural rearrangements in RNA molecules (Silverman et al., 2003). They have a core region consisting of nine conserved motifs flanked by unique N and/or C-terminal extensions which can help to guide the proteins to their sites of action via RNA or protein interactions (Huang et al., 2004). The family is named after a certain amino acid sequence in motif II which is DEAD or DEAH or other variants thereof. Figure 9 shows the build-up of the DEAD-box protein Mss116p. The conserved helicase motifs are located on the N-terminal domain and in addition it harbors an arginine-rich C-terminal domain. Several aspects like ATPase, helicase, nucleic acid binding as well as remodeling activities of some family members have been studied very well *in vitro* (Cordin et al., 2006), but only one of

the members, DbpA, was characterized in the aspect of a biological relevant substrate (Diges and Uhlenbeck, 2001). Therefore very little is known about the role these proteins play inside the cell. Mss116p is one of three DEAD-box proteins encoded in the yeast nuclear genome with Suv3p and Mrh4p being the other two which function in mitochondria (Margossian et al., 1996; Schmidt et al., 2002; Seraphin et al., 1989). Mss116p resembles another DEAD-box protein, CYT-19, encoded in the genome of the fungus *Neurospora crassa* which is needed for splicing of mitochondrial group I introns, in more detail for resolving nonnative structures which are caught in a kinetic trap on their folding pathway (Mohr et al., 2002). Furthermore, Mss116p is crucially needed for efficient splicing of both group I as well as group II introns, including ai5γ, in yeast mitochondria (Huang et al., 2004) and it was also shown that overexpression of CYT-19 can rescue Mss116p deletion mutants (Huang et al., 2004).
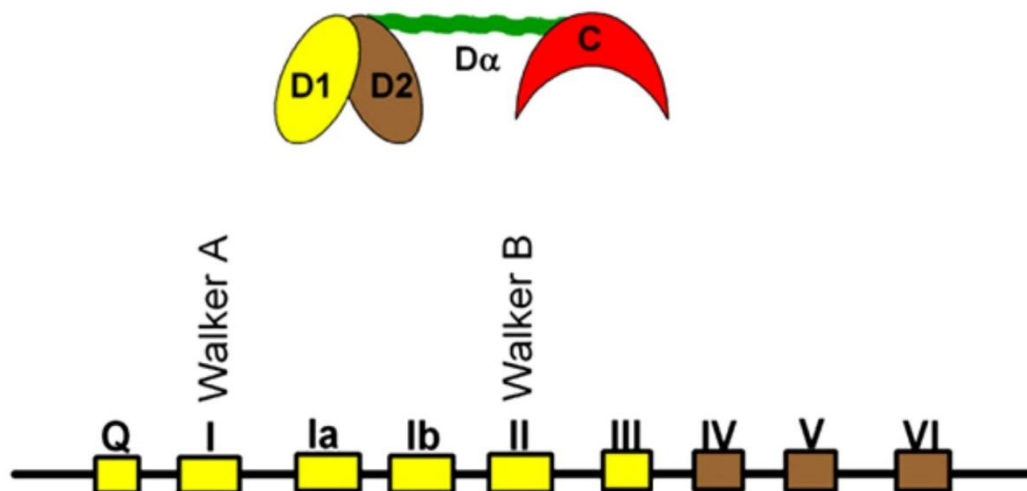

Figure 9. Schematic view of the DEAD-box protein Mss116p. The core region consists of two RecA folds (yellow D1 and brown D2) which are linked via an α helix to the arginine-rich C-terminal domain (red). The first RecA fold consists of motifs Q, I, Ia, Ib and II and contains a Walker A as well as a Walker B motif. The second domain is made of motifs IV, V and VI. Adapted from Solem, 2006.

So far, Mss116p has been described to be able to efficiently cleave ATP upon binding RNA and to unwind short duplexes which can have either a 3' single-stranded overhang, a 5' single-stranded overhang or a blunt end (Solem et al., 2006). Furthermore, Mss116p was shown to lower the Mg^{2+} requirements for splicing of the ai5γ intron *in vitro*. It has subsequently been shown that the ATPase activity is essential for promoting ai5γ intron splicing *in vitro* and *in vivo*, while the importance of helicase activity has been a matter of debate (Del Campo et al., 2007; Halls et al., 2006; Huang et al., 2005; Solem et al., 2006). Based on these findings it has been proposed both that either Mss116p has to resolve misfolded ai5γ RNA molecules or to stabilize RNA folding intermediates. Most recently, Mss116p was found to stimulate ai5γ folding by accelerating the collapse to the near-native state through stabilization of an early folding intermediate (Fedorova et al., 2010). Interestingly, Mss116p is not required to stabilize the native state of ai5γ, but stabilization comes from binding of flanking exons. Along this line, we have recently shown that in the absence of Mss116p the ai5γ intron is largely unfolded *in vivo*, suggesting

that this protein is required for the formation of an early folding intermediate *in vivo* as well (Liebeg & Waldsich, unpublished). Possibly, the folding control element mentioned before act as the biological relevant substrate for Mss116p *in vivo* (Liebeg and Waldsich, unpublished). In addition, the unwinding activity of Mss116p was observed to be essential for exon unfolding, but not for intron folding *in vitro* (Zingler et al., 2010).

1.10 Metal ion binding sites of ai5γ in vitro

Despite the important function metal ions play *in vivo* (Gregan et al., 2001) all studies until this date focused mainly on *in vitro* systems (Woodson, 2005b). As mentioned earlier the group II intron ai5γ needs high temperature and salt concentrations to splice *in vitro* (Fedorova and Zingler, 2007). One particular study by Sigel et al. in the year 2000 focused on metal ion binding sites within the ai5γ intron under such non-physiological conditions. In brief, a ribozyme construct containing domains 1, 3, 5 and 6 was probed with different lanthanide(III) ions such as Tb^{3+} or Lu^{3+} but also with Pb^{2+} and Mg^{2+}. The results are summarized in figure 10 and show the positions where metal ion-induced cleavage occurred.

Figure 10. Summary of Tb^{3+}-induced cleavage sites mapped on the secondary structure of the group II intron ai5γ *in vitro.* Based on the D135 ribozyme domains 2 and 4 are truncated to short hairpins but nomenclature corresponds to those of the intact intron. The intensity of cleavage (3' of the marked nucleotides) is color-coded. Tertiary interactions are marked by small Greek letters. Adapted from Sigel et al., 2000.

Interestingly, the locations of high affinity metal ion binding sites correlates very well with sites which are known to be involved in ribozyme chemistry; for example a cluster of cleavage sites is observed at the branch point adenosine (A880) in domain 6 and its proposed docking site in D1, the coordination loop consisting of Jd''/d''' and Jd'''/d'' (Hamill and Pyle, 2006; Sigel et al., 2000). Also, a metal ion seems to be in proximity to the γ-γ' (A587 and U887) interaction, which plays a role in 3'-splice site recognition and cleavage (Sigel et al., 2000). In addition, Tb^{3+}-induced cleavage was detected at the di-nucleotide bulge in D5 and J2/3, which both engage in a base triple interaction with the AGC-triad in D5 (REF: Toor). Consistent with these sites, cleavage was also observed at the ε' and λ region in stem c1 of D1, which are important active site constituents (Boudvillain et al., 2000; Sigel et al., 2000). Aside from metal ion binding pocket(s) close to the catalytic center, other pockets reflecting metal ions of mere structural importance are observed as well (e.g. in Jc1/c2 or residues close to the β' region). Another very interesting aspect of this study is the finding that there are no metal ion binding sites reported in the area of the folding control element, indicating that either metal ions are not required to stabilize the region in the native state or that Mg^{2+} ions are only essential during the collapse of the ai5γ intron. Indeed in a later study it had been shown that metal ions are in the vicinity of the κ–ζ element during the collapse, but cleavage can no longer be observed in the native state (Waldsich 2008). This observation can be explained by the fact that first it is only possible to detect metal ions in the minor groove of RNA (but not those located in its major groove) and second Tb^{3+}-induced cleavage requires a perfect alignment of Tb^{3+} with the 2'OH of the ribose (cf. Waldsich 2008).Thus, it is possible that the alignment is not met in the native state, which is much more rigid than the collapsed state and in turn precluding cleavage at the κ–ζ element.

2 Aims of the project

Metal ions play an important role for folding of large RNAs like group II introns under non-physiological conditions *in vitro* as well as *in vivo*; for example the Mrs2p protein which is an essential splicing factor for all group II introns by regulating mitochondrial Mg^{2+} concentration (Gregan et al., 2001). The goal of this project was to gain fundamental insight into the formation of metal ion binding pockets *in vivo* using the ai5γ intron as model system. As the DEAD-box protein Mss116p is essential for splicing of ai5γ and other yeast mitochondrial introns, I aimed at assessing the influence of Mss116p on the formation of metal ion binding pockets *in vivo*. To address these questions we had to establish a metal-induced cleavage assay *in vivo* and compared the cleavage pattern for ai5γ in wild-type yeast strain and in a Δmss116 knock-out strain.

3 Methods

3.1 In vivo metal ion-induced cleavage assay

Yeast cultures of strain #16 and #189 were grown on YPD plates at 30° C for three days and afterwards stored at 4° C up to four weeks. For RNA extraction an o/n – culture was inoculated in 5 ml of YPD medium and grown at 30° C. The next day the primary culture was set up in 100 ml YP-Raf medium and grown to OD_{600} 1. The cells were then harvested in 30ml aliquots by centrifugation (6000 rpm for 5 minutes at 25° C). The supernatant was discarded and the pellets were resuspended in 1000µl of either YPD or 80mM MOPS buffer pH 7.

The metal ion-induced cleavage assay was performed as follows:

- Pb^{2+}-induced cleavage assay
  - Cells were resuspended in YPD media.
  - Always prepare a fresh 500mM stock solution of $Pb(OAc)_2$ in ddH_2O.
  - $Pb(OAc)_2$ was added to the cells to give a desired final concentration (25-200mM) in a working volume of 1000µl.
  - Incubate samples at 30° C for 10 minutes shaking at 300 rpm.
  - Stop the cleavage reaction by adding 2.5x molar excess of 0.5M EDTA pH 8.
  - Pellet cells for 5 minutes at 6000 rpm and discard supernatant.
  - Wash cells in 1000µl of YPD media.
  - Pellet cells again (see above) and quickly freeze for 30 min at -80° C.
  - Continue with RNA isolation.

When using Tb^{3+} some adjustments were necessary which are listed below:

- Tb^{3+}induced cleavage assay
  - Cells were resuspended in 80mM MOPS buffer pH 7
  - $TbCl_3$ was dissolved in 5mM cacodylate buffer pH 5.5
  - Cells were washed in 80mM MOPS buffer pH 7 instead of TM buffer

For the assay the following solutions, buffers and media were used:

YPD plates

10 g	yeast extract	[f. c. 1%]
20 g	peptone	[f. c. 2%]
20 g	agar	[f. c. 2%]
20 g	raffinose	[f. c. 2%]
55 mg	adenine	[f. c. 55mg]
55 mg	tyrosine	[f. c. 55mg]

Fill up to 1 l with ddH_2O, autoclave and pour plates

YPD medium

10 g	yeast extract	[f. c. 1%]
20 g	peptone	[f. c. 2%]
55 mg	adenine	[f. c. 55mg]
55 mg	tyrosine	[f. c. 55mg]

Fill up to 800 µl with ddH$_2$O and autoclave it, glucose (20 g/l [f. c. 2%]) was added prior to use.

YP-Raffinose medium

10 g	yeast extract	[f. c. 1%]
20 g	peptone	[f. c. 2%]
55 mg	adenine	[f. c. 55mg]
55 mg	tyrosine	[f. c. 55mg]

Fill up to 800 ml with ddH$_2$O and autoclave it, raffinose (20 g/l [f. c. 2%]) was added prior to use.

0.5M EDTA pH 8

14.61 g   EDTA

Fill up to less than 100 ml, adjust to pH 8 with NaOH, afterwards raise the volume to 100 ml and autoclave it.

1M MOPS pH 7

16.7 g   MOPS

Fill up to less than 100 ml with ddH$_2$O, adjust pH to 6, afterwards raise the volume to 100 ml and sterile filter it. Store at 4° C (in the dark).

1M Cacodylate buffer pH 5.5

21.4 g   Sodium Cacodylate

Fill up to less than 100 ml with ddH$_2$O, adjust pH to 5.5, afterwards raise the volume to 100 ml and sterile filter it. Store at 4° C.

TM buffer

0.5 ml	2M Tris.HCl pH 7	[f. c. 10mM]
1 ml	1M MgCl$_2$	[f. c. 10mM]

Fill up to 100 ml with sterile ddH$_2$O.

2M Tris·HCl pH 8

> 24.2 g   Tris base
>
> Fill up to less than 100 ml with ddH$_2$O, adjust pH to 8 with HCl, afterwards raise the volume
>
> to 100 ml and autoclave it.

3.2   Isolation of total RNA

In order to isolate total RNA from yeast cells the *RiboPure^{TM}-Yeast Kit* from *Ambion$^{©}$* was used. The isolation was performed according to the manufacture's manual. In brief yeast cells were broken by mechanical force (15 minutes vortexing) using Zirconia beads at 4° C, followed by phenol:chloroform:isoamylalcohol extraction. The obtained solution was then filtered through glass-fibre columns, washed in order to get rid of cell debris, salts and eluted using 100µl elution solution. Potential DNA was digested using DNaseI and free nucleotides were chelated by adding 1.5µl 0.5M EDTA pH 8. Finally 250µl 0.3M NaOAc pH 5/EtOH were added for precipitation at – 20° C. The typical yield of RNA was about 200 to 250µg. The following solutions, which were not provided by the manufacturer, were used during this step:

3M NaOAc pH 5

> 204.1 g   Sodium acetate
>
> Fill up to less than 500 ml with ddH$_2$O, adjust pH to 5 with Acetic Acid, afterwards raise the
>
> volume to 500 ml and autoclave it.

3.3   Quality control of RNA

In order to assess the quality of the RNA isolated with the *RiboPure^{TM}- Yeast* Kit, a small amount of ~1µl was loaded on a native 1.5% agarose gel. High quality RNA resulted in two bands corresponding to 25S and 18S RNA. 5.8S RNA as well as tRNAs were present as smeary area at the bottom of the gels.

3.4   Reverse transcription

With the isolated total RNA a reverse transcription was performed. A typical reaction is set up as follows:

- 2.5µl RNA [16µg/µl]
- 1µl 4.5x Hybridization buffer
- 1µl ^{32}P-labelled gene-specific oligo
- Incubate for 1 minute at 95° C (in a Thermomixer).
- Cool slowly to 42° C (takes ~30 minutes).
- Add 2.2µl containing
  - 0.6µl 10x Extension buffer

- o 0.3µl 2.5mM dNTPs

- o 0.15µl AMV Reverse transcriptase [10u/µl]

- o 1.15µl ddH$_2$O

- Incubate for 60 minutes at 42° C.

- Add 1.5µl 1M NaOH in order to degrade RNA.

- Incubate for 60 minutes at 42° C.

- Add the following

  - o 1.5µl 1M HCl in order to neutralize pH

  - o 1µl 0.5M EDTA pH 8 to chelate free Mg^{2+} ions

  - o 1µl glycogen [10µg/µl]

- Precipitate with 30µl 0.3M NaOAc/EtOH for 60 minutes at – 20° C.

- Centrifuge samples at 13000 rpm for 45 minutes at 4° C.

- After drying resuspend the pellets in 8µl PAGE loading buffer.

The pool of cDNA obtained from reverse transcription was separated on an 8% denaturing polyacrylamide gel with either 0.4 or 0.5mm thickness. Running parameters were different corresponding to the different size of the gels. The gel was pre-run for 60 minutes to reach a constant distribution and equilibration of temperature. The typical electrical power for running a gel (20cm x 42cm) was 40W or 65W for gels of the size of 33cm x 52cm for 120 minutes.

Afterwards the gel was transferred to 2 pieces of Whatman 3M paper and covered with Saran wrap. Drying of the gel was done using a vacuum slab gel dryer at 80° C for 150 minutes. The dried gels were then transferred to a Phosphoimager cassette for at least 1 day. The screen was scanned using the phosphorimager *STORM 820* from *Molecular Dynamics (GE Health care)*. The file itself was then analyzed using the software *ImageQuant 5.2* from *Molecular Dynamics*.

Solutions which were used for the single steps are listed below:

4.5x Hybridization buffer

| 45 µl | 1 M K-Hepes pH 7 | [f. c. 225mM] |
| 90 µl | 3M KCl | [f. c. 450mM] |

Fill up to 200 µl with sterile ddH$_2$O and store at – 20° C.

10x Extension buffer

130 µl	2 M Tris·HCl pH 8	[f. c. 1.3M]
20 µl	1 M MgCl$_2$	[f. c. 100mM]
20 µl	1 M DTT	[f. c. 100mM]

Fill up to 200 µl with sterile ddH$_2$O and store at – 20° C.

10mM dNTPs

10 µl	dATP	[f. c. 10mM]
10 µl	dGTP	[f. c. 10mM]
10 µl	dCTP	[f. c. 10mM]
10 µl	dTTP	[f. c. 10mM]

Add 60µl of sterile ddH$_2$O and store at – 20° C.

1M NaOH

19.95 g   Sodium hydroxide

Fill up to 500 ml with sterile ddH$_2$O.

1M HCl

41.6 ml   37% HCL

Fill up to 500 ml with sterile ddH$_2$O.

Glycogen [10 mg/ml]

500 mg   glycogen

Fill up to 50 ml with ddH$_2$O and sterile filter it. Store at -20° C.

8% Acrylamide solution

100 ml	40% Acrylamide solution 19:1 [f. c. 8%]	
210 g	Urea	[f. c. 7M]
50 ml	10x TBE	[f. c. 89mM]

Fill up to 500 ml with ddH$_2$O and filter it. Store at 4° C.

10x TBE buffer

108 g	Tris	[f. c. 890mM]
55 g	Boric acid	[f. c. 890mM]
5.8 g	EDTA	[f. c. 20mM]

Fill up to 1 l with ddH$_2$O, filter and then autoclave it.

PAGE loading buffer

21 g	Urea	[f. c. 7M]
12.5 g	Sucrose	[f. c. 25%]
0.125 g	Bromphenol blue	[f. c. 0.25%]
5 ml	10x TBE	[1x]

Fill up to 50 ml with ddH$_2$O.

1M K-HEPES pH 7

> 23.8 g   HEPES
>
> Fill up to less than 100 ml with ddH$_2$O, adjust pH to 7, raise the volume to 100 ml and sterile filter it. Store at 4° C.

3M KCl

> 22.4 g   KCl
>
> Fill up to 100 ml with ddH$_2$O and autoclave it.

1M MgCl₂

> 20.3 g   MgCl$_2$
>
> Fill up to 100 ml with ddH$_2$O and autoclave it.

3.5 Purification of RT oligos

The synthesized oligo was diluted in ddH$_2$O to a stock concentration of 100µM. 40µl 100µM oligo and 40µl loading buffer were mixed and loaded on a 20% acrylamide gel (25 Watts; 1mm spacer). The gel was then wrapped in Saran wrap and transferred onto a TLC plate with a fluorophore. The band corresponding to the oligo was cut out using a sterile blade. The gel slice containing the oligo was placed in a 1.5 ml Eppendorf tube, after adding 1.5 ml elution buffer the sample was incubated at 37° C for 4 hours. After centrifugation at 13000 rpm for 2 minutes the supernatant was transferred to a new tube and precipitated at − 20 ° C using 4500µl EtOH p. A.. The pellet was then resuspended in 15µl ddH$_2$O.

The concentration of a 1:1000 dilution was measured at 260nm and calculated according to the following formula:

$$\frac{A_{260} * Dilution * 10^6}{\varepsilon} = \mu M$$

For calculating the extinction coefficient ε the following term was used:

$$\varepsilon = \left(15346{,}6 * n_A + 11751{,}8 * n_G + 7625{,}2 * n_G + 9929{,}7 * n_U\right)$$

where n represents the sum of the corresponding base in the nucleotide. List of solutions used in this step:

Elution buffer

> 250 µl 2M Tris·HCl pH 7.5   [f. c. 10mM]
>
> 4167 µl 3M NaOAc pH 5.5   [f. c. 250mM]
>
> 200 µl 0.5M EDTA pH 8   [f. c. 2mM]
>
> Fill up to 50 ml with ddH$_2$O and autoclave.

3.6   5' end-labeling of RNA and DNA

The primers which afterwards have been used for reverse transcription were labelled with ^{32}P-phosphate at the 5' end of RNA. The reaction is catalyzed by the T4 polynucleotide kinase which transfers the γ-phosphate of ^{32}P-γ-ATP to the 5'-OH of RNA and DNA, respectively.

- 1 µl oligo [10pmol/µl]
- 2 µl PNK 10x buffer from NEB
- 10 µl ddH$_2$O
- 1 µl T4 PNK [10u/µl] from NEB
- 6 µl ^{32}P-γ-ATP [10µCi/µl]
- Incubate the sample for 45 minutes at 37° C.
- Add 1 µl 0.5M EDTA pH 8 to chelate free Mg^{2+} ions.
- Denature sample for 1 minute at 95° C.
- Immediately put sample on ice for 2 minutes.
- Precipitate for 60 minutes at – 20° C after adding:
    - 1 µl glycogen [10µg/µl]
    - 50 µl 0.3M NaOAc pH 5/EtOH
- Centrifuge with 13000 rpm at 4° C for 60 minutes
- Resuspend the dried pellet in 20 µl ddH$_2$O

The labelled oligo can be used for about 10 days due to the half-time of ^{32}P which is about two weeks.

3.7   Dephosphorylation of RNA

In order to use *in vitro* transcribed RNA for 5' - end labelling the phosphate group of the 5'- end needs to be removed. This reaction is catalyzed by the Alkaline Phosphatase. The reaction was performed as described below:

- 10 µl RNA [15pmol/µl]
- 210 µl ME buffer
- 5 µl Alkaline Phosphatase [5u/µl]
- Incubate the reaction at 37° C for 30 minutes.
- Add 30 µl 3M NaOAc pH 5 and vortex!
- Extract RNA with 250 µl phenol:chloroform:isoamylalcohol (25:24:1). Vortex strongly!
- Centrifuge at 13000 rpm for 5 minutes at 4° C
- Transfer aqueous phase into new tube.
- Perform back extraction with organic phase in order to increase yield.
- Extract the aqueous phase with 250 µl CI, vortex and repeat the centrifugation step.
- Combine both aqueous phases and precipitate at – 20° C for 60 minutes after adding
    - 2 µl glycogen [10µg/µl]
    - 1200 µl EtOH p. A.

- Centrifuge at 13000 rpm for 60 minutes at 4° C.
- Dry pellet and resuspend it in 10µl ME buffer.

The concentration of a 1:1000 dilution was measured at 260nm and calculated according to the following formula:

$$\frac{A_{260} * Dilution * 10^6}{\varepsilon} = \mu M$$

For calculating the extinction coefficient ε the following term was used:

$$\varepsilon = \left(15346,6 * n_A + 11751,8 * n_G + 7625,2 * n_G + 9929,7 * n_U\right)$$

where n represents the sum of the corresponding base in the nucleotide.

List of solutions used in this step:

ME buffer

1 ml	1M MOPS pH 6	[f. c. 10mM]
200 µl	0.5M EDTA pH 8	[f. c. 1mM]

Fill up to 100 ml with ddH$_2$O and autoclave.

CI (25:1)

50 ml	Chloroform	[25x]
2 ml	Isoamylalcohol	[1x]

Mix, wrap with aluminium foil and store in the dark.

3.8 In vitro TbCl$_3$ cleavage assay

An *in vitro* Tb^{3+}-mediated cleavage assay was performed using ^{32}P-labeled D135 RNA. Labeling was performed as described in section 3.6. A protocol for the assay is given below and is based on that published by Sigel et al., 2000:

- 1 µl *in vitro* transcribed RNA [2.5pmol/µl], yielding 2.5pMol f. c.
- 4 µl 1M MOPS pH 7 [f. c. 90mM]
- 8.35 µl 1M KCl [f. c. 190mM]
- 30.65 µl ddH$_2$O
- Incubate for 1 minute at 95° C and 2 minutes at RT.
- Fold RNA with 0.5 - 5µl 1M MgCl$_2$ yielding 10 - 100mM f.c. for 20 minutes at 42° C.
- Dissolve TbCl$_3$ in 5mM cacodylate buffer pH 5.5 for 150mM stock solution.
- Induce cleavage by adding 1 µl Tb(III) solution [f. c. 2.5 - 25mM]
- Incubate for 30 minutes at 30° C.
- Stop reaction with 1 µl 0.5M EDTA pH 8.
- Precipitate at - 20° C for 60 minutes after the addition of 200µl 0.3M NaOAc pH 5/EtOH.

- Centrifuge at 13000 rpm for 30 minutes at 4° C.
- Dry pellet and resuspend in 8µl PAGE loading buffer.

3.9 Poison primer assay

This assay is used to monitor the splicing efficiency of the group II intron ai5γ. It is based on reverse transcription and therefore requires a primer which binds to the 5' end of the downstream exon. The first guanosine residue located upstream of the primer binding-site occurs at a different distance in the ligated exons and in the unspliced pre-RNA. By replacing dCTP with ddCTP in the reverse transcription reaction, the extension of the primer is stopped precisely at first guanosine, yielding extension products of distinct length depending on the template (mRNA or pre-RNA) used for reverse transcription. When the gel is dried and exposed to a Phosphorimager screen two bands can usually be detected; **a)** primer + 8nts which corresponds to unspliced pre-RNA and **b)** primer + 14nts which represents the ligated exons. By performing a quantitative analysis of both bands the percentage for *in vivo* splicing activity can be calculated. Reverse transcription is performed as described in section 3.1 with one exception for the dNTP mixture. The following dNTPs/ddNTP mixture was used:

Poison mix

3 µl	10mM dATP	[f. c. 0.88mM]
3 µl	10mM dGTP	[f. c. 0.88mM]
3 µl	10mM dTTP	[f. c. 0.88mM]
25 µl	5mM ddCTP	[f. c. 3.67mM]

3.10 Analysis

In order to perform a quantitative analysis of the phosphoimager scans ImageQuant 5.2 ™ from Molecular Dynamics (GE Healthcare) was used.

Metal ion-induced cleavage assay

Line objects were applied to the control lane (showing natural RT stops) as well as to the sample lane. Both line objects had the same length, covered approximately 100 nucleotides and were then transformed into graphs which were then exported for further analyses with Microsoft Excel. The position of cleavage sites was determined with the help of two sequencing lanes (A and C lanes) allowing assigning the cleaved nucleotide. Cleavage was defined as a peak with an at least 1.5x increased intensity compared to corresponding band in the control lane. For normalization several bands along the gel were chosen

Poison primer assay

Rectangular objects of the same size were applied to the bands corresponding to the mRNA and pre-RNA. Additionally, for the subtraction of the background rectangular objects of identical size were applied to regions above the cDNA bands stemming from the mRNA or pre-RNA. Volume reports were created with ImageQuant 5.2 ™ and analyzed in Microsoft Excel. Background was subtracted from both species and the percentage of each species was calculated.

4 Materials

4.1 Yeast strains

This section lists all the yeast strains which were used for the experiments.

Strain #16

This strain has wild-type background (MATa ade1 lys1ura3-52) and was generously provided by Phil Perlman.

Strain #189

This strain was created using strain #16 and gene knocking out of *mss116* (MATa ade1 lys1 ura3-52 mss116Δ::kan[R]). We also obtained the strain from Phil Perlman.

4.2 Oligonucleotids for mapping ai5γ

Poison Primer Oligo

ai5g −6: 24mer Tm = 64°C

5' GAAAATAGCACCCATTGATAATAC 3'

RT Oligos

Sc 82: 22mer Tm = 56˚C

5' TGTTACCATTTTAATACACTTG 3'

Sc 160: 23mer Tm = 58˚C

5' GTTCCTTCATCTTTTTTTTATAA 3

Sc 234: 21mer Tm = 56˚C

5' GTTTTCAATTAGTGGTGTAAG 3'

Sc 312: 21mer Tm = 56˚C

5' TTCCAATACATAACATCAACC 3'

Sc 312_2: 25mer Tm = 68˚C

5' CTCATTCCAATACATAACATCAACC 3'

Sc 375: 22mer Tm = 58˚C

5' GTAAATATCTAACTTAGCTCTC 3'

Sc 440: 28mer Tm = 60˚C

5' TTGGTATTATTATTATTTTTTATTATTA 3'

Sc 520: 27mer Tm = 60˚C

5' TATTTCTTAATCCAATAATTATTTATA 3'

Sc 588: 21mer Tm = 58˚C

5' GCATTAGCTTTTTATACAATC 3'

Sc 656: 22mer Tm = 58˚C

5' CACCTATAGTATAAGTTAGCAG 3'

Sc 658: 26mer Tm = 70˚C

5' GCATATCACCTATAGTATAAGTTAGC 3'

Sc 686: 25mer Tm = 62˚C

5' ATATATATAAATAAAGATAGGCACC 3'

Sc 754: 25mer Tm = 62˚C

5' ATTTATCAGTTATATATAAACCTCC 3'

Sc 833: Tm 58˚C 18mer

5' AGGAACCGTACGTGCGAC 3'

Sc 854: 21mer Tm = 62°C

5' GACCTTTACAAGTTTTCCCCC 3'

Sc 866: 20mer Tm = 60°C

5' CCCGATAGGTAGACCTTTAC 3'

4.3 Enzymes

Name	Activity	Company
AMV Reverse transcriptase	[10u/µl]	Promega
T4 Polynucleotide kinase	[10u/µl]	NEB
Alkaline phosphatase	[5u/µl]	NEB

4.4 Kits

RiboPure^{TM}-Yeast Kit from Ambion to isolate total RNA from yeast.

4.5 Table listing the extinction coefficient ε for all RNA or DNA sequences

Name	Length	ε [L·mol−1·cm−1]
ai5g -6	24mer	291832.5
Sc 82	22mer	245373
Sc 160	23mer	242646.8
Sc 234	21mer	244234.3
Sc 312	21mer	241130.3
Sc 312_2	25mer	281653
Sc 375	22mer	248483.4
Sc 440	28mer	319594.1
Sc 520	27mer	315351.4
Sc 588	21mer	235443.3
Sc 656	22mer	259851
Sc 658	26mer	302680.2
Sc 686	25mer	317209
Sc 754	25mer	287284.2
Sc 833	18mer	205219.2
Sc 854	21mer	215383.5
Sc 866	20mer	219127.9
D135	618 nucleotides	7442189

5 Results

In order to map metal ion binding sites *in vivo* I employed a metal-induced cleavage assay, a powerful technique, which has been widely used *in vitro* (Lindell et al., 2002). In order to accomplish my goal, I adapted a previously published assay (Lindell et al., 2002), but I also established a new assay. The main difference in these two assays was the type of metal ion used as a probe: Pb^{2+} vs. Tb^{3+}. These two metal ions differ in their ionic radius and therefore in their efficiency to displace a magnesium ion from its binding pocket and in turn to induce site-specific cleavage of RNA residues in the vicinity of the ion. In brief yeast cells were grown to log phase, harvested by centrifugation and incubated with lead(II)acetate or terbium(III)chloride, respectively. The reaction was stopped by addition of a molar excess of EDTA and total RNA was extracted. 5'-^{32}P labelled oligos were used to map the cleavage sites in the RNA by reverse transcription. The pool of cDNAs was then resolved on a denaturing acrylamide gel which was dried and exposed in a phosphoimager screen. As a control, RNA extracted from un-treated cells was also reverse transcribed to reveal any natural stops of the Reverse Transcriptase along the RNA template.

Due to the fact that metal ion binding pockets may vary between a catalytically active RNA and an unspliced RNA I had to make sure that the metal probes did not affect the *in vivo* splicing activity of the group II intron ai5γ. Thus, a poison primer assay, which allows monitoring splicing rates *in vivo*, was performed using RNA extracted from untreated and Pb^{2+} or Tb^{3+} treated yeast cells.

5.1 Terbium(III)-induced cleavage assay

In order to investigate metal ion binding sites without perturbing the pocket I decided to use ter-bium(III), as its ionic radius is close to that of magnesium (0.92 Å and 0.72 Å), which is known to be the most important divalent metal ion *in vivo*. This small ionic radius allows terbium to displace magnesium ions in specific metal ion binding pockets, which contribute to the structural organization of the ai5γ intron RNA. Upon displacement, Tb^{3+} induces cleavage of the RNA backbone in its vicinity at neutral pH. By using terbium rather than lead I expected a stronger cleavage pattern and therefore a better signal-to-noise ratio.

First of all it is important to mention that all optimization steps and changes in the protocol were first tested with oligo Sc 160. The reason to use this specific oligo is the presence as well as the abundance of metal ion binding sites *in vitro* in the area mapped with this primer *(Sigel et al., 2000)*. Though there is no mandatory overlap between *in vivo* and *in vitro* metal ion binding pockets chances are high that some of them might be overlapping and therefore present a unique chance for optimization of the as-say. For setting up the assay I started with concentration series in order to narrow down the concentra-tion of Tb^{3+} needed for cleavage (Figure 11 panel A). Several runs with RNA obtained from different RNA preparations did not result in any cleavage pattern. No detectable difference in the band pattern be-tween the control lane and the different sample lanes could be observed. The only bands which are of strong intensity in the sample lanes (e.g. A114 in figure 11) are also present in the sequencing- as well

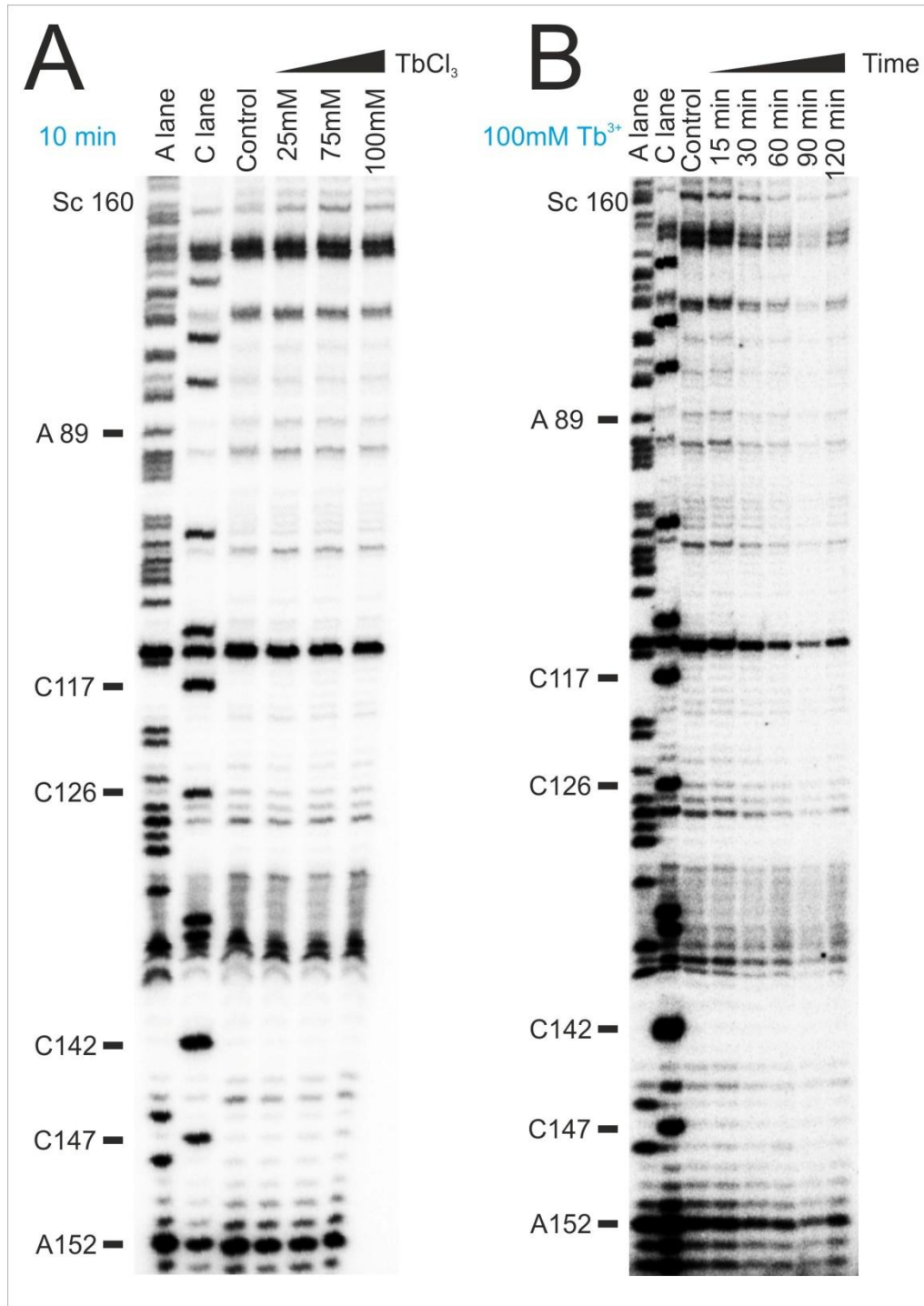as in the control lane, thus representing natural stops of the Reverse Transcriptase which are independent of Tb^{3+}.

Figure 11: Varying the concentration of and incubation time with $TbCl_3$ during the metal-induced cleavage experiment. Representative gel showing the c-c1-c2 three-way junction of the ai5γ group II intron. (**A**) Yeast cells were incubated with increasing [$TbCl_3$] for 10min at 30°C. (**B**) Yeast cells were incubated with 100mM $TbCl_3$ for increasing periods ranging from 15 to 120 minutes. Lanes A and C denote the sequencing lanes. In the control lane, RNA which was extracted from untreated cells was reverse transcribed to detect natural, Tb^{3+}-independent stops of the Reverse Transcriptase. In lanes 25mM to 75mM in (A) or lanes 15min to 120min in (B) RNA extracted from cells treated with Tb^{3+} was reverse transcribed to detect any Tb3+-induced cleavage sites.

For this set of experiments I would have expected a cleavage pattern of certain bands increasing in strength with higher Tb^{3+} concentration. Such pattern had been observed previously when using terbium as a probe *in vitro* (Sigel et al., 2000). Although *in vivo* terbium-induced cleavage is likely to need much higher concentrations when compared to *in vitro* work (this was observed when working with lead(II)acetate but this might be true as well for terbium chloride (Lindell et al., 2005)), it is strange that no cleavage pattern at all was observed during several repeats of the experiments. One potential explanation for this lack of Tb^{3+}-induced cleavage might an incubation time too short to induce cleavage. The next step was therefore to investigate and extend the time period during which yeast cells were incubated with terbium.

Elongation of the incubation time did not lead to a detectable cleavage pattern as well (figure 11B). Due to the longer incubation time I expected increasing cleavage intensity over time but this was obviously not the case. Quantification of each lane and normalization in order to counteract loading differences resulted in almost identical values. A possible explanation for the negative results is the usage of a terbium chloride reagent different than that described in the literature. *In vitro* terbium-induced cleavage was performed using $TbCl_3 \cdot 6H_2O$, while I was using anhydrous $TbCl_3$. In order to rule out potential differences between the two compounds I decided to perform an *in vitro* cleavage assay with both chemicals and compare the results. I used an *in vitro* transcribed the ai5γ ribozyme construct D135, in which the exons and domain 6 were deleted and domains 2 and 4 were shortened to small UUCG-capped hairpins (Swisher et al., 2001). This *in vitro* assay was carried out with much lower Tb^{3+} concentration [2.5mM] than usually used for the *in vivo* assays. Figure 12 shows a direct comparison of cDNA transcribed from RNA treated with a solution of either $TbCl_3 \cdot 6H_2O$ or anhydrous $TbCl_3$. For this experiment the RNA was folded in the presence of $MgCl_2$ and then treated with 2.5mM terbium for 90 minutes. As can be easily judged from the gel, there is no detectable difference between the control and the sample lanes independent of the type of $TbCl_3$. The experiment was performed two times in order to rule out difficulties which might not have been recognized if the experiment would have been performed only once. The small differences between the lanes are due to loading differences and not due to cleavage. Thus it appears that the problems with the *in vivo* terbium-induced cleavage assay are not necessarily caused by the different terbium stock, which we initially used.
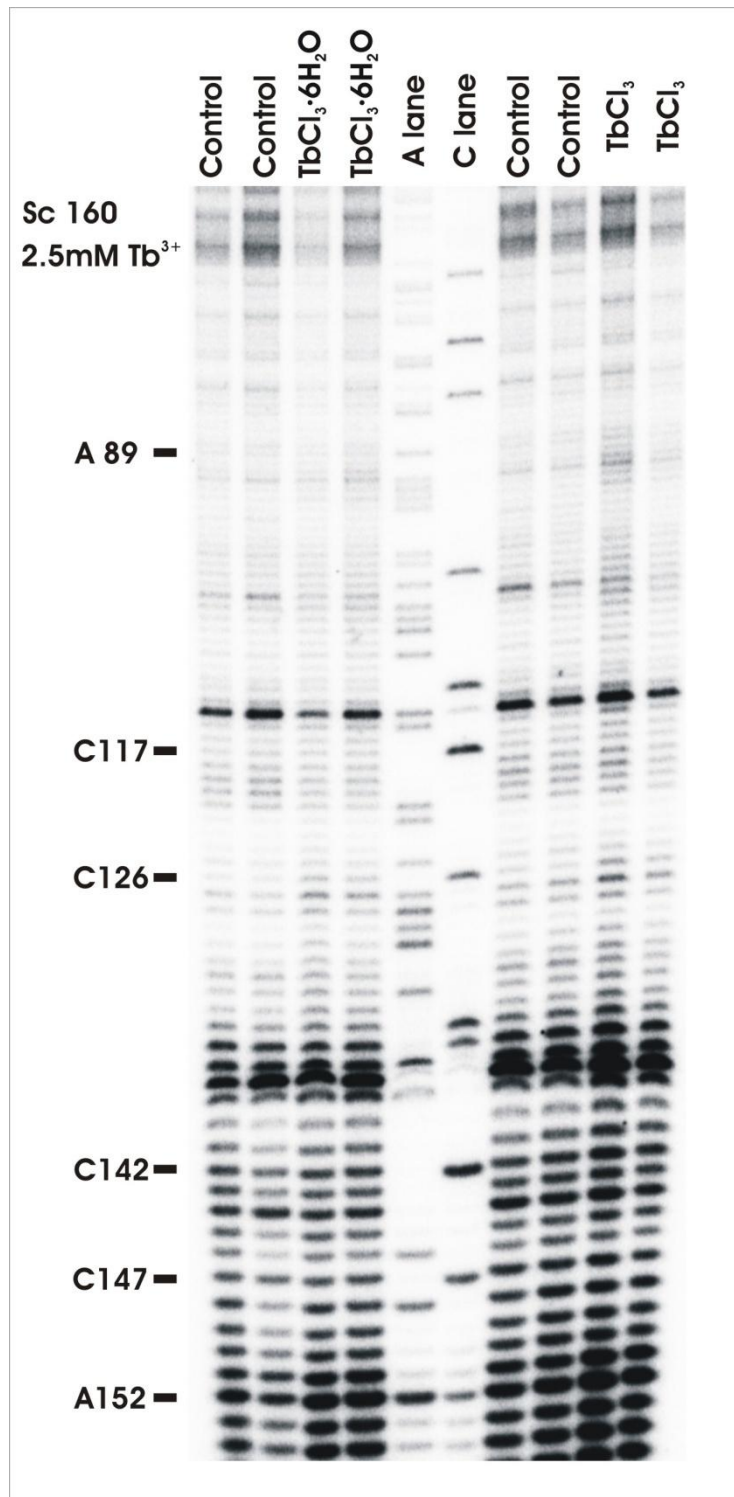
Figure 12: Comparison of TbCl$_3$·6H$_2$O and anhydrous TbCl$_3$ as metal probes for metal ion-induced cleavage assay. Representative gel showing the c-c1-c2 three-way junction of the ai5γ group II intron. Lanes A and C denote the sequencing lanes. In the control lanes, RNA which was extracted from untreated cells was reverse transcribed to detect natural, Tb^{3+}-independent stops of the Reverse Transcriptase. In lanes TbCl$_3$·6H$_2$O (left side) or TbCl$_3$ (right side) RNA extracted from cells treated with TbCl$_3$·6H$_2$O or TbCl$_3$ was reverse transcribed in order to detect stops of the Reverse Transcriptase due to Tb^{3+}-dependent cleavage.

.

In order to determine if the problem is related to the fact that I map potential cleavage sites with the reverse transcription, I decided to perform an *in vitro* cleavage assay using 5'endlabelled D135 RNA. After folding the radioactively labelled RNA was incubated with either 2.5mM or 25mM $TbCl_3$. The reaction was quenched and the samples were immediately loaded on a standard denaturing PAGE. In the first set of these experiments I folded D135 at varying Mg^{2+} concentration to yield unfolded state (0mM Mg^{2+}), the compact intermediate state (10mM Mg^{2+}) and the native state (100mM Mg^{2+}; figure 13, compare also to figure 8 for folding pathway). In the unfolded state, Tb^{3+} cleaves the RNA at almost every position as it displaces delocalized ions, but the cleavage pattern becomes more and more defined with increasing Mg^{2+} ions, as specific metal ion binding pockets are formed (figure 13). Most importantly, the Tb^{3+}-induced cleavage pattern is in very good agreement with the previously published ones (Sigel et al., 2000; Waldsich and Pyle, 2008). Secondly, the signal intensity is much stronger compared to figure 11, although the experimental set up was identical except for how the cleavage sites were detected: either directly via 5'-endlabel of the RNA (figure 13) or indirectly via reverse transcription (figure 11). Although I did not load a sequencing lane the gel in figure 13 represents the same region as all other gels shown until now (nucleotides 155 to 50, domain 1). This result was quite satisfying, as unlike with reverse transcription a clear and distinct difference between the control and the sample lanes is visible (figure 13). It is worth noting though that the background in the control lane in figure 13 is much smaller than that observed for reverse transcription (e.g. figure 11). As expected some cleavage positions are decreasing with increasing Mg^{2+} concentrations and therefore make it quite clear that the cleavage pattern observed is due to terbium. Another interesting finding is the fact that there is only minor difference in the cleavage pattern between 2.5mM and 25mM terbium. This may indicate that the *in vitro* assay might be already saturated at 2.5mM terbium.

Having now proven that terbium-induced cleavage is working *per se* - at least *in vitro* for 5'-labeled RNAs I decided to investigate if the cleavage pattern is still detectable when using reverse transcription rather than 5'-labeling (figure 14).
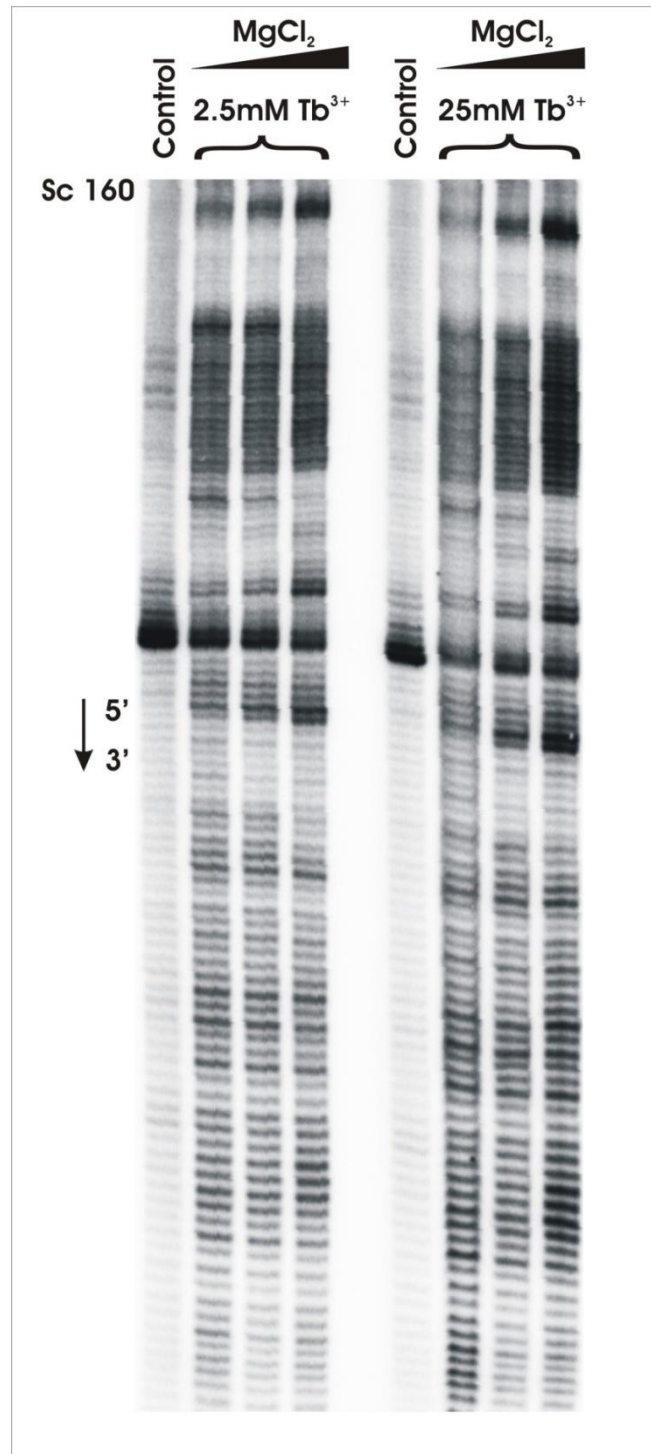
Figure 13. *In vitro* Tb^{3+}-induced cleavage assay of 5'-end labelled D135 construct with different Tb^{3+} concentrations. Representative gel showing the c-c1-c2 three-way junction of the ai5γ group II intron. The control lanes show folded but untreated RNA. The sample lanes show RNA treated with 2.5mM $TbCl_3$ (left side) or 25mM $TbCl_3$ (right side) and 0mM (unfolded); 10mM (compact intermediate) or 100mM (native state) Mg^{2+}. With increasing Mg^{2+} concentration the cleavage pattern becomes more defined.

Therefore, D135 RNA was again radioactively labelled at its 5' end, folded under optimal splicing conditions (Qin and Pyle, 1998) and subsequently incubated with terbium *in vitro*. Again two different concentrations were used: 2.5mM and 25mM, respectively. This time however, I decided to map the cleavage site in two different ways: one half of the samples was loaded directly on a 8% denaturing polyacrylamide gel to resolve the cleavage products, while the second half of the samples was used to perform reverse transcription with radioactively labelled oligos to map the cleavage sites. If the lack of Tb^{3+}-induced cleavage was indeed caused by the reverse transcription reaction a clear difference should be seen between the two sample halves. Fortunately, figure 14 shows the exact opposite because in both cases cleavage was observed. I should however point out that I took great care to obtain an extremely clean RT reaction (e.g. cleaning pipettes, new aliquot of required solutions, etc.). The results are very promising because the cleaved positions are exactly the same for 5'-labeled RNA as well as for reverse transcribed cDNA – as expected. This is shown in figure 14 by coloured ellipses which are appearing in reversed order on panel B than in panel A due to the different orientation of the cDNA based on the method. Although the cleavage sites match, it also became evident that the different intensity of the bands reflecting the same cleavages is very different. In case of the end-labelled RNA the band intensity is very strong with a pronounced difference between the sample and control lane (figure 14A). Unfortunately, this is not the case for mapping the cleavage site via reverse transcription (figure 14B); only few bands of the sample lane show a large difference in intensity compared to the control lane. I therefore concluded that mapping metal ion binding pockets via reverse transcription decreases the sensitivity of detection. Considering this fact it is obvious that particular attention has to be paid to working in an absolutely clean manner and to handling the RNA samples with optimal care. It is also worth mentioning that both sample lanes containing 25mM of Tb^{3+} were completely degraded. This could be explained by too low concentrations of EDTA before heating the sample during reverse transcription.
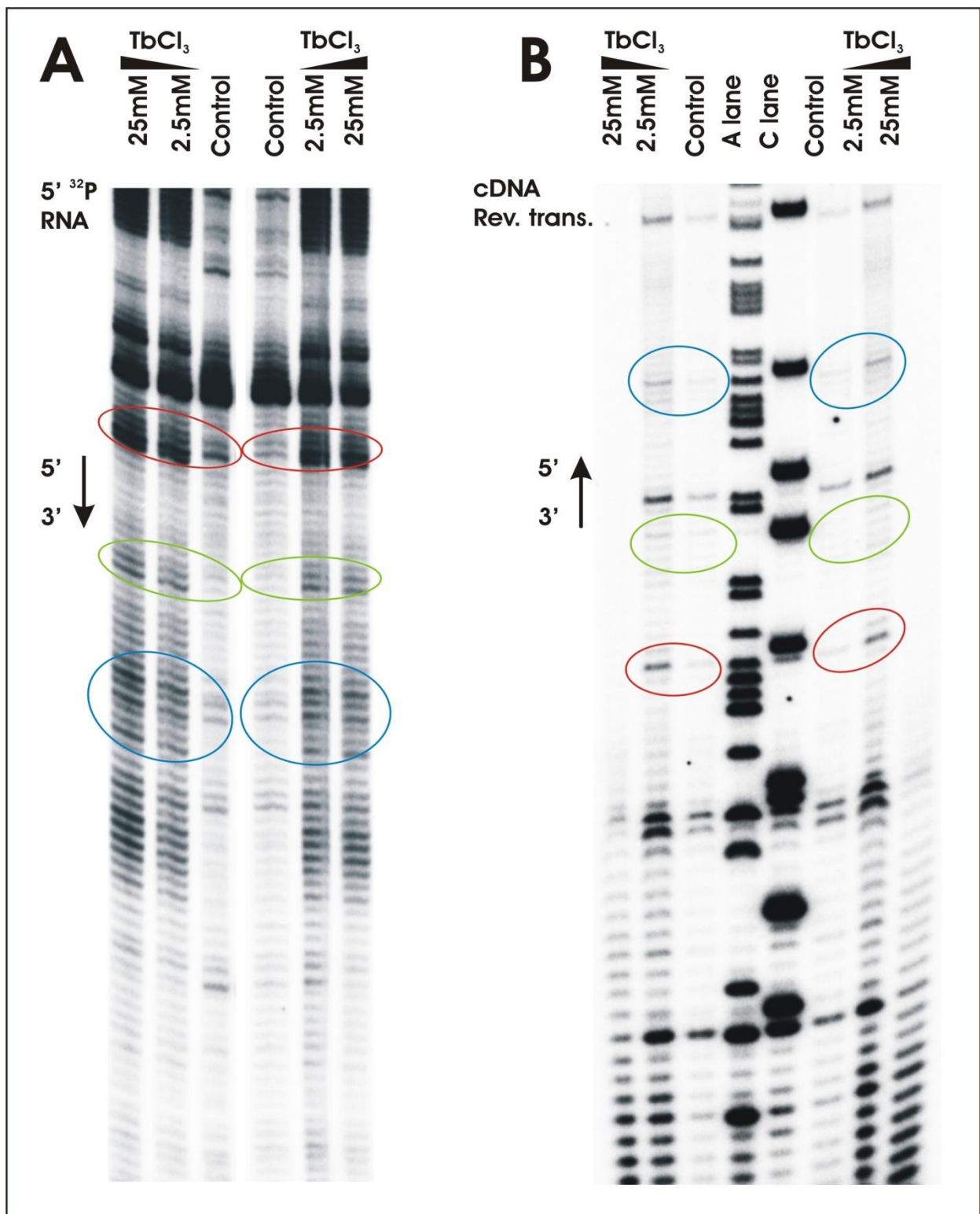
Figure 14. Comparison of a Tb[3+]-induced cleavage assay with either 5'-labeled RNA or reverse transcribed RNA. Representative gel showing Representative gel showing the c-c1-c2 three-way junction of the ai5γ group II intron. **(A)** Control lanes show untreated RNA which was folded. Sample lanes show RNA treated with 2.5mM or 25mM TbCl₃, respectively. Coloured ellipses indicate positions where Tb[3+]-dependent cleavage has taken place. **(B)** A and C lanes denote the sequencing lanes, the control lanes show untreated and reverse transcribed RNA in order to detect Tb[3+]-independent stops of the Reverse Transcriptase. Sample lanes show RNA treated with 2.5mM or 25mM TbCl₃ which was reverse transcribed to detect any Tb[3+]-dependent cleavage positions. Such positions are marked by coloured ellipses.

The next logical step in developing the novel *in vivo* technique was now to address the question why mapping of metal ion binding pockets *in vivo* via reverse transcription is not working. Obviously it must have something to do with the reverse transcription reaction because I have shown previously that *in vitro* mapping using RT is working without problems. Another aspect that differs between *in vitro* and *in vivo* mapping is the buffer used for incubation Tb^{3+} with yeast or with RNA. In case of *in vivo* assay we incubated the yeast cells with Tb^{3+} in prewarmed YPD medium, while the cleavage reaction is carried out *in vitro* in the presence of MOPS pH7 (Waldsich and Pyle, 2008). Therefore I decided to switch to MOPS buffer which was used successfully for the *in vitro* control experiments. In brief, cells were grown to log phase, harvested by centrifugation and resuspended in different buffers followed by incubation for 10 minutes with 150mM terbium. The high concentration was chosen in order to make sure that strong cleavage will be observed. The different buffers that were tested were the following: **a)** 80mM MOPS pH 7; **b)** 80mM MOPS pH 7, 10mM $MgCl_2$; **c)** 80mM MOPS pH 7, 250mM NaCl and **d)** 80mM MOPS pH 7, 1% Glucose. The results of this experiment are shown in figure 15. It is quite obvious that the choice of right buffer has a profound impact on the obtained results. When using MOPS buffer pH 7 alone without any additives a pronounced cleavage pattern is observed. This is not the case with the other buffer systems. Additional metal ions (Mg^{2+}, Na^+) in the buffer might increase the chances that Tb^{3+} ions are not able to displace the bound Mg^{2+} in the RNA or these ions interfere with the uptake of Tb^{3+} into the cell. Glucose on the other hand might quench the reaction or scavenge the Tb^{3+} ions. Until now all the *in vivo* experiments have been performed with incubation in YPD-media and quenching of glucose would explain why no cleavage pattern had been observed. In addition it must be mentioned that the control lane and the sample lane for 80mM MOPS pH7 are not equally loaded but nevertheless cleavage has taken place (normalized quantitative analyses was carried out).
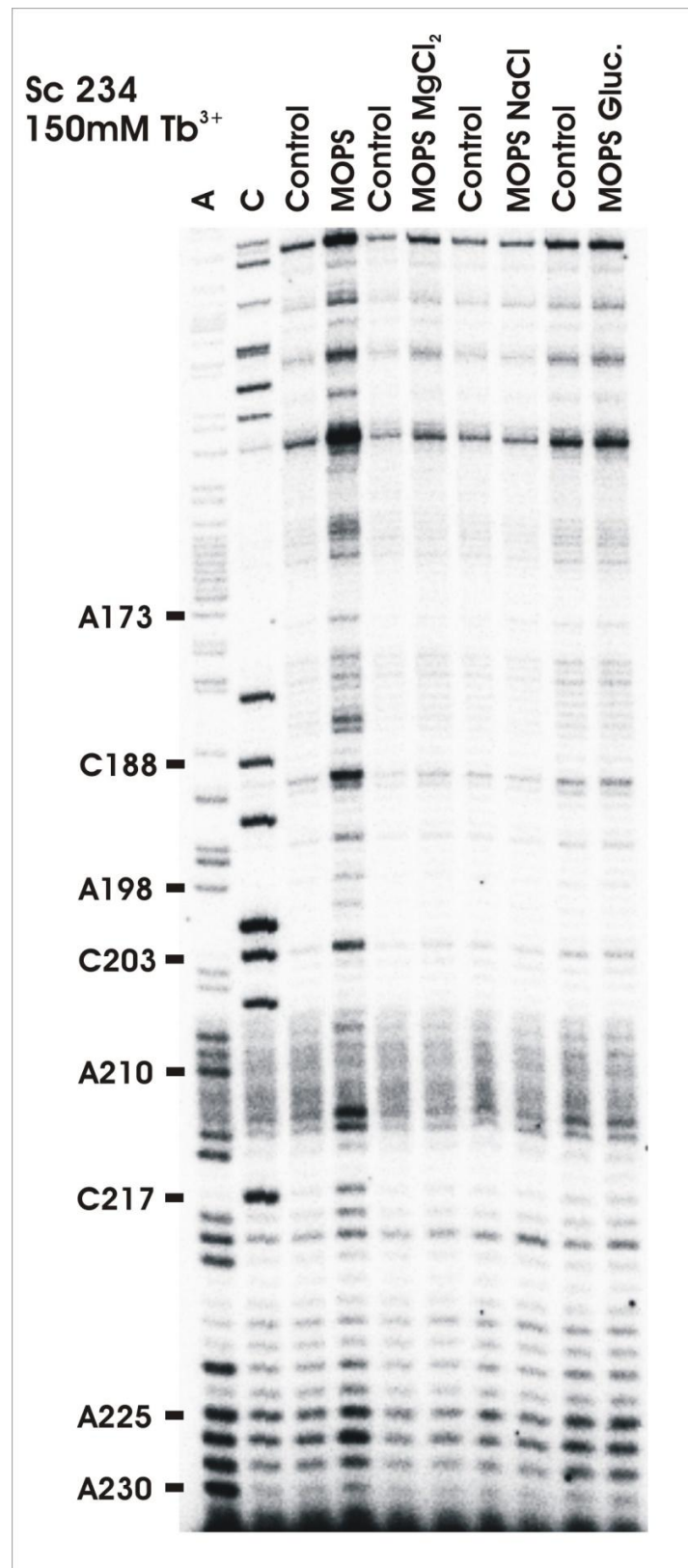
Figure 15. Influence of different buffer systems on the Tb^{3+}-induced cleavage assay. Representative gel showing the κ – ζ element and c-d region of the five-way junction of the ai5γ group II intron. A and C lanes denote the sequencing lanes. For the control lanes RNA extracted from untreated cells was reverse transcribed in order to detect natural, Tb^{3+}-independent stops of the Reverse Transcriptase. The other lanes correspond to the buffer system used *in vivo*. RNA extracted from cells treated with 150mM $TbCl_3$ was reverse transcribed to detect any Tb^{3+}-dependent cleavage positions. The buffer systems used were: **a)** 80mM MOPS pH 7; **b)** 80mM MOPS pH 7, 10mM $MgCl_2$; **c)** 80mM MOPS pH 7, 250mM NaCl and **d)** 80mM MOPS pH 7, 1% Glucose.

Now knowing which buffer to use for performing metal ion-induced cleavage assays *in vivo*, I performed several independent experiments to map Tb^{3+} cleavage sites throughout the entire ai5γ intron (requiring reverser transcription reactions with 13 different oligos). The final goal was to obtain three independent data sets which can be used to do quantification and final analysis.

After completing three data sets I discovered that the reproducibility of terbium-induced cleavage was not as high as expected and required for a reliable analysis. This fact is best seen in figure 16 which compares the same region of the group II intron ai5γ, the so-called κ–ζ element. The three gels were loaded with cDNA obtained from terbium-treated RNA, which was isolated independently. Despite showing exactly the same region ranging from nucleotide 223 to approximately 140 the cleavage pattern do not correlate very well. Only the gel on the left side shows significant cleavage whose positions are marked by light blue arrows. In order to make a reliable conclusion it would be necessary to observe the same cleavage pattern also for the gels shown in the middle as well as on the right panel of figure 16.

Overall the impossibility to get good reproducibility strengthened my decision to employ the previously published lead-induced cleavage assay. Nevertheless, I am convinced that the Tb^{3+}-induced cleavage assay *in vivo* might work out at some point, requiring a lot more of optimization. Unfortunately, I did not have time to do so.
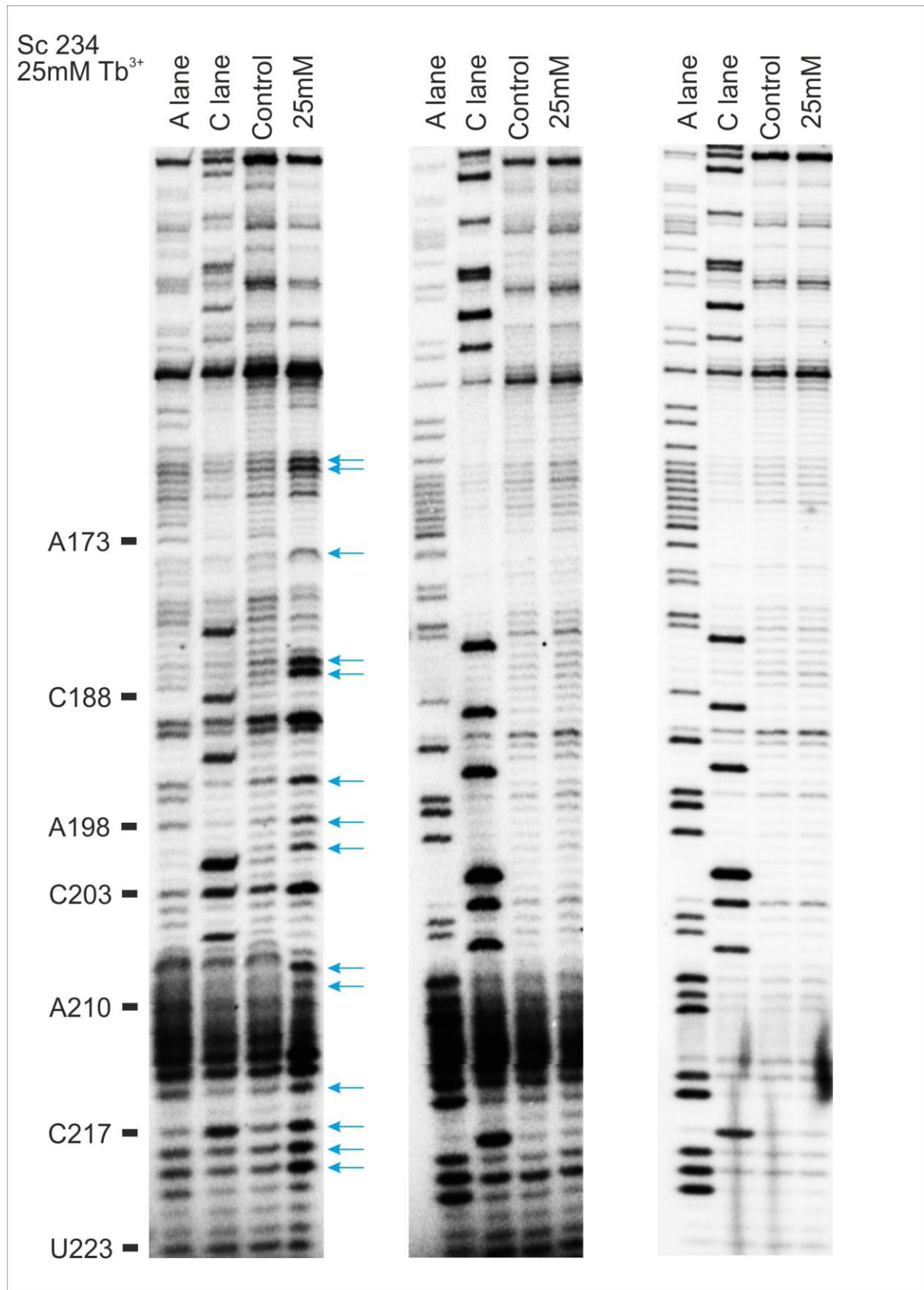
Figure 16. Comparison of three representative gels showing the κ – ζ element and stems c and d of the five-way junction of the ai5γ group II intron. Sequencing lanes (A lane and C lane, respectively) as well as the control lane show reverse transcribed RNA extracted from untreated cells and. For the sample lanes RNA extracted from cells treated with 25mM TbCl$_3$ was reverse transcribed to detect any Tb^{3+}-dependent cleavage positions. These positions are marked with light blue arrows.

5.2   Lead(II)-induced cleavage assay

With the discovery that the reproducibility of the results obtained by terbium(III)-induced cleavage assays is insufficient I decided to go back to lead(II)-induced cleavage assays. The assay itself is based on that published by Lindell et al. in 2002, but some modifications were of course required to meet the different situation. One aspect which turned out to be of critical importance is the *in vivo* splicing activity of ai5γ. A poison primer assay was used to monitor splicing of the ai5γ group II intron in the cell, which has previously been reported to be highly efficient yielding approximately 80% of ligated exons (figure 17).
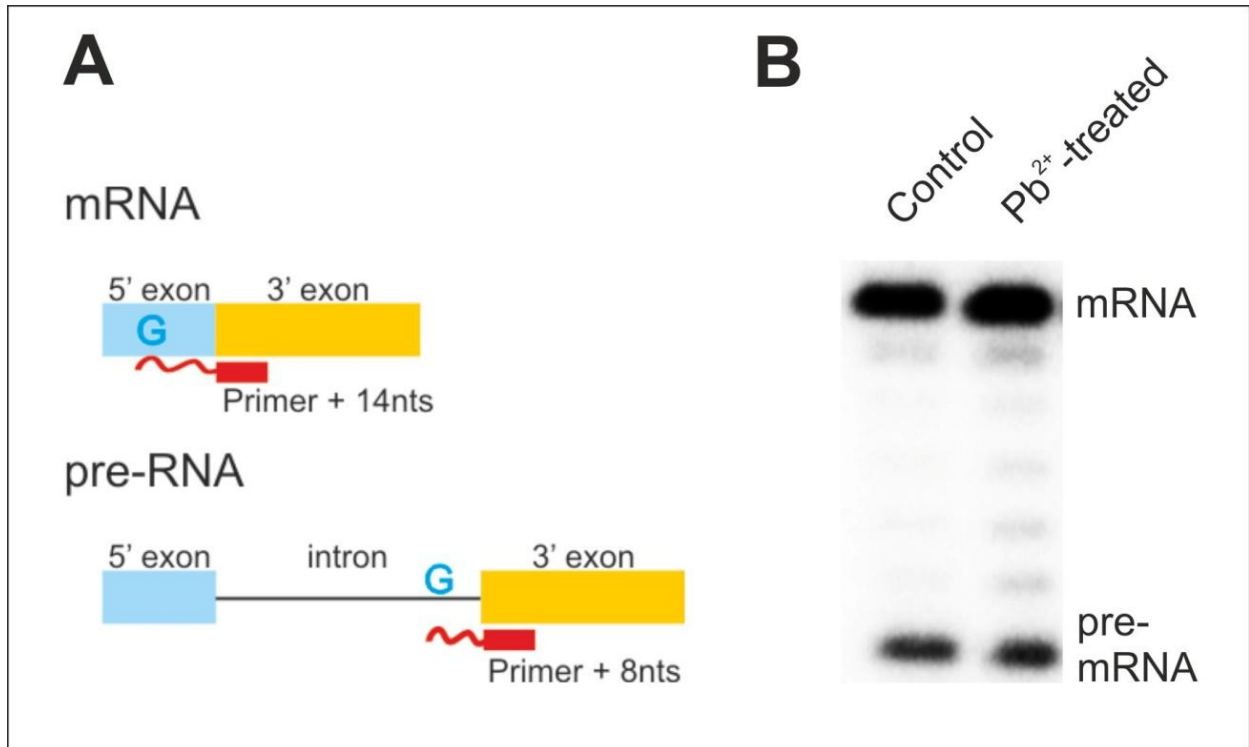
Figure 17. Representative gel showing a poison primer assay for monitoring splicing activity of the group II intron ai5γ *in vivo*. (A) In this assay a radiolabelled primer hybridizing to the 5' end of the downstream exon is extended with a poisoned dNTP mix containing one nucleotide as a ddNTP. The incorporation of the ddNTP results in termination of the extension by the reverse transcriptase. The extension signal differs with respect to the template of the extension, as the first residue complementary to the ddNTP is at a different position in the pre-RNA or ligated exon RNA, thereby resulting in distinct extension lengths of the primer. The ratio of the different extension signals reflects the *in vivo* splicing activity. In case of ai5γ ddCTP is used to detect the first G in pre-RNA (primer + 5nts) or in the ligated exons (primer + 14 nts). (B) Representative gel of a Poison Primer Reaction. Quantification reveals that 75% of the control RNA are spliced (upper band on the left) versus 25% unspliced (lower band on the left) in the wt yeast strain. When treated with lead(II)acetate (25mM) the ratio is shifted to 83% unspliced (upper band on the right) versus 17% unspliced (lower, right band). In both cases about ¾ of all ai5γ RNA is spliced and therefore available in sufficient quantities to perform lead(II)-induced cleavage assays. Notably, higher lead(II)acetate concentrations did also not affect the splicing efficiency.

With the RNA extracted from untreated and Pb^{2+}-treated cells a reverse transcription was performed with different DNA primers to map potential cleavage site throughout the entire intron. On the next pages I will show a representative gel for each primer and the corresponding area with its secondary structure elements. Cleavage sites are marked with light blue colour. The cleavage sites shown are summarized from at least 3 independent experiments.
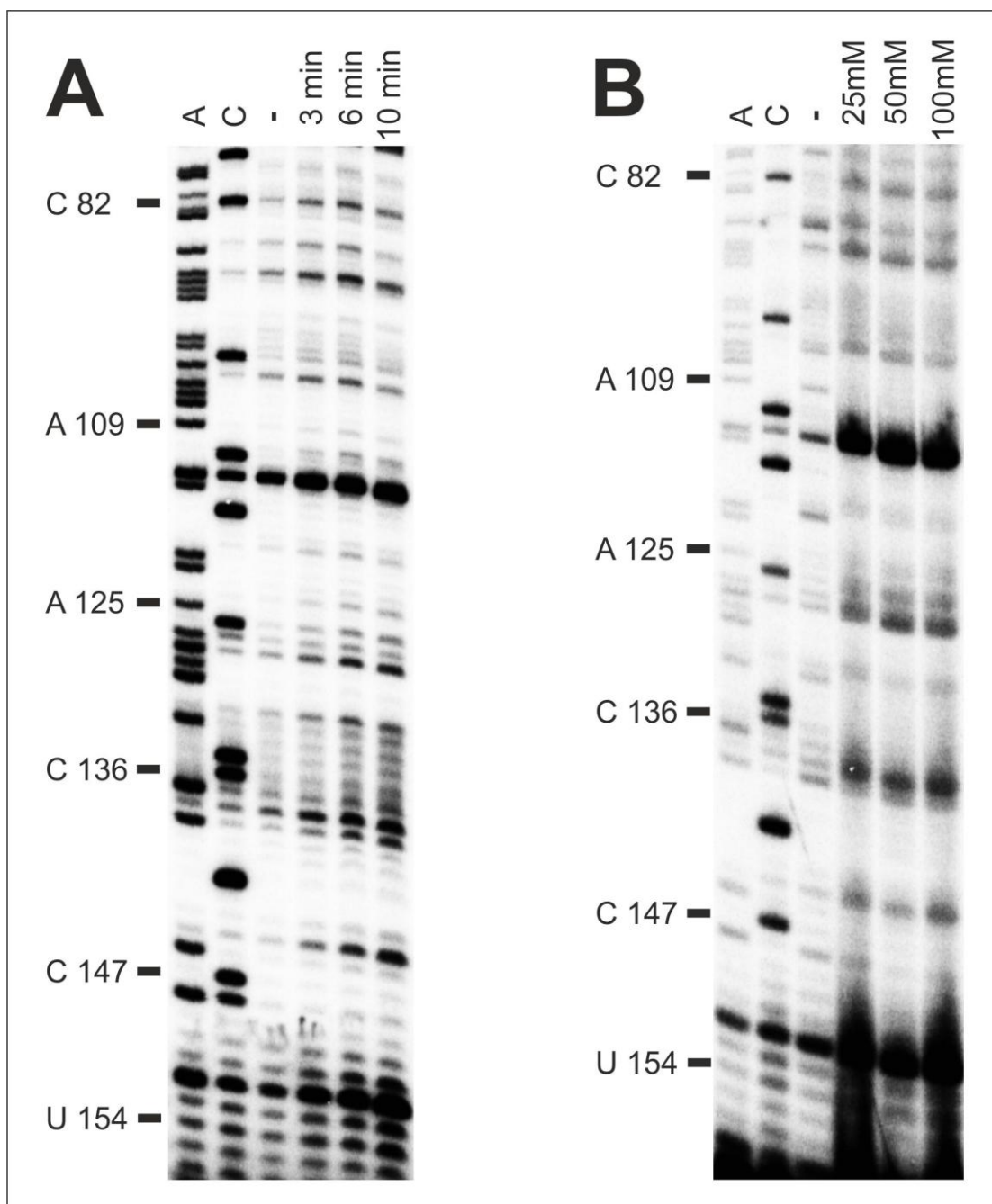

Figure 18. Varying the concentration of and incubation time with lead(II)acetate during the metal-induced cleavage experiment. Representative gel showing the c-c1-c2 three-way junction of the ai5γ group II intron. (**A**) Yeast cells were incubated with 50mM Pb(OAc)$_2$ for increasing periods ranging from 3 to 10 minutes. (**B**) Yeast cells were incubated with increasing [Pb(OAc)$_2$] for 10min at 30°C. Lanes A and C denote the sequencing lanes. In the control lane, RNA which was extracted from untreated cells was reverse transcribed to detect natural, Pb^{2+}-independent stops of the Reverse Transcriptase. In lanes 3 min to 10 min in (A) or lanes 25mM to 100mM in (B) RNA extracted from cells treated with Pb^{2+} was reverse transcribed to detect any Pb^{2+}-induced cleavage sites.

At the beginning several experiments were performed in order to find the optimal concentration of lead(II)acetate as well as incubation time (figure 18). I found out that a concentration of 25mM Pb(OAc)$_2$ is sufficient to observe a good cleavage pattern (figure 18) and that higher concentrations of lead(II)acetate did not result in a stronger cleavage pattern. Incubation time did show to have an influence on the intensity of the cleavage pattern as the bands became more intense with time (figure 18). After 10 minutes the cleavage pattern was really strong and also showed cleavage sites which were only faint when incubating for 10 minutes only. Based on these findings I choose the following conditions to perform the experiments: i) 25mM lead(II)acetate and ii) 10 minutes of incubation at 30° C. Note that these values are only valid for RNA which had been spliced efficiently in the cell. When I encountered a severe decrease in ai5γ splicing efficiency the concentration of lead(II)acetate was increased to 125mM in order to detect Pb^{2+}-dependent cleavage sites (see page 69 for further discussion).

Domain 1 – residues close to the 5'SS including stems c and i

The region detectable with primer Sc 82 starts by nucleotide A78 and goes down to C8. It is part of domain 1, which is the largest domain of the intron. At position U60 to C66 the long-range intra-domain interaction α-α' takes place which has its target at position G343 to A348 in the d3 stem of domain 1. It is considered to play a role for both binding of the 5'-exonic substrate (Qin and Pyle, 1998) as well as intron folding (Waldsich and Pyle, 2007). Figure 19 shows four sites where Pb^{2+}-induced cleavage has taken place in the majority of independent experiments; at nucleotides A71, A72 and G73 as well as A77. All four cleavage sites are spatially close and located at or near the five-way junction: A71 is part of stem b, which harbours the α-α' interaction, and the other three, A72, G73 and A77 are part of stem c and Jc/c1, which then divides itself further into stem c1 and c2. This three-way junction is important for the orientation of the active-site constituents ε' and λ in stem c1. However, none of the four bases are directly involved in either tertiary interactions or catalysis of the ribozyme. Interestingly, these cleavage sites were not observed by Sigel et al (2000), when they mapped the metal ion binding sites within the ai5γ intron *in vitro* (compare figures 12 and 19). On the other hand I did not observe cleavage in the hairpin around nucleotides A38 to A44 located in stem a where Sigel et al. did observe weak cleavage. At this point it should be mentioned that using primer extension the distinction between weak cleavage sites and strong ones is far more demanding than with 5'-end labelled RNA due to the higher background of natural RT stops. Therefore I decided to not distinguish between different cleavage intensities.
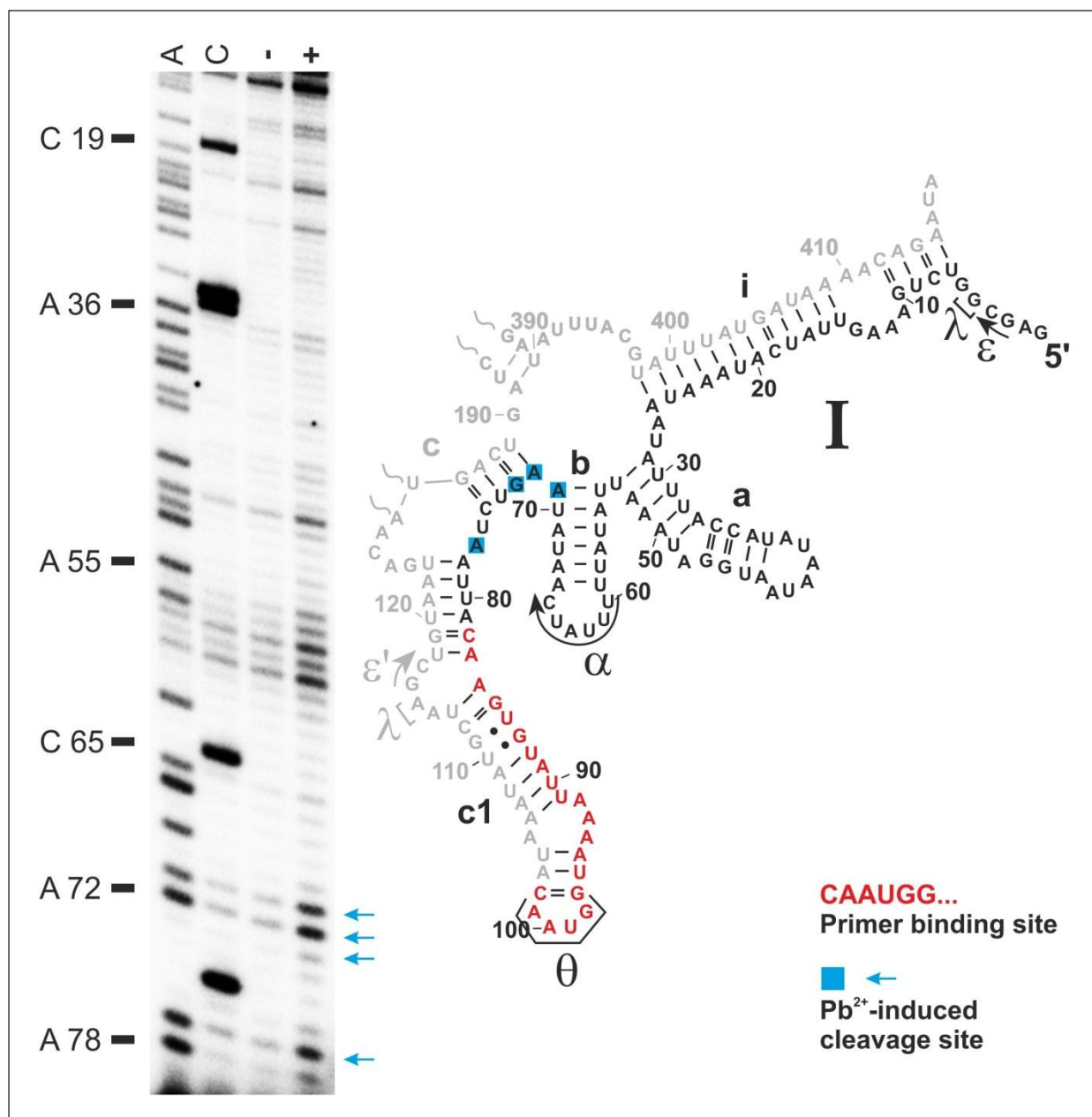
Figure 19. Overview of Pb^{2+}-induced cleavage in domain 1 of ai5γ - residues close to the 5'SS including stems c and i - mapped with primer Sc 82. A representative gel is shown on the left panel. A and C lanes denote sequencing lanes, while the lane labelled "-" contains RNA extracted from untreated cells which was reverse transcribed in order to see natural, Pb^{2+}-independent stops of the Reverse Transcriptase. Lane "+" shows reverse transcribed RNA which was extracted from cells treated with 125mM lead(II)acetate. Arrows in light blue show the cleavage sites on the gel. On the right panel the corresponding section of the secondary structure map is shown; cleavage sites are highlighted with light blue boxes. Red nucleotides show the area where the primer is binding to the RNA.

Domain 1 – the c-c1-c2 three-way junction

The next primer for mapping domain 1 completely is Sc 160. It covers half stem c2 and all of stem c1. The area also has some interesting elements like long-range tertiary interactions as well as secondary structure elements like bulges. The β-β' intra-domain interaction between positions G145 to G149 in stem c2 and C254 to C259 in the d2a stem of domain 1 is known to play a role in intron compaction during folding (Waldsich and Pyle, 2007). The tetraloop GUAA at positions 98 to 101 interacts with its receptor in D2 (two Watson-Crick base pairs: C422-G581; C423-G580), thereby forming the θ-θ' interaction. This contact is thought to play a role in recruiting D3 as well as the joiner J2/J3 (Fedorova et al., 2003). In contrast to the last one the λ-λ' tertiary interaction is directly involved in catalysis and positioning of D5 in close proximity of the 5'- splice site (Boudvillain et al., 2000). Its position in D1 is A115
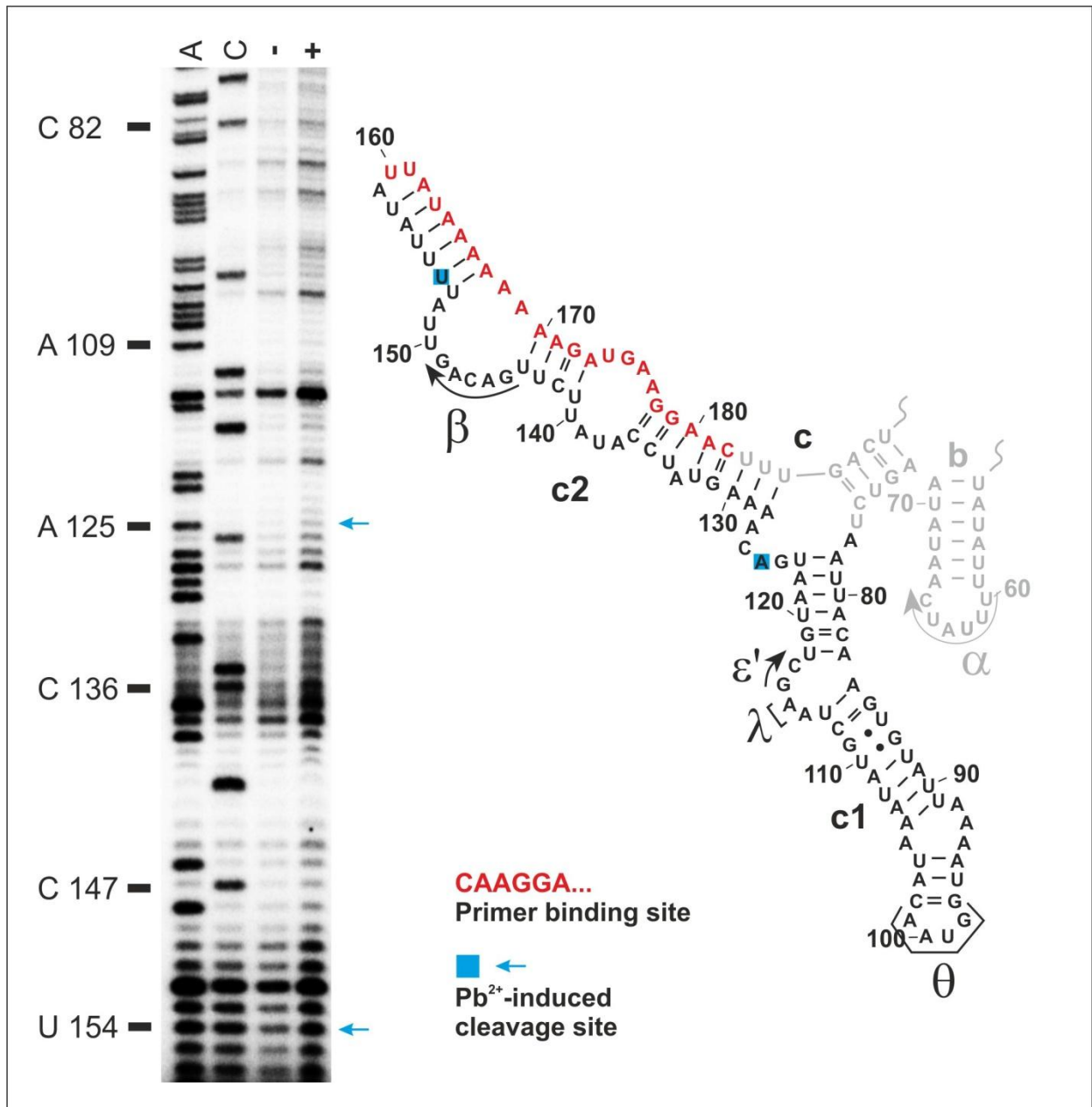

Figure 20. Overview of Pb^{2+}-induced cleavage in domain 1 of ai5γ – the c-c1-c2 three-way junction – mapped with primer Sc 160. Figure legend is the same as for figure 19.

and G5 (which could not be mapped – see figure 19), respectively. The last interaction present in the area which can be mapped by Sc 160 is ε-ε' and this is also an intra-domain interaction located at G116, C117 and G3, C4 (which could not be mapped – see figure 19). It is thought to play also a role in catalysis (Jacquier, 1990). All of the latter three, ε-ε', θ-θ' and λ-λ' are located in the c1 stem of domain 1. In contrast to Sigel et al. I did not observe many cleavage sites in this region except for A125 and U154 (figure 20). Although the gel in figure 20 shows a lot more cleavage sites, most of them had to be considered as non-cleavage sites because they did not show up in the majority of independent experiments (see p. 66 for further discussion). A125 is located at the same position where cleavage sites had been observed *in vitro*, and a metal ion at the site might be of importance to coordinate the structural organization of the three-way junction. Several metal ion binding sites observed *in vitro* are completely missing in my data, for example a really strong cleavage at A137, the region containing λ-λ' and ε-ε' (A114 to C117) and also the bulge region around U90 to A93.

Particularly the absence of cleavage sites in the region involved in tertiary interactions is potentially a very interesting difference to the previously published *in vitro* data (Sigel et al., 2000).

Domain 1 – the core of D1 including the κ–ζ element and the coordination loop

Continuing in domain 1 the next primer Sc 234 covers the by far most interesting region of this domain; a part of the folding control element (κ-ζ element). Besides its central role for intron compaction and folding it is interesting that *in vitro* no metal ion binding sites were discovered when folded to its native state. The long-range tertiary interactions of the folding control element, κ-κ' and ζ-ζ' are crucial for docking of the catalytic centre of the intron, domain 5, onto D1 (Boudvillain and Pyle, 1998; Costa and Michel, 1995). κ-κ' is at positions G213 to A215 in domain 1 and interacts with the Watson-Crick base pairs C818-G845 and C819-G844 in domain 5. ζ is a conserved 11-nt tetraloop-receptor located in D1 at bases U197 to C201 (partly base-pairing on the complementary strand with nucleotides G379 to G384) and targets the canonical GNRA tetraloop ζ' at positions G829 to A832 in D5. Unlike with the terbium(III)-induced cleavage assay I did not observe any metal ion induced cleavage in this area using a lead(II) as a probe (compare to figure 16 left panel). If the Tb^{3+}-induced cleavage sites are an artefact or if they are really important is hard to say due to the bad reproducibility I encountered using a terbium(III)-based cleavage assay. Figure 21 shows a representative gel as well as the corresponding secondary structure map of this area. Sigel et al. did observe some strong cleavage sites located in the internal loop around nucleotides A228 to A225 (Sigel et al., 2000). In line with their data, I also detected Pb^{2+}-induced cleavage in the coordination loop at residues A228 and U224.
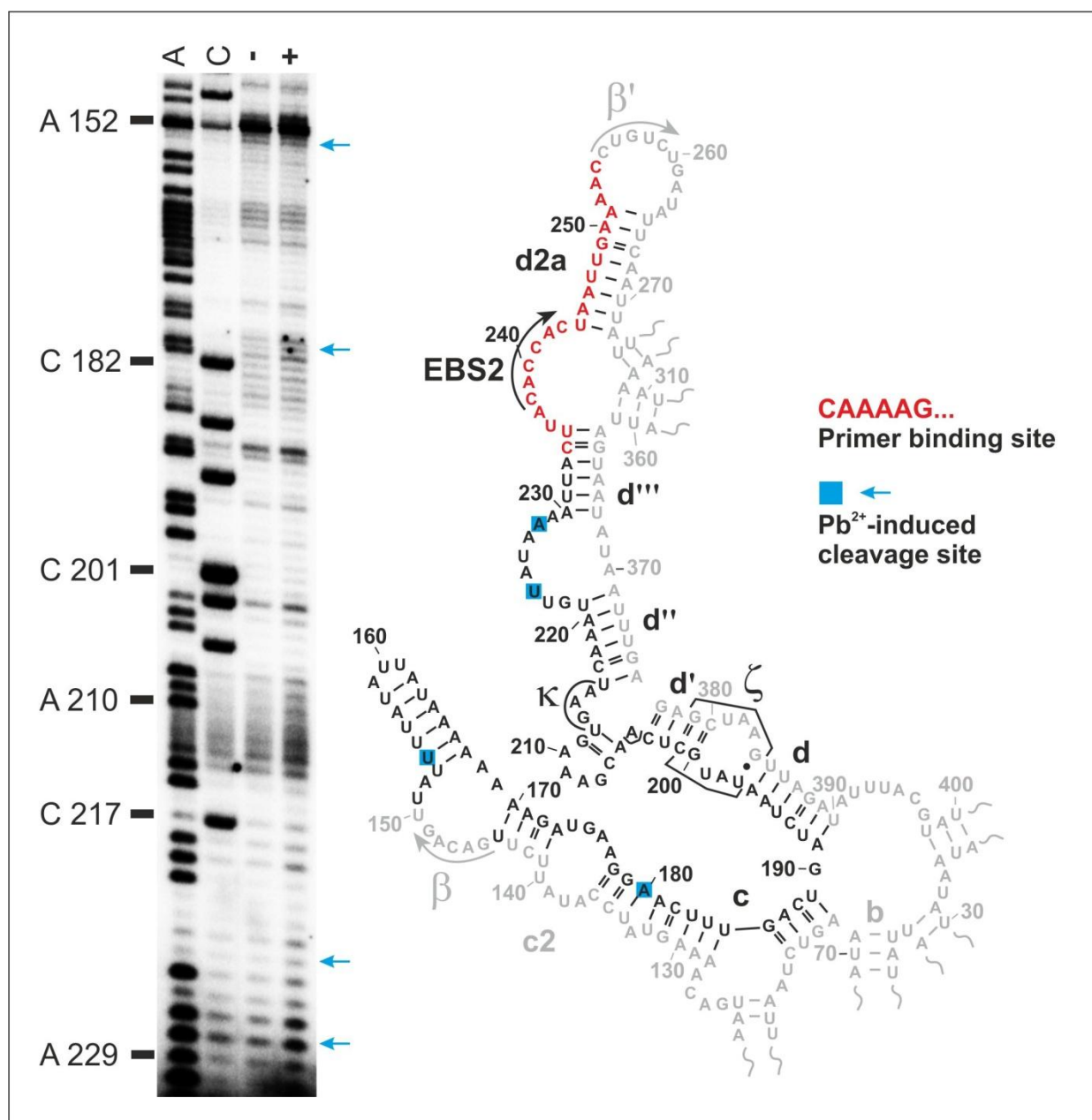
Figure 21. Overview of Pb²⁺-induced cleavage in domain 1 of ai5γ – the core of D1 including the κ–ζ element and the coordination loop – mapped with primer Sc 234. Figure legend is the same as for figure 19.

Domain 1 – elements β' and EBS2

For mapping stems d2a and d2b within domain 1 the primer Sc 312 was used. This area proved to be tricky due to the weak signal of the Sc 312 which was often only ⅓ of the signal intensity of the other primers. The low signal-to-noise ratio therefore made quantification very difficult. The region itself has only two interesting elements to point out; firstly the high A and U content of stem d2b with not a single G or C and secondly the second part of the β-β' interaction which was already described previously on pages 48 and 49. Furthermore exon binding site 2 (EBS2) is also located in this area (C238 to C243) which interacts with its counterpart, the intron binding site located at the 3'-end of the 5'-exon. This interaction helps to tether the exon to the intron and to define the 5'-splice site (Pyle and Lambowitz,

2006; Qin and Pyle, 1998). I could observe four cleavage sites located at A289 and U291 which corresponds to the hairpin at the end of d2b and A303 and A304 (figure 22). The latter ones are at located the beginning of d2b. Sigel et al. reports here additional metal ion binding sites in this area, in particular at EBS2 and residues close to the β' region (Sigel et al., 2000). As can be seen in Figure 22 I do observe some faint cleavage at these important structural elements, but these need additional data sets to be verified.
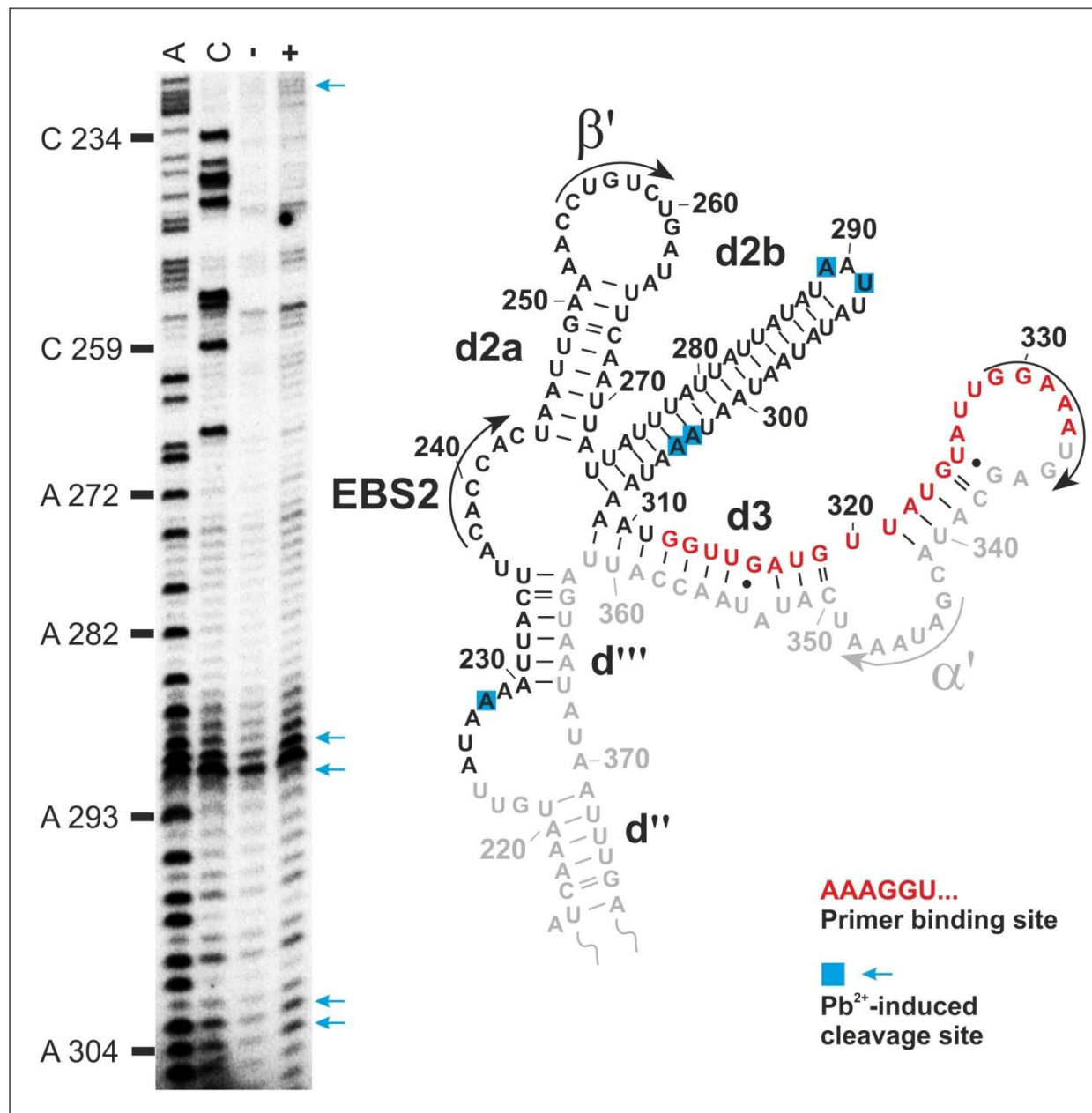

Figure 22. Overview of Pb^{2+}-induced cleavage in domain 1 of ai5γ – elements β' and EBS2 – mapped with primer Sc 312. Figure legend is the same as for figure 19.

Domain 1 – the coordination loop and stem d3 with EBS1 and α'

The region which was mapped using primer Sc 375 did reveal some very interesting aspects. First of all it is the region with the most metal ion-induced cleavage sites observed and secondly this does also correlate very well with the data reported by Sigel et al. In terms of interesting structure elements there are the α-α' interaction as well as the exon binding site 1 (EBS1) at positions G329 to A333.

Figure 23 gives an overview of all cleavage sites detected in this region *in vivo*. Starting with G330, A332 and G335 which are all located in the EBS1 site, additional cleavage sites are found at U340 and A347, A348 located in the α' region. While cleavage sites were observed at EBS1 *in vitro* (Sigel et al., 2000) as
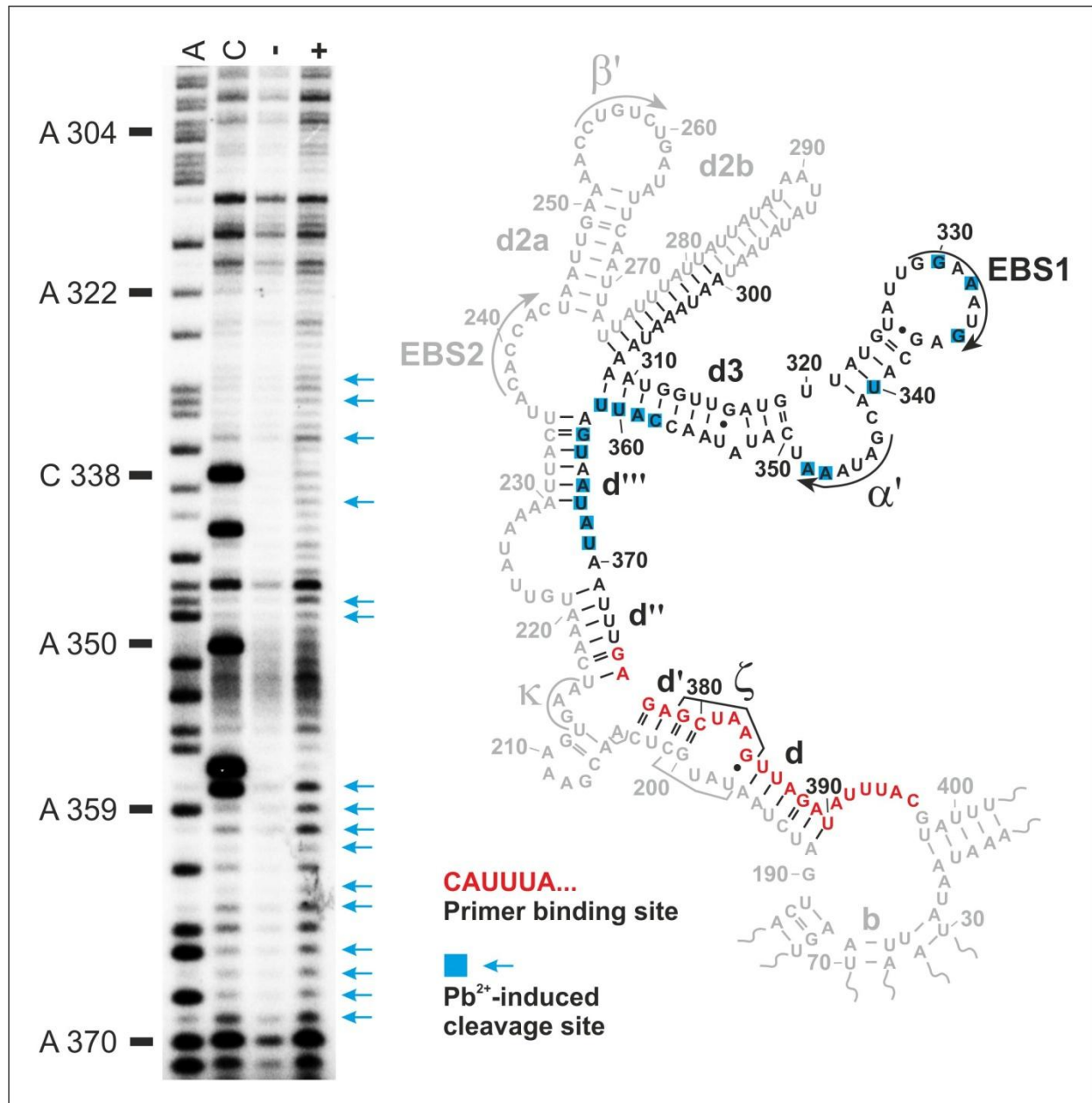

Figure 23. Overview of Pb^{2+}-induced cleavage in domain 1 of ai5γ - the coordination loop and stem d3 with EBS1 and α' - mapped with primer Sc 375. Figure legend is the same as for figure 19. Note that in this area the most cleavage sites were detected.

well, the exact location are slightly different compared to my *in vivo* data (see Figures 12 and 23). This could be due to the difference in radius and coordination geometry between Tb^{3+} and Pb^{2+}. Along the same line, Sigel et al., (2000) detected metal ion-induced cleavage at the bulged A and neighbouring residues in d3, but not directly within α' as I did *in vivo*. The implications which subsequently follow for these differences will be discussed later on. The next cluster of cleavage sites is located around nucleotides C358, A359, U360 and U361 and immediately followed by the next ones, G363 and U364. These residues are at the junction of stems d3 and d''', which are part of a four-way junction, which might also require a metal ion for adopting its proper conformation. The last region where cleavage was observed is within the coordination loop and flanking nucleotides, i.e. A366, U367, A368 and U369. Importantly, cleavage was also observed for the 5' part of the coordination loop, which plays an important role for the positioning of the branch point adenosine of D6 (Hamill and Pyle, 2006). All of these three clusters are also observed *in vitro* with some minor differences in length and/or positions but overall correlation is very good for this area of domain 1.

Domain 1 – the κ–ζ element and stem i

The last primer necessary to cover all of domain 1 is Sc 440 which is technically already starting in domain 2. It maps the basal stem of D2 and stem I in D1 which are both protruding from the central wheel. In addition, the 3' part of the κ–ζ folding control element embedded in stems d and d' as well as junction Jd/I part of the D1 five-way junction are shown in figure 24. The long-range tertiary interactions which are part of these regions, θ-θ' and ζ-ζ' are both inter-domain interactions. The first one confers stability to the structure by forming a tetraloop-receptor interaction between D1 and D2 (Costa et al., 1997). The other one, ζ-ζ' anchors D5 onto the D1 scaffold and ζ is also is part of the folding control element mentioned already earlier. The region itself does not show significant metal binding sites *in vitro* (except for weak cleavage at Jd/i) and domain 2 is also shortened to a hairpin in the D135 ribozyme used by Sigel et al., (2000). Due to this truncation of domain 2 the two metal ion binding sites identified *in vivo* do not show up on the *in vitro* map. Their location is A429 and U430. Figure 24 shows that no other cleavage sites were identified in this region.
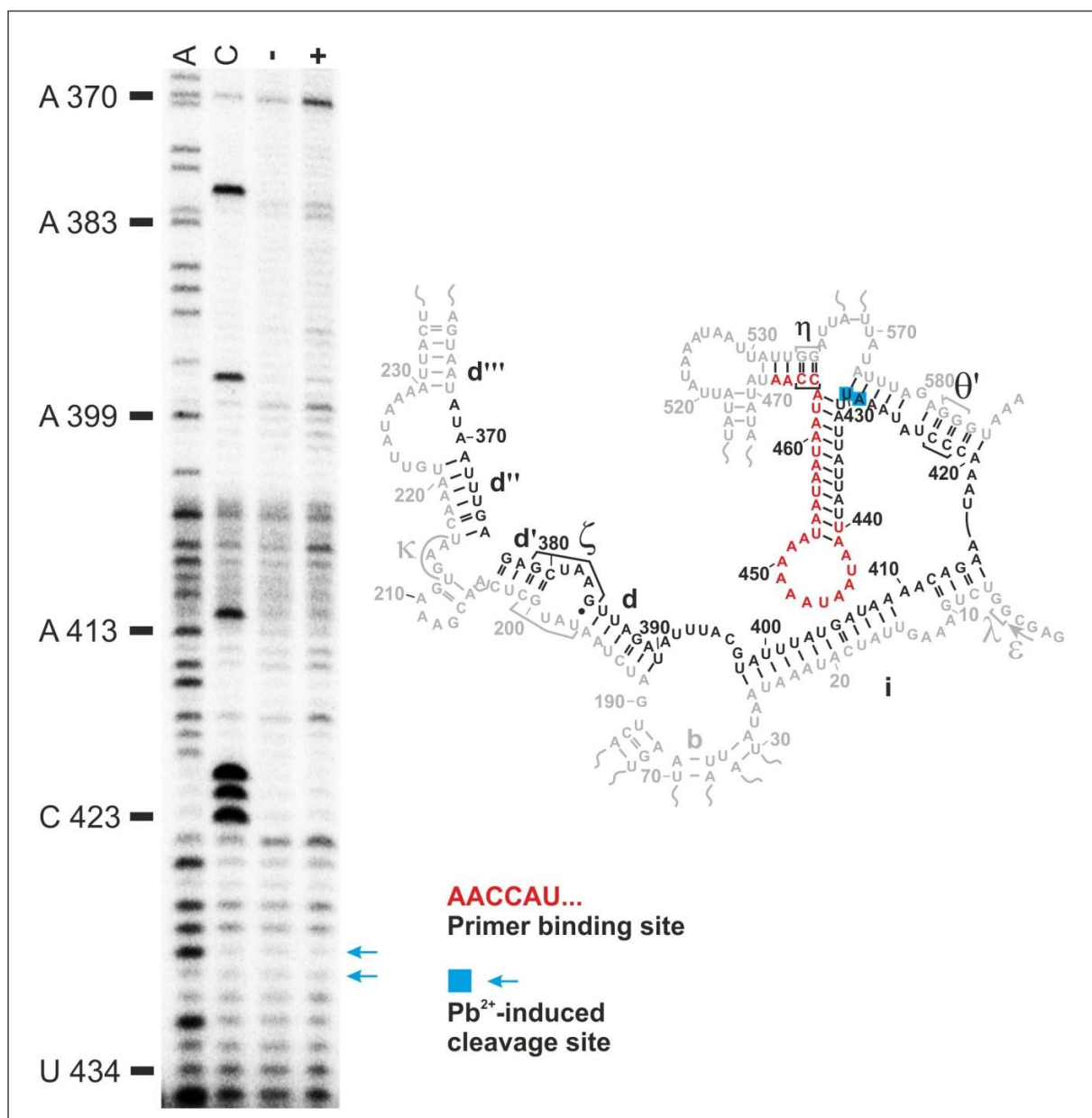
Figure 24. Overview of Pb^{2+}-induced cleavage in domain 1 of ai5γ – the κ–ζ element and stem i – mapped with primer Sc 440. Figure legend is the same as for figure 19. Note the strong strand break at position A370 probably makes it hard to map the cleavage sites at A368 and U369. This is also the last part of domain 1.

Domain 2 – 5' part of D2

The next domain in which metal ion binding sites were mapped is domain 2. It is a rather short one with approximately 160 nucleotides and it is therefore possible to map it with only two primers. Its presence is not required for catalysis and it harbours two long-range tertiary interactions, θ-θ' and η-η' which form a tetraloop-receptor interaction with between D2 and D1 or D6, respectively (Chanfreau and Jacquier, 1996; Costa et al., 1997).

Unlike for domain 1 there is no *in vitro* data available for domain 2 because in the D135 ribozyme used *in vitro* the whole domain is truncated to a short hairpin. The first part of domain 2 is covered by primer

Sc 520 and does not reveal any metal ion binding sites as one can see in figure 25. At a first glimpse the gel might imply that there are cleavage sites within D2, e.g. at positions 447-452 or 504-507. However, due to a significant loading difference in all gels additional data sets are required to verify these potential cleavage sites.
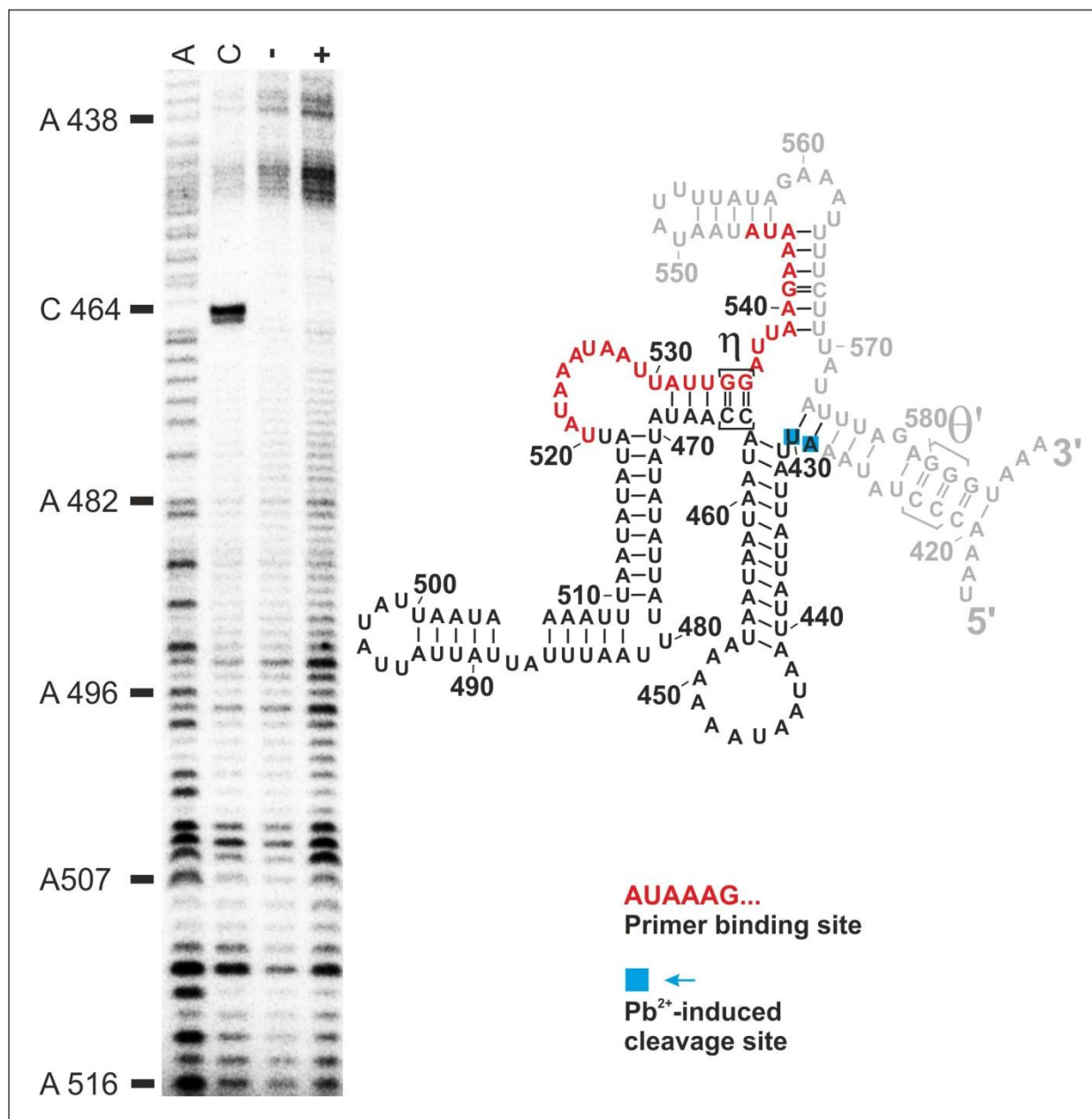

Figure 25. Overview of Pb^{2+}-induced cleavage in the 5' part of domain 2 of ai5γ mapped with primer Sc 520. Figure legend is the same as for figure 19. Note that no cleavage sites are observed in this area.

Domain 2 – 3' part of D2 and J2/3

With domain 2 being a small domain only a second primer besides Sc 520 is needed to finish mapping: Sc 588. It anneals to a sequence which is located outside of domain 2 but except for four bases of J2/3 (A584 to A587) the remainder of its template sequence is located within domain 2. As already mentioned no *in vitro* data by Sigel et al. is available for this region except for the very beginning of the

primer, J2/3. Here, in the area of the γ–γ' interaction a very prominent cluster of cleavage sites is observed *in vitro* (Sigel et al., 2000) and it has been shown that metal ions bound to this site contribute to reaction chemistry (Sontheimer et al., 1999). However, *in vivo* there is no cleavage observed at this region. A possible explanation could reside in the fact that in the wt yeast strain the majority of the intron population is in the spliced form and thus these catalytic metal ions are no longer needed and their original binding pockets are not accessible anymore. Alternatively, Pb^{2+} is too large to be able to displace the catalytic metal ions, which would be in line with the fact that the spliced intron maintains its active site *in vivo* (Liebeg and Waldsich, unpublished). As can been seen in figure 26 cleavage sites have also not been detected within the 3' part of D2. The two long-range tertiary interactions of this domain have been already described (θ-θ' and η-η').

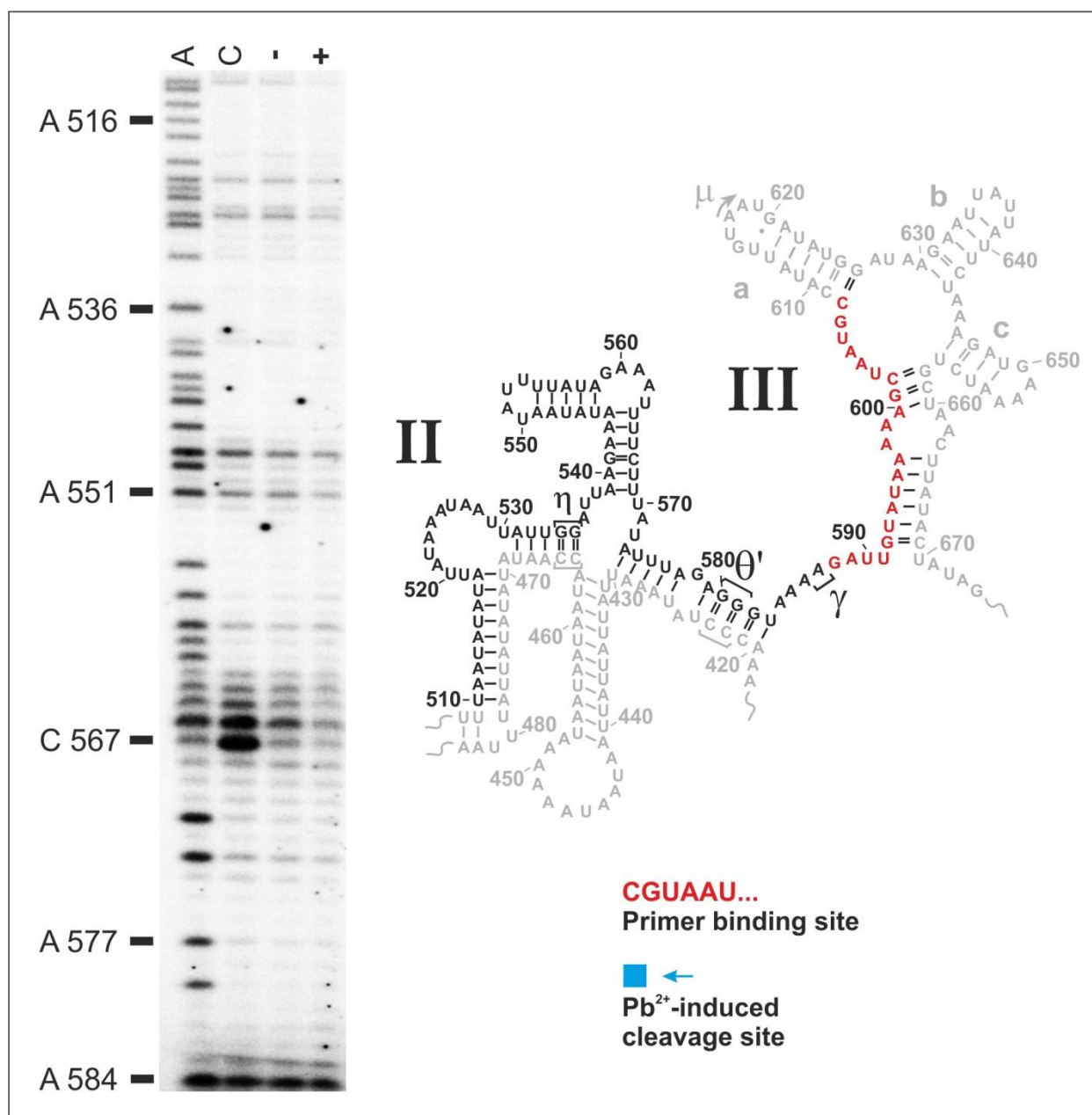**Figure 26.** Overview of Pb^{2+}-induced cleavage in the 3' part of domain 2 and J2/3 of ai5γ mapped with primer Sc 588. Figure legend is the same as for figure 19. Note that no cleavage sites are observed in this area.

Domain 3 – stems a, b and c as well as the D3 basal stem with the A-rich internal loop

Mapping of metal ion binding pockets continues with domain 3 and the primer annealing to this region, Sc 658. Like domain 2 this domain is also short with approximately 80 nucleotides but its function in catalysis is way more important than domain 2. Domain 3 is usually referred to as a catalytic effector, as it is not strictly required for catalysis but its presence enhances reaction rates tremendously (Fedorova and Zingler, 2007). However, with my *in vivo* Pb^{2+}-induced cleavage assay no metal ion binding were detected which is in sharp contrast to the *in vitro* study by Sigel et al. (2000) where several clusters of cleavage sites had been observed. Notably, most of these *in vitro* cleavages sites (Sigel et al., 2000) are of weak intensity and only two of medium or strong intensity. These two sites are the loop b and the junction connection stems b and c (U634-U638; U643-A644). A representative gel is shown in figure 27 as well as the corresponding region of the secondary structure map.
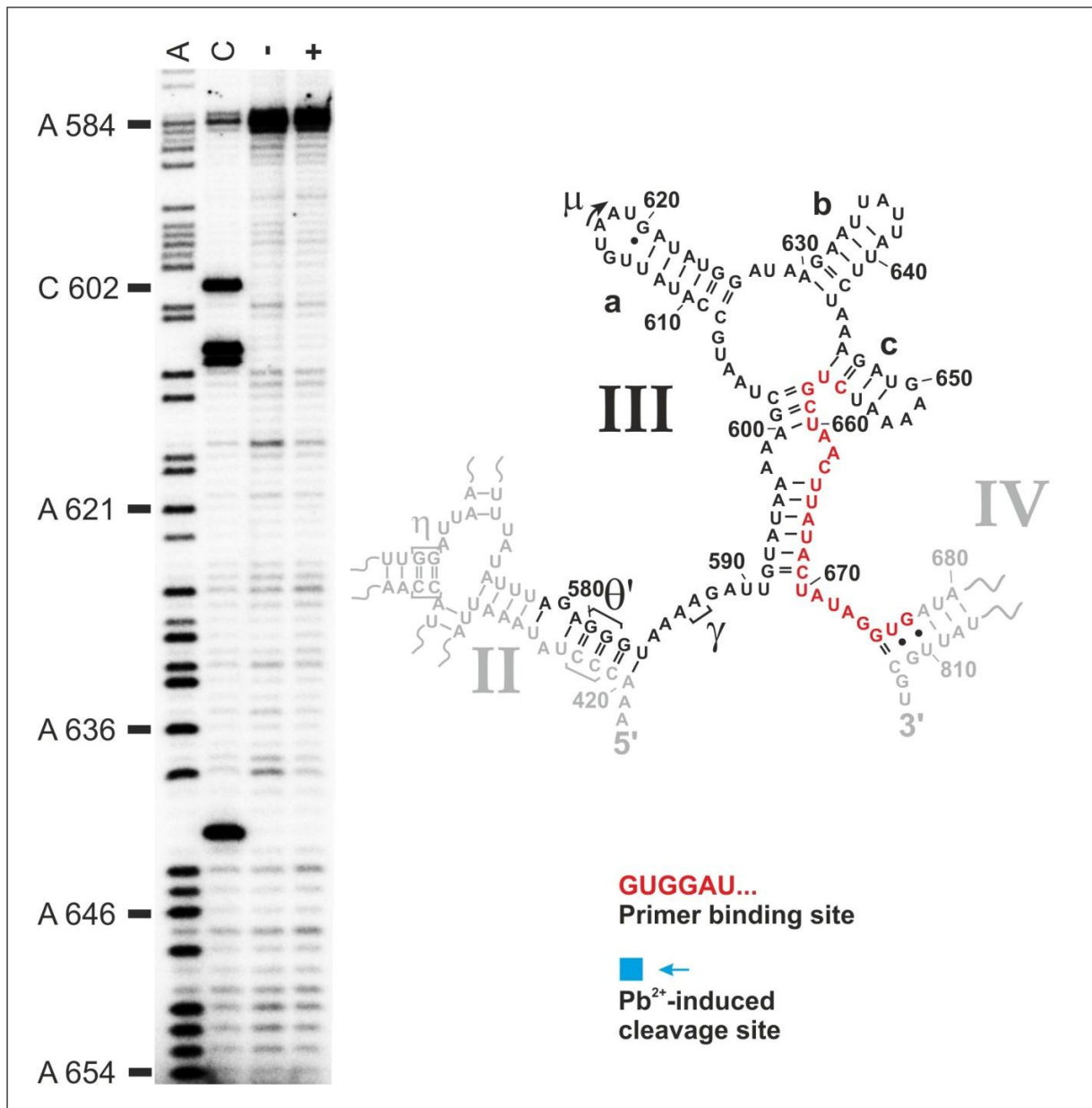

Figure 27. Overview of Pb^{2+}-induced cleavage in domain 3 of ai5γ – stems a, b and c as well as the D3 basal stem with the A-rich internal loop – mapped with primer Sc 658. Figure legend is the same as for figure 19. Note that no cleavage sites are observed in this area.

Domain 4 – the A-rich loop of the D3 basal stem and the 5' part of D4

The next region is domain 4 which often contains an open-reading frame for the maturase protein involved in intron splicing and intron mobility (Lambowitz and Zimmerly, 2004). However this is not the case for ai5γ. Knowledge about domain 4 is limited and no tertiary contact has yet been reported for this domain. For *in vitro* studies of group II introns domain 4 is often truncated; this is also the case for the study by Sigel et al., (2000). I mapped this area using primer Sc 754 which also covers the short rest of domain 3.
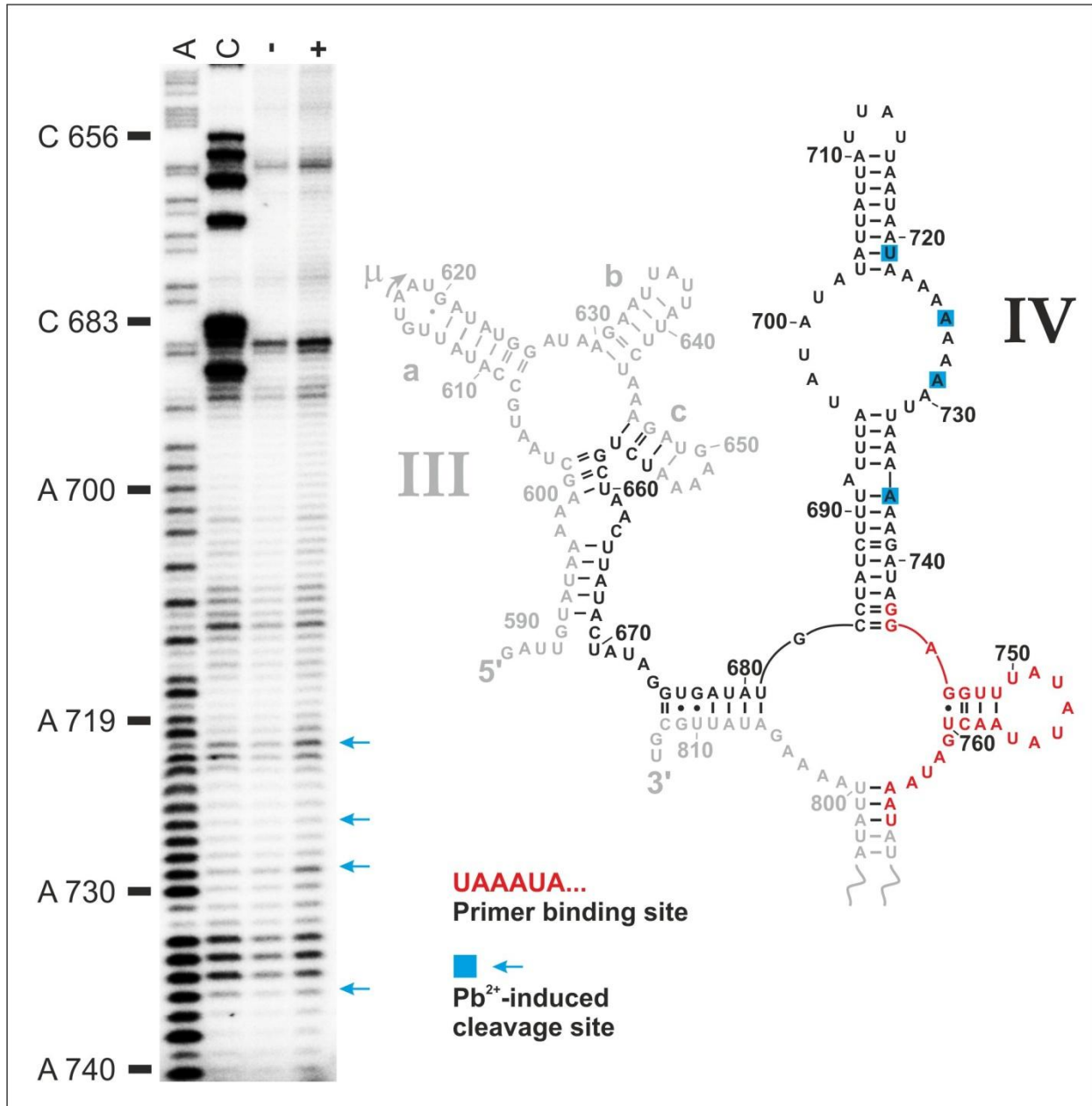

Figure 28. Overview of Pb^{2+}-induced cleavage in domain 4 of ai5γ – the A-rich loop of the D3 basal stem and the 5' part of D4 – mapped with primer Sc 754. Figure legend is the same as for figure 19. Note that cleavage was observed at positions U721, A726, A729 and A736.

In terms of interesting results this region has more to offer than the previous primers as I observed several cleavage sites in stem a of D4. The first one is U721 which is located in a double-stranded region, followed by two additional cleavage sites (A726 and A729) which are part of a close internal asymmetric loop. The last one, A736 is again located in a double-stranded region. Consistent with the *in vitro* data (Sigel et al., 2000) I did also not detect any lead cleavage in the 3' part of the basal stem of D3 *in vivo*. The results are summarized in figure 28.

Domain V – the active site centre

The last primer which was used for mapping metal ion binding sites within the group II intron ai5γ is Sc 866. It covers half of domain 6 as well as the whole of domain 5 and the last part of domain 4. This is possible due to the relative small size of domains 5 and 6 which are only 30 nucleotides in length each. Domains 5 and 6 are by far the most interesting ones due to their roles in reaction chemistry. Starting with domain 5 which is the catalytic heart of ai5γ and harbours the κ-κ' as well as the ζ-ζ' long-range tertiary interaction. Both of them dock on their target sites within domain 1 (Boudvillain and Pyle, 1998; Costa and Michel, 1995). Furthermore, the λ-λ' interaction, which is part of the active site, is also formed between domains 1 and 5 (Boudvillain et al., 2000). Domain 6 harbours the branch point adenosine which important for the branching pathway of splicing and the η-η' inter-domain interaction (Fedorova and Zingler, 2007). For more detailed information on each interaction please refer to section 1.3.2 on page 14.

Technically speaking this area proofed to be difficult to map via primer extension due to some very strong breaks at the beginning of the extension, which are mainly caused by long stretches of G/Cs and stable secondary structure. Therefore the signal-to-noise ratio was low and around the same level like primer Sc 312. This of course is unfortunate compared to the importance of this particularly region. Quantification did not reveal any metal ion induced cleavage sites (figure 29) but in my opinion this is more due to technical problems than the actual absence of metal ion binding sites. *In vitro* the situation is different because Sigel et al. did observe several clusters of cleavage sites located for instance around G811 to U814 and G836 to G840. Like for domain 3, only a single cleavage site of medium intensity was observed (all others are weak sites) centred around the di-nucleotide bulge, which forms a triple helix with the conserved AGC catalytic triad and is thus part of the active site centre (Toor et al., 2008). Accordingly, this cleavage site may result from one of the catalytic metal ions (Toor et al., 2008). The lack of lead-induced cleavage *in vivo* could possibly be explained by the fact the catalytic metal ions are no longer needed in the spliced intron variant and their original binding pockets might not be accessible anymore. Alternatively, Pb^{2+} is too large to be able to displace the catalytic metal ions.

It should also be mentioned that domain 6 could not be mapped with the method of primer extension and therefore there is also no *in vivo* data about this domain. Furthermore the 3' part of D4 (U788 to A740) has not been entirely mapped due to insufficient separation of the bands on the gel; here a second primer covering this area is needed for future analysis.
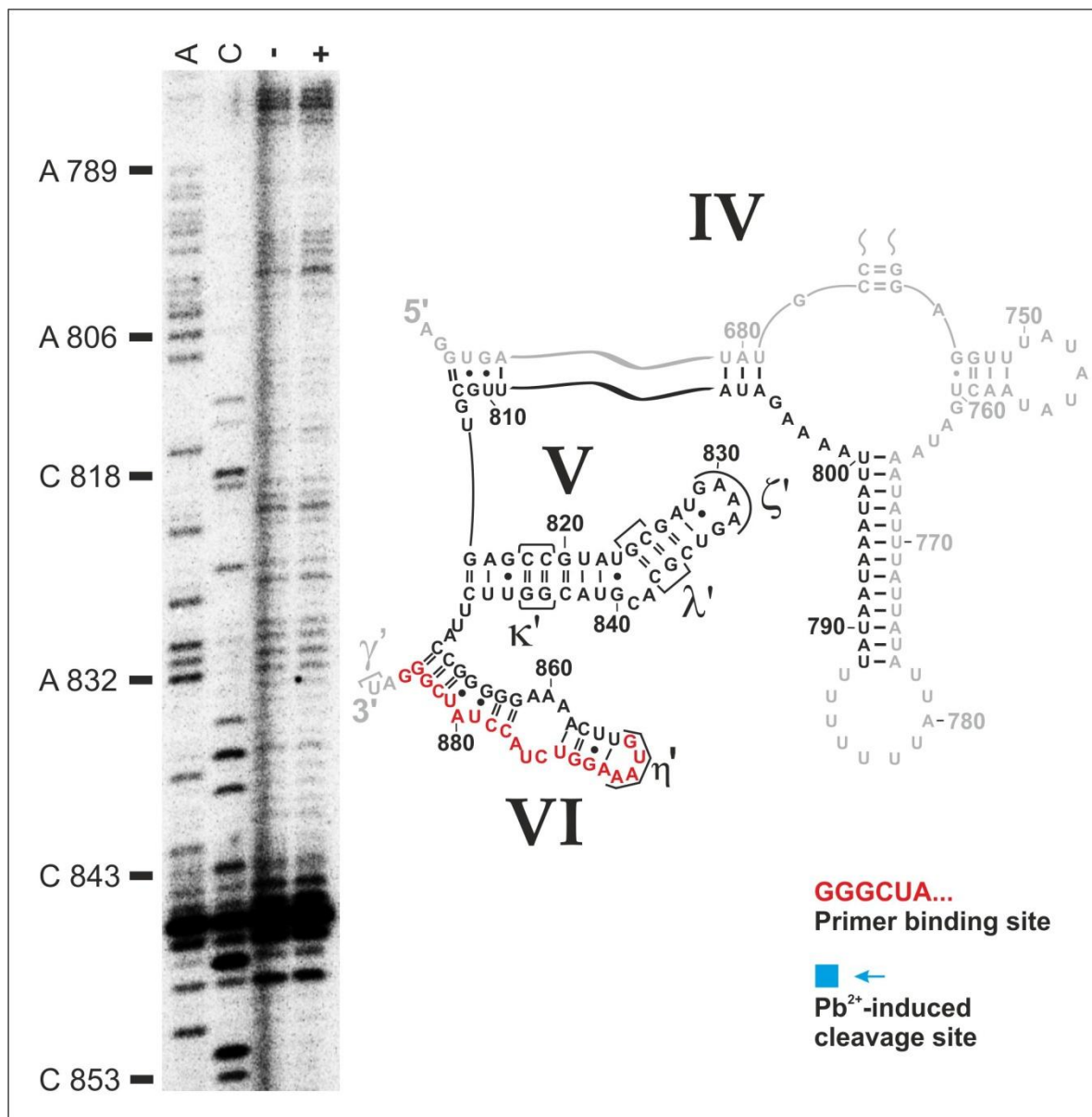
Figure 29. Overview of Pb^{2+}-induced cleavage in domain 4 and 5 of ai5γ mapped with primer Sc 866. Figure legend is the same as for figure 19. Note that no cleavage sites were observed. Also, as I mapped the spliced intron form, it was not possible to map D6.

5.3   Summary

Using a Pb^{2+}-induced cleavage assay and primer extension I mapped the metal ion binding pockets within the group II intron ai5γ in yeast mitochondria. The method detects metal ion binding sites which might be important for intron splicing (catalysis and folding). The cleavage sites observed are not equally distributed throughout the intron, the majority of them is clustered in domain 1 which is also the largest of all six domains. For the last domain, D6, mapping was not possible due to the fact that the primer binds in domain 6 and therefore most of the domain is covered by the primer itself leaving no space for extension in this area. All cleavage sites observed so far are summarized in figure 30.
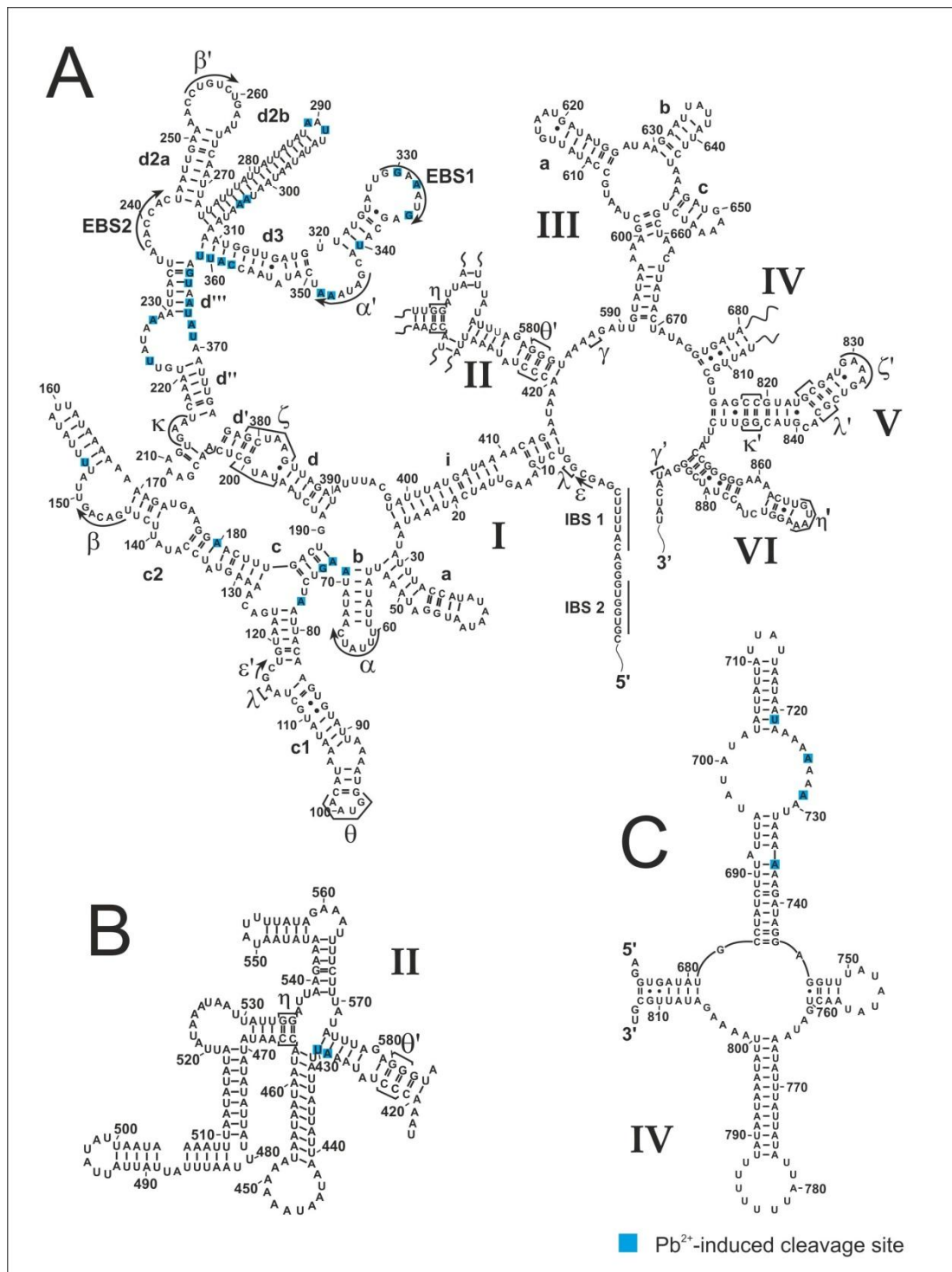
Figure 30. Pb²⁺-induced cleavage sites superimposed on a secondary structure map of the group II intron ai5γ. Panel A shows domains 1, 3, 5 and 6, panel B shows domain 2 and in panel C domain 4 is shown. Cleavage sites are marked by light blue boxes; tertiary interactions with Greek letters.

6   Discussion

Lead(II)acetate proved to be a powerful probe for investigating metal ion binding sites *in vivo*. By comparing newly collected *in vivo* data to already known *in vitro* data from Sigel et al., (2000) new insights into the intracellular formation of metal ion binding sites can be gained.

6.1   Does Mss116p compensate for metal ion binding sites in vivo?

The group II intron ai5γ is known to require high salt conditions and elevated temperature to fold correctly *in vivo* (Fedorova and Zingler, 2007). The folding pathway *in vitro* is a two-step mechanism with the slow formation of a compact intermediate, which contains most of the D1 structural elements followed by a fast transition to the native state (Fedorova et al., 2007; Su et al., 2005; Waldsich and Pyle, 2007; Waldsich and Pyle, 2008) . Divalent metal ions like Mg^{2+} play an important role for the compaction of the intron and subsequently folding (Swisher, 2000) SU 2003. However the situation is completely different *in vivo*; first of all the concentration of divalent ions is about two orders of magnitude lower than for *in vitro* experiments (1mM vs. 100mM) (Fedorova and Zingler, 2007) and secondly the temperature of the host organism is also lower than 42° C (for yeast it is 30° C). In order to enable the group II intron ai5γ to fold correctly in its natural habitat, the mitochondria of yeast, additional help is required. It has been shown that the DEAD-box protein Mss116p promotes efficient splicing of the ai5γ intron *in vivo* (Huang et al., 2005). In other words, in a Δmss116 knock-out strain splicing of ai5γ is barely detectable (Huang et al., 2005; Liebeg and Waldsich, unpublished), as the intron is largely unfolded in the absence of Mss116p (Liebeg and Waldsich, unpublished), implying that the protein confers stability to the intron RNA.

Analysis and quantification for a Pb^{2+}-induced cleavage assay *in vivo* did reveal that metal ion binding sites decrease significantly compared to *in vitro* data (see figure 10 and 30). By losing these metal ion binding pockets and the interactions conferring stability to the intron it becomes even harder for the group II intron ai5γ to adopt its native fold. A possible escape from that unpleasant situation is a protein co-factor that compensate for lost metal ions and confers again stability to the intron. Thus, my data would suggest a model in which Mss116p substitutes certain metal ions allowing ai5γ to reach its native fold. However, it has recently been shown in our lab (Liebeg and Waldsich, unpublished) that the spliced intron maintains its native structure. Importantly, it appears that Mss116p is not bound to native intron *in vivo* and *in vitro*, but is required for the formation of a folding intermediate centred on D1 (Liebeg and Waldsich unpublished, (Fedorova et al., 2010). In light of these findings, it is rather unexpected and perhaps unlikely that Mss116p substitute for metal ions in the native state of the ai5γ intron. Accordingly, other positively charged small ligands, like polyamines, provide additional support for ai5γ folding *in vivo*. Alternatively, it is possible that there is a potential problem with the data sets which I have acquired so far (see section 7.2), as I came to realize that in the wt yeast strain ai5γ spliced only

to reduced levels compared to my initial studies and as reported by co-workers (Liebeg and Waldsich, unpublished).

Lastly, it should be noted that the reduced number of metal ion cleavage site can also be the results of i) Pb^{2+} being unable to displace Mg^{2+} from its binding site (due to affinity or size of the pocket); Tb^{3+}, which was used to map metal ion binding sites *in vitro* (Sigel et al., 2000) has an ionic radius closer to that of Mg^{2+}. ii) *In vitro*, an ai5γ ribozyme construct was used that lacked D2 and D4 (Sigel et al., 2000). In yeast mitochondria the natural ai5γ intron, which contains full-length domains 2 and 4, was studied and thus the presence of D2 and D4 could potentially influence metal ion binding *in vivo* as well.

6.2   Impact of splicing of ai5γ on metal ion binding sites

It seems very important to have a pool of ai5γ RNAs which must be spliced to a high percentage. Quantification revealed that the lower the part of correctly spliced ai5γ the fewer the amount of cleavage sites observed. These two things seem to be connected in a reciprocal way. Starting with my experiments in June 2009 to collect data sets needed for quantification the splicing values were all above 75% meaning 75% of all ai5γ RNAs was spliced correctly with only 25% unspliced. In contrast to this I encountered a much lower splicing activity for the last round of experiments in June 2010 (see figure 31 for comparison).
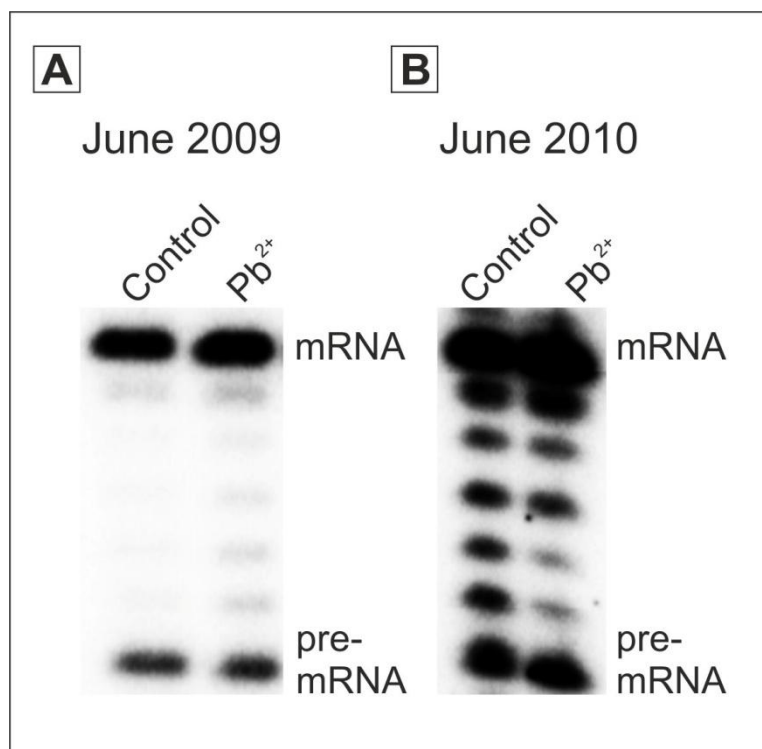

Figure 31. Comparison of two poison primer assays for monitoring splicing values *in vivo*. (**A**) RNA extracted from untreated cells as well as from Pb^{2+}-treated cells was reverse transcribed. Samples date from June 2009. Note that there are no additional bands between the upper and lower ones. (**B**) RNA extracted from untreated cells as well as from Pb^{2+}-treated cells was reverse transcribed. Samples date from June 2010. In contrast to (**A**) there are strong bands between upper and lower band visible. This indicates a severe change in splicing and/or RNA quality compared to (**A**).

It is obvious that splicing did not work very well for the samples prepared around June 2010 (figure 31B), the additional bands between mRNA and pre-mRNA show that other products have been made, indicating a decrease in correctly spliced ai5γ RNA or a general decrease in RNA quality. Importantly, this is not caused by increased lead concentration, as the same bands are visible in the control, in which RNA extracted from untreated cells was reverse transcribed. The reason for this circumstance is un-

clear, as neither the protocol nor solutions were changed. However one possibility which should investigated further is the fact that chemicals for preparing YPD- as well as YP-Raffinose media were re-ordered but from a different company. In any case, it seems that a significant amount of the ai5γ intron does currently not splice in the wt strain. This is most likely due to a defect in folding of the intron, thereby explaining that many of the originally cleavages site were not seen in the recent data sets. As we only summarized those metal-induced cleavages that were seen in at least 2/3 of all experiments, many of the initially observed cleavage sites had therefore to be excluded from the final list summarized in figure 30.

The impact of splicing on Pb^{2+}-induced cleavage can also be observed on the representative gels and is best seen when comparing primer Sc 160 (c-c1-c2 three-way junction) from two independently RNA isolations (figure 32). RNA extracted from Pb^{2+}-treated cells can either show a lot of metal-induced cleavage sites (figure 32A) or none (figure 32B) depending on the splicing activity of the intron.
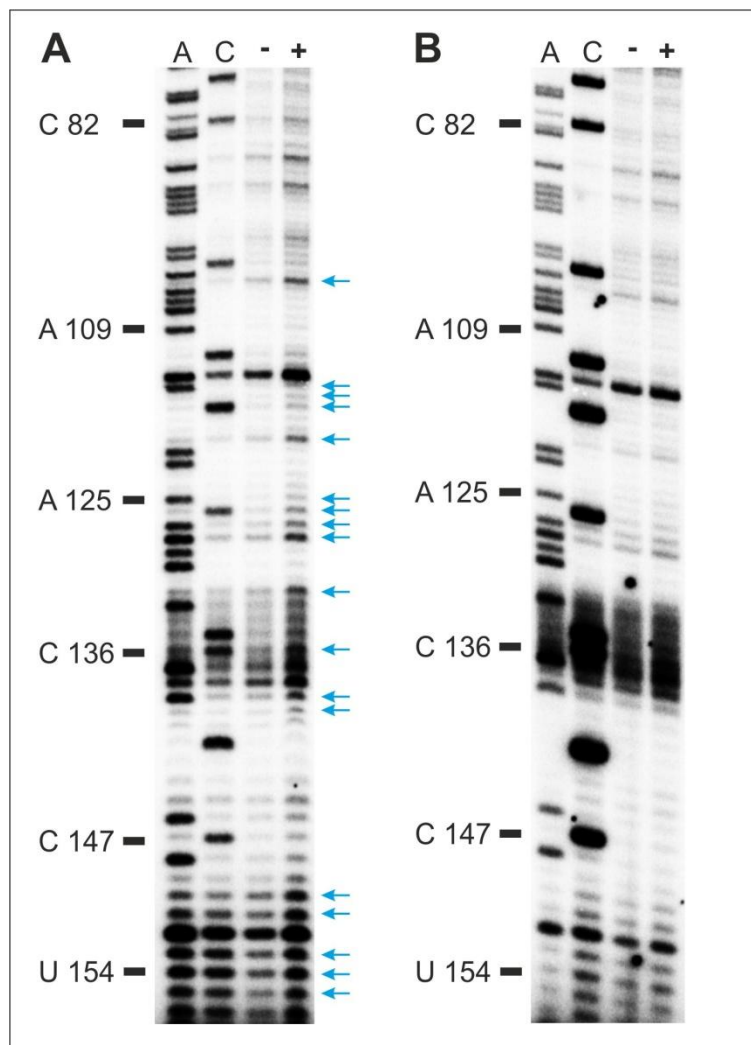

Figure 32. Impact of splicing of ai5γ on detecting metal ion binding sites *in vivo*. Representative gels showing the c-c1-c2 three-way junction of the ai5γ group II intron. Lanes A and C denote the sequencing lanes. In the control lanes (-), RNA which was extracted from untreated cells was reverse transcribed to detect natural, Pb^{2+}-independent stops of the Reverse Transcriptase. In the sample lanes (+), RNA which was extracted from Pb^{2+}-treated cells was reverse transcribed to detect stops of the RT due to Pb^{2+}-dependent cleavage. (**A**) shows an experiment, in which ai5γ spliced as expected in the wt strain (> 75%; from June 2009). Note that cleavage sites are observed (marked by light blue arrows). (**B**) In contrast, an experiment, in which ai5γ spliced as expected in the wt strain (< 60%; from June 2010). No cleavage was detected.

The direct connection between splicing values and detectable metal ion binding sites makes it necessary to re-evaluate my results. When comparing my *in vivo* data to the *in vitro* study already available (Sigel et al., 2000) the correlation for metal ion binding sites was rather low (compare figure 30 with

figure 10) except for the region of primer Sc 375 in domain 1. Taking into account the drastically affected splicing activity and reduced RNA quality and the detectable amount of cleavage sites the question arises of how the picture would change when only considering the data obtained from the well-spliced population of ai5γ RNAs. I expected that that the majority of metal ion binding sites *in vivo* would correlate very well with those *in vitro* and I therefore summarized all cleavage sites (figure 33) observed only in Pb^{2+}-induced cleavage assays which showed normal splicing activity (about ¾ of spliced ai5γ).
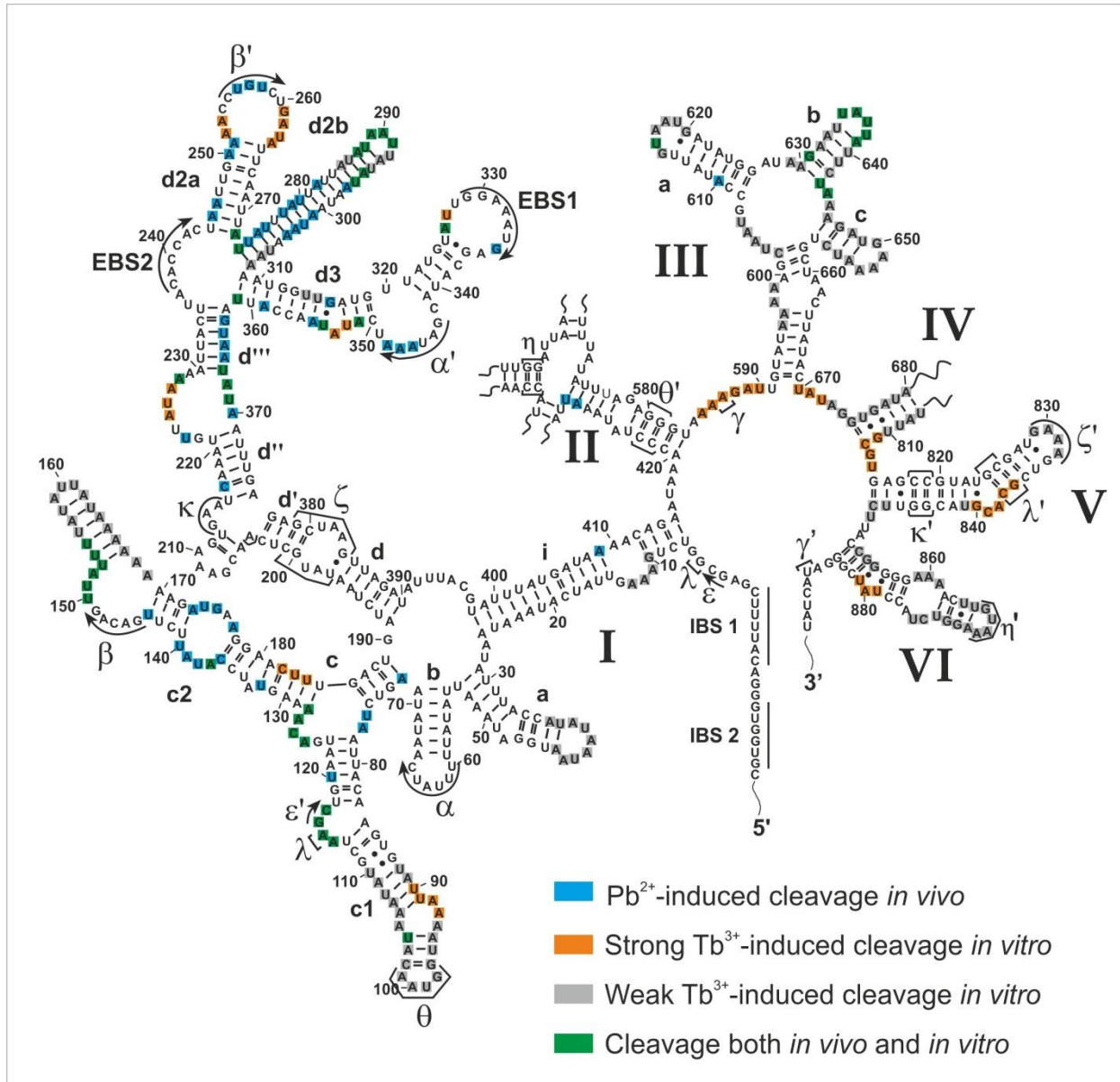
Figure 33. Comparison between *in vivo* and *in vitro* metal ion binding sites based on only those data, in which the ai5γ pre-RNA had spliced efficiently. Blue boxes show Pb^{2+}-dependent cleavage sites only observed *in vivo*, orange (strong) and gray boxes (weak) show Tb^{3+}-induced cleavage sites observed *in vitro* (Sigel et al., 2000) and cleavages sites which are observed both *in vivo* and *in vitro* are marked with green boxes. Domain 2 and 4 are not shown.

When only taken into account data generated by experiments with correctly spliced ai5γ RNA the situation changes dramatically as seen in figure 33. The position of cleavage sites now shows very good correlation with data by Sigel et al., 2000, at least for the strong cleavage sites; the majority of them are also observed *in vivo.* Firstly, as mentioned already above, the number of cleavage sites observed increases significantly throughout domains 1, 2 (not shown), 3 and 4 (not shown). For domain 5 and the 3'-part of domain 4 it is difficult to suggest if the situation may change due to technical problems with primer Sc 866. However I think it is quite possible that in these regions some cleavage sites will be observed as well.

Interestingly some clusters of cleavage sites are not observed at all *in vivo*, for example in domain 1 nucleotides U90 to A93 in stem c1 and nucleotides C182 to U184 in stem c2. Also, the linker region between domains 2 and 3 (J2/3) around bases A585 to U590,which are involves the γ-γ' site of interaction and in the base triple interaction with the AGC triad in D5 (Toor et al., 2008), does not seem to be cleaved *in vivo*. This is of particular interest due to metal ions which are involved in the recognition and cleavage of the 3'-splice site (Sontheimer et al., 1999) as well as the fact that two potentially catalytic metal ions were detected close to D5 and J2/3 in the crystal structure of the *Oi*. group II intron (Toor et al., 2008). For the other metal ion binding sites located U670 to U672, G811 to U814, G836 to G840 (involves λ-λ' interaction) no *in vivo* equivalent was observed, but this could change when optimization for primer Sc 866 are finished. On the other hand, as mentioned before, it is well possible that the binding pocket for the catalytic Mg^{2+} ions is narrow compared to others and thus Pb^{2+} cannot enter these sites and in turn it does not induce cleavage within D5 and J2/3. The last main difference is that due to experimental limitations I cannot map D6 and therefore assess whether a metal ion is close to U879 to U881 *in vivo* as previously reported *in vitro*.

Of all other cleavage sites most of them are detected both *in vivo* and *in vitro,* although their position is sometimes shifted for one or two nucleotides. This can be seen best in domain 1, stem d2a for positions A250 and A251 *in vivo* which are shifted to A252 and A253 *in vitro.* Another possibility is an increased length of the cleavage sites *in vivo* like this is the case for the region between G363 and A370. *In vitro* cleavage is only observed around bases U367 to U369 which extends *in vivo* and involves five more nucleotides (G363 to A366 on the 5'-end and A370 on the 3'-end). However, there are also some cleavage sites which are only detected *in vivo*: e.g. nucleotides A347 to A349 in domain 1 which participate in the α-α' interaction. In contrast, A72, U76 and A77 are not involved in long-range tertiary interactions, but these residues are located next to the five-way junction in D1 and a metal ion might be necessary to co-ordinate and stabilize folding of this complex substructure.

Having now seen the severe impact of correct splicing of ai5γ on the detection of metal ion binding sites *in vivo*, it remains possible that the cleavage sites observed *in vivo* correlate very well with those already published by Sigel et al., (2000). Therefore, I also consider an alternative model inasmuch the stabilizing cofactor Mss116p influence the formation of metal ion binding pockets *in vivo* (section 7.3).

6.3   Mss116p and its role for folding of ai5γ

The ability to detect metal ion binding sites with a Pb^{2+}-dependent cleavage assay *in vivo* is strongly connected to the splicing values of the yeast strain used. I observed a tremendous decrease of Pb^{2+}-dependent cleavage sites when the group II intron ai5γ spliced at reduced efficiency. Therefore I decided to analysis only data from experiments with a high percentage of correctly spliced ai5γ RNA molecules ending up with a summary map (figure 33), which shows a completely contrary picture than figure 30. Summing up, my *in vivo* results now correlate very nicely with those already published by Sigel et al. indicating a different role for Mss116p than compensating for metal ion binding sites in the native state of ai5γ. *In vitro* it has been shown that the rate-limiting step for folding (compaction of domain 1) involves the formation of an unstable intermediate which is then captured by Mg^{2+} and the balance is shifted towards the native state (Waldsich and Pyle, 2007) and Waldsich 2008. The main element involved in this process is the folding control element embedded centrally in domain 1 (Waldsich and Pyle, 2007). Indeed, metal ion binding was observed at the κ–ζ element in the collapsed state of the ai5g intron, which were no longer detectable in the native state (Waldsich and Pyle, 2008). Based on the data shown in figure 33 I suggest a mode of action for Mss116p, in which the protein specifically targets and stabilizes this κ–ζ element during compaction and therefore shifting the balance towards the second step of folding, the fast transition to the native state, which is by now known to be independent of Mss116p (Fedorova 2010, Liebeg and Waldsich, unpublished). In other words, Mss116p appears to substitute for Mg^{2+} during the collapse of the intron by facilitating charge neutralizations and/or stabilizing a folding intermediate *in vivo*.

6.4   Perspectives

In order to investigate which one of the above models is correct additional experiments must be carried out. First of all, an *in vitro* Pb^{2+}-dependent cleavage assay in presence of Mss116p is needed in order to investigate possible differences in position and/or length of metal ion binding sites when compared to *in vitro* data without Mss116p (Sigel et al., 2000) as well as to my *in vivo* data. This will be the diploma project of Eva-Maria Steiner.

Of course, some aspects must be improved for the present work; new primers must be designed for the regions which were not mapped at all or only badly (5'-region of domain 4 and the whole domain 5), and for the Pb^{2+}-induced cleavage assay I suggest a minor change to the protocol: Instead of having only a control lane and one sample lane, a concentration series should prove much helpful for quantification due to the increasing strength of the cleavage sites with increasing concentrations of lead(II)acetate. Lastly, once I have gathered all data, I will compare the observed *in vivo* metal ion binding pockets with those identified in the crystal structure of the *Oi*. group II intron (Toor et al., 2008).

7   Literature

Augustin S., Müller M.W., Schweyen R.J. (1990) Reverse self-splicing of group II intron RNAs *in vitro*. Nature 343:383-386.

Bengston S. (1994) Early life on earth Columbia University Press, New York.

Boudvillain M., Pyle A.M. (1998) Defining functional groups, core structural features and inter-domain tertiary contacts essential for group II intron self-splicing: a NAIM analysis. EMBO J. 17:7091-7104.

Boudvillain M., Delencastre A., Pyle A.M. (2000) A new RNA tertiary interaction that links active-site domains of a group II intron and anchors them at the site of catalysis. Nature 406:315-318.

Boulanger S.C., Belcher S.M., Schmidt U., Dib-Hajj S.D., Schmidt T., Perlman P.S. (1995) Studies of point mutants define three essential paired nucleotides in the domain 5 substructure of a group II intron. Mol. Cell. Biol. 15:4479-4488.

Braunlin W.H. (1995) NMR Studies of cation-binding environments on nucleic acids, Advances in Biophysical Chemistry, JAI Press Inc. pp. 89-139.

Chanfreau G., Jacquier A. (1996) An RNA conformational change between the two chemical steps of group II self-splicing. EMBO 15:3466-3476.

Cordin O., Banroques J., Tanner N.K., Linder P. (2006) The DEAD-box protein family of RNA helicases. Gene 367:17-37.

Costa M., Michel F. (1995) Frequent use of the same tertiary motif by self-folding RNAs. EMBO J. 14:1276-1285.

Costa M., Christian E.L., Michel F. (1998) Differential chemical probing of a group II self-splicing intron identifies bases involved in tertiary interactions and supports an alternative secondary structure model of domain V. Rna 4:1055-68.

Costa M., Deme E., Jacquier A., Michel F. (1997) Multiple tertiary interactions involving domain II of group II self- splicing introns. J. Mol. Biol. 267:520-36.

Daniels D.L., Michels W.J., Jr., Pyle A.M. (1996) Two competing pathways for self-splicing by group II introns: a quantitative analysis of in vitro reaction rates and products. J. Mol. Biol. 256:31-49.

Del Campo M., Tijerina P., Bhaskaran H., Mohr S., Yang Q., Jankowsky E., Russell R., Lambowitz A.M. (2007) Do DEAD-box proteins promote group II intron splicing without unwinding RNA? Mol. Cell 28:159-166.

Diges C.M., Uhlenbeck O.C. (2001) Escherichia coli DbpA is an RNA helicase that requires hairpin 92 of 23S rRNA. EMBO 20:5503-5512.

Draper D.E. (2004) A guide to ions and RNA structure. RNA 10:335-343.

Faye G., Simon M. (1983) Processing of the oxi-3-pre-messenger RNA in yeast, in: R. J. Schweyen, et al. (Eds.), Mitochondria 1983 - Nucleo-Mitochondrial Interactions, Schliersee, Germany. pp. 432-439.

Fedorova O., Pyle A.M. (2005) Linking the group II intron catalytic domains: tertiary contacts and structural features of domain 3. Embo J 24:3906-16.

Fedorova O., Zingler N. (2007) Group II introns: structure, folding and splicing mechanism. Biol Chem 388:665-78.

Fedorova O., Mitros T., Pyle A.M. (2003) Domains 2 and 3 interact to form critical elements of the group II intron active site. J. Mol. Biol. 330:197-209.

Fedorova O., Waldsich C., Pyle A.M. (2007) Group II intron folding under near-physiological conditions: collapsing to the near-native state. J. Mol. Biol. 366:1099-1114.

Fedorova O., Solem A., Pyle A.M. (2010) Protein-facilitated folding of group II intron ribozymes. J Mol Biol. 3:799-813.

Feig A.L., Uhlenbeck O.C. (1999) The role of metal ions in RNA biochemistry, in: R. F. Gesteland, et al. (Eds.), The RNA World, Cold Spring Harbor Laboratory Press, Cold Spring Harbor. pp. 287-319.

Gill M.L., Strobel S.A., Loria J.P. (2006) Cystallization and characterization of the thallium form of the *Oxytricha nova* G-quadruplex. Nucleic Acids 34:4506-4514.

Gregan J., Kolisek M., Schweyen R.J. (2001) Mitochondrial Mg(2+) homeostasis is critical for group II intron splicing in vivo. Genes Dev 15:2229-37.

Griffin E.A., Jr., Qin Z., Michels W.J., Jr., Pyle A.M. (1995) Group II intron ribozymes that cleave DNA and RNA linkages with similar efficiency, and lack contacts with substrate 2'-hydroxyl groups. Chem. Biol. 2:761-770.

Guerrier-Takada C., Gardiner K., Marsh T., Pace N., Altman S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell 35:849-57.

Halls C., Mohr S., Del Campo M., Yang Q., Jankowsky E., Lambowitz A.M. (2006) Involvement of DEAD-box Proteins in Group I and Group II Intron Splicing. Biochemical Characterization of Mss116p, ATP Hydrolysis-dependent and -independent Mechanisms, and General RNA Chaperone Activity. J. Mol. Biol. 365:835-855.

Hamill S., Pyle A.M. (2006) The receptor for branch-site docking within a group II intron active site. Mol Cell 23:831-40.

Harrison B.R., Yazgan O., Krebs J.E. (2009) Life without RNIi: noncoding RNAs and their functions in *Saccharomyces cerevisiae*. Biochem. Cell Biol. 87:767-779.

Hebbar S.K., Belcher S.M., Perlman P.S. (1992) A maturase-encoding Group IIA intron of yeast mitochondria self-splices in-vitro. NAR 20:1747-1754.

Herschlag D. (1995) RNA chaperones and the RNA folding problem. J. Biol. Chem. 270:20871-20874.

Holbrook S.R., Kim S.H. (1997) RNA crystallography. Biopolymers 44:3-21.

Huang H.R., Rowe C.E., Mohr S., Jiang Y., Lambowitz A.M., Perlman P.S. (2004) The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box protein with RNA chaperone function. Proc Natl Acad Sci U S A 102:163-8.

Huang H.R., Rowe C.E., Mohr S., Jiang Y., Lambowitz A.M., Perlman P.S. (2005) The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box protein with RNA chaperone function. Proc Natl Acad Sci U S A 102:163-8.

Jacquier A. (1990) Self-splicing Group II and nuclear pre-mRNA introns: how similar are they? TIBS 15:351-354.

Jarrell K., Dietrich R., Perlman P. (1988a) Group II Intron Domain 5 facilitates a trans-splicing reaction. Mol. Cell Biol. 8:2361-2366.

Jarrell K.A., Peebles C.L., Dietrich R.C., Romiti S.L., Perlman P.S. (1988b) Group II intron self-splicing: Alternative reaction conditions yield novel products. J. Biol. Chem. 263:3432-3439.

Jurica M.S., Moore M.J. (2003) Pre-mRNA splicing: awash in the sea of proteins. Molec. Cell 12:5-14.

Koch J.L., Boulanger S.C., Dib-Hajj S.D., Hebbar S.K., Perlman P.S. (1992) Group II Introns deleted for multiple substructures retain self-splicing activity. Mol. Cell Biol. 12:1950-1958.

Koculi E., Lee N.K., Thirumalai D., Woodson S.A. (2004) Folding of the Tetrahymena ribozyme by polyamines: importance of counterion valence and size. J Mol Biol 341:27-36.

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 31:147-157.

Lambowitz A.M., Zimmerly S. (2004) Mobile group II introns. Annu Rev Genet 38:1-35.

Lambowitz A.M., Caprara M.G., Zimmerly S., Perlman P.S. (1999) Group I and group II ribozymes as RNPs: Clues to the past and guides to the future, in: R. F. Gestland and J. F. Atkins (Eds.), The RNA world, Cold Spring Harbor Laboratory Press, Cold Spring Harbor. pp. 451-485.

Lehmann K., Schmidt U. (2003) Group II introns: structure and catalytic versatility of large natural ribozymes. Crit. Rev. Biochem. Mol. Biol. 38:249-303.

Lindell M., Romby P., Wagner G.H. (2002) Lead (II) as a probe for investigating RNA structure in vivo. RNA 8:534-541.

Lindell M., Brannvall M., Wagner E.G., Kirsebom L.A. (2005) Lead(II) cleavage analysis of RNase P RNA in vivo. Rna 11:1348-54.

Margossian S.P., Li H., Zassenhaus H.P., Butow R.A. (1996) The DExH box protein Suv3p is a component of a yeast mitochondrial 3'- to-5' exoribonuclease that suppresses group I intron toxicity. Cell 84:199-209.

Matsuura M., Noah J.W., Lambowitz A.M. (2001) Mechanism of maturase-promoted group II intron splicing. Embo J 20:7259-70.

Michel F., Ferat J.-L. (1995) Structure and activities of group II introns. Annu. Rev. Biochem. 64:435-461.

Michel F., Umesono K., Ozeki H. (1989) Comparative and functional anatomy of group II catalytic introns--a review. Gene 82:5-30.

Misra V.K., Shiman R., Draper D.E. (2002) A Thermodynamic Framework for the Magnesium-Dependent Folding of RNA. Biopolymers 69:118-136.

Mohr S., Stryker J., Lambowitz A. (2002) A DEAD-box protein functions as an ATP-dependent RNA chaperone in group I intron splicing. Cell 109:769-79.

Mörl M., Schmelzer C. (1990) Integration of group II intron bI1 into a foreign RNA by reversal of the self-splicing reaction in vitro. Cell 60:629-636.

Mörl M., Niemer I., Schmelzer C. (1992) New reactions catalyzed by a Group II intron ribozyme with RNA and DNA substrates. Cell 70:803-810.

Osiewacz H.D., Esser K. (1984) The mitochondrial plasmid of *Podospora anserina:* a mobile intron of a mitochondrial gene. Curr. Genet. 8:299-305.

Padgett R.A., Podar M., Boulanger S.C., Perlman P.S. (1994) The stereochemical course of group II intron self-splicing. Science 266:1685-1688.

Peebles C.L., Benatan E.J., Jarrell K.A., Perlman P.S. (1987) Group II intron self-splicing: development of alternative reaction conditions and identification of a predicted intermediate. Cold Spring Harbor Symp. Quant. Biol. 52:223-232.

Peebles C.L., Zhang M., Perlman P.S., Franzen J.F. (1995) Identification of a catalytically critical trinucleotide in Domain 5 of a group II intron. PNAS 92:4422-4426.

Podar M., Chu V.T., Pyle A.M., Perlman P.S. (1998) Group II intron splicing in-vivo by first step hydrolysis. Nature 391:915-918.

Pyle A.M., Lambowitz A.M. (2006) Group II introns: ribozymes that splice RNA and invade DNA., in: R. F. Gesteland, et al. (Eds.), The RNA World., Cold Spring Harbor Laborartory Press, Cold Spring Harbor, New York. pp. 469-534.

Pyle A.M., Fedorova O., Waldsich C. (2007) Folding of group II introns: a model system for large, multidomain RNAs? Trends Biochem. Sci. 32:138-145.

Qin P.Z., Pyle A.M. (1998) The architectural organization and mechanistic function of group II intron structural elements. Curr. Opin. Struct. Biol. 8:301-308.

Rajkowitsch L., Chen D., Stampfl S., Semrad K., Waldsich C., Mayer O., Jantsch M.F., Konrat R., Blasi U., Schroeder R. (2007) RNA chaperones, RNA annealers and RNA helicases. RNA Biol 4:118-30.

Robart A.R., Zimmerly S. (2005) Group II intron retroelements: function and diversity. Cytogenet Genome Res. 110:589-597.

Roitzsch M., Pyle A.M. (2009) The linear form of a group II intron catalyzes efficient autocatalytic reverse splicing, establishing a potential for mobility. RNA 15:473-482.

Römer R., Hach R. (1975) tRNA conformation and magnesium binding. A study of Yeast Phenylalanine-specific tRNA by a fluorescent indicator and differential melting curves. Eur. J. Biochem. 55:271-284.

Schmidt U., Lehmann K., Stahl U. (2002) A novel mitochondrial DEAD box protein (Mrh4) required for maintenance of mtDNA in Saccharomyces cerevisiae. FEMS Yeast Res. 2:267-276.

Schmidt U., Podar M., Stahl U., Perlman P.S. (1996) Mutations of the two-nucleotide bulge of D5 of a group II intron block splicing in vitro and in vivo: Phenotypes and suppressor mutations. RNA 2:1161-1172.

Schmidt U., Riederer B., Morl M., Schmelzer C., Stahl U. (1990) Self-splicing of the mobile group II intron of the filamentous fungus Podospora anserina (COI I1) in vitro. EMBO Journal 9:2289-2298.

Schroeder R., Barta A., Semrad K. (2004) Strategies for RNA folding and assembly. Nat. Rev. Mol. Cell. Biol. 5:908-919.

Schroeder R., Grossberger R., Pichler A., Waldsich C. (2002) RNA Folding in vivo. Curr. Opin. Struct. Biol. 12:296-300.

Seraphin B., Simon M., Boulet A., Faye G. (1989) Mitochondrial splicing requires a protein from a novel helicase family. Nature 337:84-7.

Sigel R., Vaidya A., Pyle A. (2000) Metal ion binding sites in a group II intron core. Nat Struct Biol 7:1111-6.

Silverman E., Edwalds-Gilbert G., Lin R.J. (2003) DExD/H-box proteins and their partners: helping RNA helicases unwind. Gene 312:1-16.

Solem A., Zingler N., Pyle A.M. (2006) A DEAD Protein that Activates Intron Self-Splicing without Unwinding RNA. Mol. Cell 24:611-617.

Sontheimer E.J., Gordon P.M., Piccirilli J.A. (1999) Metal ion catalysis during group II intron self-splicing: parallels with the splicosome. Genes & Development 13:1729-1741.

Stein A., Crothers D.M. (1976) Conformational changes of transfer RNA. The role of magnesium(II). Biochemistry 15:160-8.

Steitz T.A., Steitz J.A. (1993) A general two-metal-ion mechanism for catalytic RNA. Proc. Natl. Acad. Sci. USA 90:6498-6502.

Su L., Qin P., Michels W., Pyle A. (2001) Guiding Ribozyme Cleavage Through Motif Recognition: The Mechanism of Cleavage Site Selection by a Group II Intron Ribozyme. Journal of Molecular Biology 306:665-668.

Su L.J., Brenowitz M., Pyle A.M. (2003) An alternative route for the folding of large RNAs: apparent two-state folding by a group II intron ribozyme. J. Mol. Biol. 334:639-652.

Su L.J., Waldsich C., Pyle A.M. (2005) An obligate intermediate along the slow folding pathway of a group II intron ribozyme. Nucleic Acids Res 33:6674-87.

Swisher J. (2000) Folding studies of the group II intron ai5γ, Cellular, Molecular, and Biophysical Studies, Columbia University, New York. pp. 145.

Swisher J., Duarte C., Su L., Pyle A. (2001) Visualizing the solvent-inaccessible core of a group II intron ribozyme. EMBO J. 20:2051-2061.

Swisher J., Su L., Brenowitz M., Anderson V., Pyle A. (2002) Productive Folding to the Native State by a Group II Intron Ribozyme. J. Mol. Biol. 315:297-310.

Tinoco I.J., Bustamante C. (1999) How RNA folds. Journal of Molecular Biology 293:271-281.

Toor N., Hausner G., Zimmerly S. (2001) Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. RNA 7:1142-1152.

Toor N., Keating K.S., Taylor S.D., Pyle A.M. (2008) Crystal structure of a self-spliced group II intron. Science 320:77-82.

Treiber D.K., Williamson J.R. (1999) Exposing the kinetic traps in RNA folding. Curr. Opin. Struct. Biol. 9:339-345.

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. Nucleic Acids:201-227.

Valadkhan S., Manley J.L. (2001) Splicing-related catalysis by protein-free snRNAs. Nature 413:701-7.

van der Veen R., Kwakman J.H.J.M., Grivell L.A. (1987) Mutations at the lariat acceptor site allow self-splicing of a Group II intron without lariat formation. EMBO Journal 6:3827-3831.

van der Veen R., Arnberg A.C., van der Horst G., Bonen L., Tabak H.F., Grivell L.A. (1986) Excised Group II introns in yeast mitochondria are lariats and can be formed by self-splicing *in vitro*. Cell 44:225-234.

Vogel J., Borner T. (2002) Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing. EMBO J. 21:3794-3803.

Waldsich C., Pyle A.M. (2007) A folding control element for tertiary collapse of a group II intron ribozyme. Nat. Struct. Mol. Biol. 14:37-44.

Waldsich C., Pyle A.M. (2008) A kinetic intermediate that regulates proper folding of a group II intron RNA. J. Mol. Biol. 375:572-580.

Waldsich C., Grossberger R., Schroeder R. (2002) RNA chaperone StpA loosens interactions of the tertiary structure in the *td* group I intron in vivo. Genes & Dev. 16:2300-2312.

Watanabe K., Lambowitz A.M. (2004) High-affinity binding site for a group II intron-encoded reverse transcriptase/maturase within a stem-loop structure in the intron RNA. Rna 10:1433-43.

Weeks K.M. (1997) Protein-facilitated RNA folding. Curr. Opin. Struct. Biol. 7:336-342.

Woodson S.A. (2005a) Metal ions and RNA folding: a highly charged topic with a dynamic future. Curr. Opin. Chem. Biol. 9:104-109.

Woodson S.A. (2005b) Structure and assembly of group I introns. Curr Opin Struct Biol.

Xiang Q., Qin P.Z., Michels W.J., Freeland K., Pyle A.M. (1998) The sequence-specificity of a group II intron ribozyme: Multiple mechanisms for promoting unusually high discrimination against mismatched targets. Biochemistry 37:3839-3849.

Zingler N., Solem A., Pyle A.M. (2010) Dual roles for the Mss116 cofactor during splicing of the ai5{gamma} group II intron. Nucleic Acids:1-8.

8   Anhang

8.1   Zusammenfassung der vorliegenden Arbeit

Das richtige Falten von RNA-Molekülen *in vivo* ist eine der Grundvoraussetzungen für das Funktionieren einer jeden Zelle. Dabei muss jedes RNA-Molekül eine definierte, dreidimensionale Struktur einnehmen, wobei oft Regionen, die in der Basenabfolge weit entfernt von einander sind, nahe zusammen kommen. Da das Rückgrat eines RNA-Moleküls viele negativgeladene Phosphorgruppen enthält, entsteht bei räumlicher Annäherung dieser Gruppen elektrostatische Abstoßung. Um dennoch den nativen, gefalteten Zustand erreichen zu können, spielen mono- und divalente Metallionen eine große Rolle. Sie können einerseits diskrete Metallionen-Bindestellen innerhalb der RNA ausfüllen oder als delokalisierte Ionen negative Ladungen kompensieren. Da in der Zelle meist sehr geringe Konzentrationen von Metallionen vorliegen, wird deren Rolle oft durch Co-Faktoren wie Proteine übernommen. Die vorliegende Arbeit beschäftigt sich mit dem mitochondrialen Gruppe II Intron ai5γ aus der Bäckerhefe *Saccharomyces cerevisiae,* das *in vitro* bereits sehr gut untersucht ist. So ist etwa bekannt, dass das Falten von ai5γ hohe Konzentrationen von divalenten Metallionen benötigt bzw. das Protein Mss116p, das als Co-Faktor für dieses Gruppe II Intron identifiziert worden ist. Weiteres sind die Metallionen-Bindestellen in diesem Intron *in vitro* bekannt. Das Ziel dieser Arbeit ist das Gewinnen von neuen Einblicken in den Vorgang der RNA-Faltung von Gruppe II Introns und der Einfluss von Metallionen-Bindestellen auf die Faltung von ai5γ *in vivo*. Zu diesem Zweck wurde von mir eine neue Methode etabliert, mit der es mir möglich ist Metallionen-Bindestellen *in vivo* zu detektieren. Die Methode basiert auf Bleiacetat oder Terbiumchlorid, die beide in der Lage sind Magnesiumionen in ihrer Bindestelle zu ersetzen und weiteres das Rückgrat der RNA an dieser Stelle zu schneiden. Diese Schnitte werden dann durch eine reverse Transkription detektiert.

Die Ergebnisse der vorliegenden Arbeit haben gezeigt, dass das Detektieren von Metallionen-Bindestellen sehr stark davon abhängt, ob die ai5γ RNA-Moleküle effizient gespleisst sind oder nicht. Daher wurden die Spleisswerte der einzelnen RNA-Proben während der gesamten Dauer der vorliegenden Arbeit kontrolliert. Unerklärlicherweise hat der eingesetzte Hefestamm gegen Ende der vorliegenden Arbeit einen rapiden Abfall bei den Spleisswerten gezeigt und ist wahrscheinlich auf ein Problem bei der Intronfaltung zurückzuführen. Ein geringerer Grad an gespleissten RNA-Molekülen ist gleichbedeutend mit dem Verlust aller Metallionen-Bindestellen im Molekül, zumindest waren diese nicht mehr zu detektieren. Im Gegenzug dazu gehen hohe Spleisswerte einher mit der Detektierbarkeit von vielen Metallionen-Bindestellen. Dieser reziproke Zusammenhang kann an Hand der ausgewerteten Daten sehr klar nachvollzogen werden.

Die Positionen der Metallionen-Bindestellen wurde dann mit den bereits bekannten Bindestellen *in vitro* verglichen, um daraus Schlüsse über deren Einfluss auf die RNA-Faltung ziehen zu können. Meine Ergebnisse müssen auf Grund von unterschiedlichen Spleisswerten differenziert betrachtet werden; einerseits wurde eine sehr starke Abnahme von Metallionen-Bindestellen festgestellt (in den Proben,

die einen geringen Anteil an korrekt gespleissten ai5γ aufwiesen), andererseits haben die Experimente mit Proben welche gute Spleisswerte aufwiesen eine sehr gute Korrelation mit den bereits bekannten Metallionen-Bindestellen von Sigel et al. (2000) aufgezeigt.

Daraus können folgende Schlüsse gezogen werden: das Fehlen von bestimmten Metallionen-Bindestellen deutet darauf hin, dass der Co-Faktor, das Protein Mss116p, die Rolle der Metallionen übernimmt und somit ausschlaggebend für die Stabilisierung des Introns im nativen Zustand ist.

Der zweite Schluss, basierend auf der Annahme, dass sich die Metallionen-Bindestellen *in vivo* nicht sehr stark unterscheiden von denen *in vitro*, deutet auf eine etwas andere Wirkungsweise von Mss116p hin. Es ist bekannt, dass der Faltungsweg von ai5γ zweiteilig ist, beginnend mit dem langsamen Kollaps von Domäne 1 und gefolgt vom schnellen Übergang zum nativen Zustand. Der Kollaps von Domäne 1 wird durch das Folding-control element gesteuert und dieses Element benötigt eine hohe Magnesium-konzentration. Basierend auf meinen Daten schlage ich einen Mechanismus vor, bei dem Mss116p das „folding control element" stabilisiert bzw. in der richtigen Konformation gefangen hält und dadurch letztendlich das Erreichen des nativen spleiß-kompetenten Zustandes ermöglicht. Um diese Hypothesen untersuchen zu können, muss erstens noch eine Optimierung der Methode erfolgen und zweitens wer-den *in vitro* Experimente mit Mss116p benötigt.

8.2 Summary of the presented work

Correct folding of RNA molecules *in vivo* is one prerequisite for sustaining life. During the folding process a RNA molecule must adopt a well-defined, three-dimensional structure by bringing together regions which are far apart in sequence. This cannot happen until the RNA backbone undergoes twists and bending which brings the negative charges of the phosphate groups of the RNA backbone in close proximity generating repulsive electrostatical forces. In order to reach the native state the RNA molecule must overcome these forces; e. g. with the help of metal ions which play a very important role for folding *in vivo.* Mono- and divalent metal ions can either fill out discrete binding pockets or neutralize the negative charges of the RNA backbone forming an "ion atmosphere" consisting of delocalized ions which surround the molecule. One problem encountered during RNA folding *in vivo* are low concentrations for metal ions inside the cell forcing nature to partly replace metal ions with protein co-factors. The thesis presented here focus on the group II intron ai5γ from *Saccharomyces cerevisiae* which is very well studied *in vitro*. For example it is known that folding of ai5γ is dependent on high salt concentration and temperature as well as the protein co-factor Mss116p. Furthermore metal ion binding sites within the intron are also known. The goal of this thesis is to get new insights into RNA folding *in vivo* as well as to study the influence of metal ion binding pockets on folding of ai5γ. In order to address this topical question I had to establish a new method which allows me to detect metal ion binding pockets within a RNA molecule *in vivo*. The method is based on either lead(II)acetate or terbium(III)chloride which are able to displace magnesium ions in their binding pockets and induce cleavage of the RNA backbone at these sites. This cleavage sites can then be detected by a reverse transcription and are superimposed on the secondary structure map of the ai5γ intron.

The results of this work did show that the ability to detect metal ion binding sites is strongly dependent on splicing rates of ai5γ. Therefore splicing rates were monitored throughout the whole working period, but inexplicably the yeast strain used did show a severe decrease in splicing rates towards the end. The reason for this might be a problem of intron folding or other sources unknown. The decrease of ai5γ splicing efficiency has a severe consequence for the work presented here because most of the metal ion binding pockets cannot be detected anymore. It therefore seems that splicing efficiency and the ability to detect metal ion binding pockets are connected in a reciprocal manner and can be seen best when looking at the results of this work.

The detected metal ion binding sites *in vivo* were then compared to the metal ion binding sites already known *in vitro* and blotted on a secondary structure map of the ai5γ intron. This allowed me to draw conclusions about the influence metal ion binding pockets have on folding of ai5γ. Due to the different splicing efficiency my results must be interpreted with care; on the one hand experiments with a low percentage of efficiently spliced ai5γ did show a tremendous decrease of metal ion binding sites *in vivo* compared to those already known *in vitro* (Sigel et al., 2000), on the other hand when working with samples containing a high percentage of efficiently spliced ai5γ I could show that metal ion binding sites correlate very well with those observed by Sigel et al. (2000). This leads me to different conclusions: first, the absence of metal ion binding pockets *in vivo* suggests that the protein Mss116p compensates

in vivo for the metal ions and is the primary source conferring stability to the intron in the native state. Secondly, assuming that there is only minor difference between the position of metal ion binding sites *in vivo* and *in vitro*, this suggests a different mode of action for Mss116p. It is known that the folding pathway of ai5γ can be subdivided into two steps, starting with the slow collapse of domain 1 and followed by the quick transition to the native state. Step one, the collapse, is triggered by a small structural element termed folding control element and requires high concentration of magnesium ions in order to do so. Based on my data I suggest a mechanism for Mss116p stabilizing the folding control element and therefore retaining it in the correct conformation allowing ai5γ to end its transition to the native state. For verifying this hypothesis optimization of the method as well as *in vitro* experiments with Mss116p are needed.

Michael Wildauer

Persönliche Daten

Geburtsdatum: 13. Juni 1984
Geburtsort: Linz, Oberösterreich
Staatsbürgerschaft: Österreichisch
Familienstand: Partnerschaft

Ausbildung:

Oktober 2004 -   Studium der Molekularen Biologie an der Universität Wien
                          Diplomarbeit im Bereich Biochemie (Labor von Dr. Christina Waldsich)

Jänner 2008   Abschluss des 1. Abschnitt im Studium Molekulare Biologie

Sept. 2003 – April 2004 Präsenzdienst

Juni 2003   Matura am Khevenhüllergymnasium, 4020 Linz

Praktika und Berufserfahrung

29. März – 1. April 2010 RNA Structure and Function – Theoretischer Kurs, organisiert von IC-GEB in Triest, Italien

März – April 2009   Wahlbeispiel (Mapping metal ion binding sites *in vivo*)
                           Department für Biochemie, Dr. Christina Waldsich, Universität Wien

Nov. 2008, Feb. 2009   Wahlbeispiel (Arp2/3 complex in fish fibroblasts and its role in actin filament synthesis)
                           IMBA – Institute of Molecular Biotechnology, Vic Small
                           Wien

August 2008   Praktikum bei Boehringer Ingelheim RCV im Labor von Dr. Walter Spevak

März - April 2008   Wahlbeispiel (Clp-proteins and their role in *Streptococcus pyogenes*)
                           Department für Mikrobiologie, Immunologie und Genetik, Prof. Emmanuelle Charpentier, Universität Wien

August-September 2007   Praktikum bei Boehringer Ingelheim RCV im Labor von Dr. Walter Spevak

August-September 2006   Praktikum bei Boehringer Ingelheim RCV im Labor von Dr. Walter Spevak