The Dissertation Committee for Christopher Allan Myers
Certifies that this is the approved version of the following dissertation:

Interaction of the *Neurospora crassa* mitochondrial tyrosyl-tRNA synthetase with group I intron RNAs

Committee:

_____
Alan M. Lambowitz, Supervisor

_____
Ellen Gottlieb

_____
David L. Herrin

_____
Kenneth A. Johnson

_____
Robert M. Krug

Interaction of the *Neurospora crassa* mitochondrial tyrosyl-tRNA synthetase with group I intron RNAs

by

Christopher Allan Myers, B.S., M.S.

Dissertation

Presented to the Faculty of the Graduate School of

the University of Texas at Austin

In Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May, 2002

Dedication

To my parents,

Lloyd Leon Myers

and

Lynda Jane Fraley

Acknowledgements

First and foremost, I thank my supervisor, Dr. Alan M. Lambowitz, for guidance, encouragement and support throughout my graduate studies. I would also like to thank Drs. Mark Caprara and Georg Mohr for serving as mentors during my early graduate career, and for advice and discussion throughout. Their guidance, help and encouragement were invaluable.

I thank Dr. Robert Fox for the Fe-EPD used in this study. I thank Dr. Stephen Cusack, Mikhail Tukalo and Anya Yaremchuk for providing the crystal structure of the *T. thermophilus* TyrRS prior to publication. I am also grateful to the members of my dissertation committee, Drs. Ellen Gottlieb, David Herrin, Kenneth Johnson, and Robert Krug, for their time, guidance and effort.

I thank those who helped out or assisted me throughout my research: Tony Kaiser, Birte Kuhla and Jim Williams, and especially Stacy Crain for making things so enjoyable near the end. Additionally, I thank all the members of the Lambowitz Lab, past and present, for advice and discussion. Special thanks go to Drs. Sabine Mohr and Roland Saldanha in this respect.

Thanks to mac, laura, jon and jim who have been here throughout and have matured as much as I (hopefully) have.

I thank my dear friend, Dr. Torsten Moeller, who is not only a great scientist and traveling companion, but also someone we should all aspire to emulate.

Finally, I would like to thank my parents, Lloyd Leon Myers, and Lynda Jane Fraley, for their support and patience throughout the entire length of my education.

NOTE:

Because of the University of Texas at Austin thesis requirements, most figures contained here are too small to be useful. I suggest checking the original references.

Interaction of the *Neurospora crassa* mitochondrial tyrosyl-tRNA synthetase with group I intron RNAs

Publication No._____

Christopher Allan Myers, Ph. D.

The University of Texas at Austin, 2002

Supervisor: Alan M. Lambowitz

The *Neurospora crassa* mitochondrial tyrosyl-tRNA synthetase, or CYT-18 protein, functions in splicing group I introns by promoting the formation of the catalytically active structure of the intron RNA. To investigate how CYT-18 stabilizes the active RNA structure, I used an *Escherichia coli* genetic assay with the phage T4 *td* intron to systematically test the ability of the CYT-18 protein to compensate for structural defects in the P7 region of the group I intron core. P7 is a long-range base-pairing interaction of the P3-P9 domain that forms the binding site for the splicing co-factor guanosine. My results show that CYT-18 can suppress numerous mutations that impair the self-splicing of the *td* intron, including mutations that disrupt base-pairing within the P7 region. CYT-18 suppressed mutations of

phylogenetically conserved nucleotide residues at all positions tested, except for the universally conserved G-residue at the guanosine-binding site. Structure mapping experiments with some selected mutant introns showed that the P7 mutations impaired the formation of both P7 and P3, thereby grossly disrupting the P3-P9 domain.

Previous work suggested that CYT-18 recognizes a conserved tRNA-like structure of the group I intron catalytic core. I used directed hydroxyl radical cleavage assays to show that the nucleotide-binding fold and C-terminal domains of CYT-18 interact with the expected group I intron cognates of the aminoacyl-acceptor stem and D-anticodon arms, respectively. Further, three-dimensional graphic modeling, supported by biochemical data, shows that conserved regions of group I introns can be superimposed over interacting regions of the tRNA in a *Thermus thermophilus* TyrRS/$tRNA^{Tyr}$ cocrystal structure.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

My research has focused on the *Neurospora crassa* mitochondrial tyrosyl-tRNA synthetase (mt TyrRS: CYT-18 protein) and how it interacts with its RNA substrates. Interestingly, the CYT-18 protein functions not only in aminoacylation (tRNA charging), but also in the splicing of group I introns (Akins and Lambowitz, 1987). The dual function of the CYT-18 protein obligates the protein to two types of RNA substrates: tRNAs and group I introns. My research goals have been to study how CYT-18 functions in promoting the splicing of group I introns and to examine the structural similarities between CYT-18's two different types of RNA substrates. How such structural similarities may have evolved is an interesting aspect of this work. In this section, I will briefly introduce group I introns, known splicing factors and aminoacyl tRNA synthetases.

1.1 GROUP I INTRONS

In 1981, Dr. Thomas Cech and coworkers showed that the intervening sequence in the large ribosomal subunit (LSU) from *Tetrahymena thermophila* spliced itself from an RNA transcript in the absence of proteins, requiring only a monovalent salt, a divalent cation and guanosine (Cech et al., 1981). This was the first demonstration of a molecule with genetic and catalytic properties, and contributed to the revitalization of the RNA world hypothesis.

Group I intron splicing involves two sequential trans-esterification reactions (Figure 1.1). In the first reaction, the 3'OH group of an exogenous guanosine molecule initiates nucleophilic attack at the 5' splice site, and becomes covalently attached to the 5' end of the intron. In the second step, the 3'OH group of the 5' exon attacks at the 3' splice site, yielding ligated exons, and the liberated intron with the non-coded guanosine at its 5' end. It is important to note that RNA splicing is not driven by GTP hydrolysis. Although many group I introns are self-splicing *in vitro*, it is believed most, if not all, rely on protein factors for efficient *in vivo* splicing (Lambowitz and Perlman, 1990; Lambowitz et al., 1999).

Since their discovery, nearly 1500 group I introns have been identified in bacterial and eukaryotic organisms and their organelles (Cannone et al., 2002). Group I introns share little sequence homology, however they do share a common core secondary structure (see. Figure 1.2) that consists of base-paired regions and

loops numbered 5' to 3' and intervening joining regions that connect specific structures (*e.g.* J3/4, which connects P3 and P4; Burke et al., 1987). The secondary structure model is based on sequence co-variation analysis, mutational analysis, and chemical probing experiments. Several highly conserved elements of the group I intron secondary structure make up the 'catalytic core' of the ribozyme, specifically the structures P3, J3/4, P4, J4/5, P5, P6, P6a, J6/7, P7 and P9. Many group I introns have insertions peripheral to the catalytic core, which can serve as structural stabilization elements, or contain open-reading frames for various proteins (see below). The tertiary structure of the group I intron catalytic core is also conserved among group I introns and a three-dimensional model for the core of the group I intron was proposed by Michel and Westhof in 1990, based upon sequence co-variation analysis, mutational analysis, and chemical probing experiments. In this model, the catalytic core of the group I intron consists of two extended helical domains, the P4-P6 structural domain, and the P3-P9 catalytic domain (see Figure 1.3). Juxtaposition of the two domains forms the active site cleft, where the substrate helices, P1 and P10, and the guanosine cofactor bind. Tertiary interactions between the two domains help stabilize their interaction (Michel and Westhof, 1990). In 1998, Golden et al. published a 5 Å crystal structure of the core of the Tetrahymena intron. Although not at high enough resolution to see atomic-level interactions, the structure did validate Michel and Westhof's model, and showed that the active site of the ribozyme is largely pre-formed.

The P4-P6 domain is formed by the coaxial stacking of elements P4, P5, P6 and P6a, (Michel and Westhof, 1990). Base triples between P4 and J6/7 and P6 and J3/4 help stabilize and orient this coaxial stacking (Michel and Westhof, 1990). In the case of the *Tetrahymena* intron, folding of the P4-P6 domain is independent of the P7-P9 domain and a high resolution structure of the isolated *Tetrahymena* P4-P6 domain has been determined by X-ray crystallography (Murphy and Cech, 1993; Cate et al., 1996).

The P3-P9 domain consists of elements P3, P7, P8, J8/7, and P9. As seen in the crystal structure of the group I intron core, the P3-P9 domain wraps around the P4-P6 domain forming the active site cleft (Golden et al., 1998). There is evidence that the P3-P9 is the catalytic domain of the group I intron, or at least contains many of the residues required for the catalytic activity (Ikawa et al., 2000). Additionally, a nearly universal conserved base-pair within P7 has been shown to bind the guanosine cofactor that initiates intron catalysis (Michel et al., 1989). In contrast to the hairpin-like P4-P6 domain, the P3-P9 domain has two long-range base-pairing interactions, P3 and P7, which impact its rate of folding (see below, Figure 1.2).

1.2 FOLDING OF THE GROUP I INTRON

Like proteins, group I introns and other large RNAs face the problem of correctly folding into a catalytically active structure. Proper secondary and tertiary interactions must be established, often in the presence of equally stable non-active interactions. The *Tetrahymena* intron has been a useful model system for studying

the folding of large RNAs. Studies have shown that the folding of the *Tetrahymena* intron into a compact globular structure requires the presence of magnesium or other cations that negate repulsion forces from the close packing of the negatively charged phosphate backbone (Celander and Cech, 1991; Thirumalai et al., 2001). Counter-ions promote the first step in RNA folding, a collapse of the extended RNA polymer into a more compact form (reviewed in Thirumalai et al., 2001). This 'collapse' partitions the RNA molecules into two populations: 1) those that rapidly achieve the native state, and 2) those that form stable non-native structures. These non-native states, or 'kinetic traps', must cross an activation energy barrier in which the RNA becomes partially or completely unfolded before achieving the active state. In the specific case of the *Tetrahymena* intron, chemical structure probing and oligonucleotide-accessibility assays, have shown that the P4-P6 domain folds first, followed by formation of the P3-P9 domain, and finally, formation of specific tertiary interactions that stabilize the association of the two domains (Murphy and Cech, 1993; Zarrinkar and Williamson, 1996; Ralston et al., 2000). Time-resolved Fe(II)-EDTA cleavage experiments revealed that association of the two domains is the rate-limiting step for formation of the active state, occurring on the minute time scale (Thirumalai et al., 2001). It is likely that all group I introns follow a similar folding pathway, since the P4-P6 domain is mainly comprised of local hairpins structures that can form in the absence of magnesium, and is fully transcribed by RNA polymerase prior to completion of the entire intron. Additionally, the P3-P9 domain forms on a

longer time-scale, reflecting, the long-range nature of elements P3 and P7 (Zarrinkar and Williamson, 1996, Sclavi et al., 1998).

1.3 GROUP I INTRON SPLICING FACTORS

Although many group I introns are self-splicing *in vivo*, most, if not all, are believed to require protein factors *in vivo* (Lambowitz and Perlman, 1990; Lambowitz et al., 1999). In a manner similar to ribosomal proteins, group I splicing factors are expected to structurally stabilize the intron RNA, while the chemical reactions remain RNA-catalyzed. To date, numerous group I intron splicing factors have been identified and can be divided into two classes: 1; intron-encoded splicing factors (‘maturases’) and 2; host-encoded RNA binding proteins (both reviewed in Lambowitz and Perlman, 1990; Lambowitz et al., 1999).

Many group I introns contain open reading frames whose protein products promote the splicing of the group I intron that encodes them. These ‘maturases’ are related to DNA endonucleases with characteristic LAGLI-DADG amino-acid motifs. The endonucleases function in intron mobility by causing a double-stranded break in an intronless allele. A double-stranded break repair mechanism leads to 'intron homing', where the intron and flanking sequences are transferred into the intronless genome (Dujon, 1989). Additionally, maturases may help group I introns reverse splice into unrelated mRNAs, which could then be reverse transcribed, and the resulting cDNA genomically integrated (Cech, 1985; Mohr and Lambowitz, 1991; Roman and Woodson, 1998). One possible mechanism of maturase function in

splicing could be by recognizing and properly docking the P1/P10 substrate helix of the group I intron (Lambowitz et al., 1999). The nucleotide sequence of the substrate helix is identical to the DNA target site of the endonuclease, and thus may be recognized by the same active site, as suggested for the yeast aI4α protein (Wernette et al., 1990; Henke et al., 1995).

Host-encoded intron splicing factors tend to be pre-existing RNA-binding proteins that have adapted to function in group I intron splicing. Two of these splicing factors, the *cyt-18* and *NAM2* gene products, are mitochondrial aminoacyl tRNA synthetases (mt aaRSs), which may use their tRNA binding domains to promote group I intron splicing (Lambowitz and Perlman, 1990; Houman et al., 2000). One apparent exception to this trend is the yeast CBP2 protein, which promotes the splicing of the yeast bI5 intron (Gampel et al., 1989). CBP2 has no appreciable homology with other proteins, suggesting that it is a unique splicing factor, which may have evolved by gene duplication (Hill et al., 1985; Li et al., 1996).

The best characterized host-encoded group I intron splicing factor is the *Neuropsora crassa* mitochondrial tyrosyl-tRNA synthetase, or CYT-18 protein, which promotes the splicing of 3 of the 10 group I introns in *Neurospora* mitochondria (Figure 1.4, Akins and Lambowitz, 1987). Although homologous to well-studied bacterial TyrRSs, only the TyrRS from *Neurospora* and that from the closely related fungus *Podospora anserina* promote group I intron splicing, while the yeast and bacterial TyrRSs do not. Remarkably, in addition to its endogenous

substrates, the CYT-18 protein promotes the splicing of a variety of group I introns from diverse sources, such as those found in yeast or bacteriophage T4 (Guo and Lambowitz, 1992). The splicing of exogenous substrates by CYT-18 reflects its ability to recognize conserved structural features of the group I intron catalytic core, rather than specific sequence elements. As CYT-18 is the focus of my research, it will be discussed in much greater detail later.

The yeast mt leucyl-tRNA synthetase, also functions in group I intron splicing (Herbert et al., 1988; Houman et al., 2000). To date, the yeast mt LeuRS protein has only been shown to promote the splicing of two closely related group I introns from yeast mitochondria. Additionally, the yeast mt LeuRS requires a maturase as a protein co-factor for splicing (Labouesse, 1990; Li et al., 1996). Two- and three-hybrid system experiments with the yeast mt LeuRS and the maturase suggest that a ternary complex is formed between the two proteins and the intron RNA (Rho and Martinis, 2000). Recently, it has been shown that both a prokaryote and a human LeuRS can complement a yeast strain lacking *NAM2,* suggesting that splicing activity comes from widely conserved features of LeuRSs (Houman et al., 2000).

The yeast CPB2 protein promotes the splicing of the bI5 group I intron at low (5 mM) magnesium levels, a condition where self-splicing is impaired (Gampel et al., 1989). CBP2 does not promote the formation of the catalytically active group I intron core as CYT-18 does, but rather 'traps' the RNA once it folds into the active form

(Weeks and Cech, 1995;1995;1996). Additionally, CBP2 helps associate the substrate helix P1 with the core of the intron (Weeks and Cech, 1995;1996).

1.4 AMINOACYL-tRNA SYNTHETASES

AaRSs catalyze the aminoacylation of their cognate tRNAs by two sequential reactions. First, the amino acid is activated with the nucleotide ATP to form an enzyme-bound aminoacyl intermediate (adenylation). The activated amino acid is subsequently transferred to the 3' end of the cognate tRNA's acceptor arm (charging). The fidelity of aaRSs is crucial for accurate protein synthesis and aaRSs have evolved mechanisms to prevent mischarging of non-cognate tRNAs. These mechanisms of structure recognition include identity, in which specific elements must be present for charging, as well as elements that block non-cognate interactions, and editing activities to correct charging errors.

AaRSs are believed to be among the most ancient proteins extant, as their activity would be required for an transition from an 'RNA world' to and 'RNA-protein world' (Schimmel and Ribas De Pouplana, 2000). The venerable nature of aaRSs has caused two interesting evolutionary phenomena. First, aaRSs have had ample time to develop mechanisms which improve their substrate discrimination efficiency, as well as acquiring additional, sometimes unrelated, functionalities (see below). Secondly, numerous non-synthetase proteins have synthetase-like domains, suggesting early gene duplication events in which modular domains of aaRSs were

acquired by proteins and then adapted for new functions (Schimmel and Ribas De Pouplana, 2000).

AaRSs are divided into two classes, of 10 synthetases each, based on structural features, conserved sequences and attachment position of the amino acid. As my research focuses on the *N. crassa* mt TyrRS, a class I synthetase, I will briefly introduce class I synthetases.

1.5 CLASS I AMINOACYL tRNA SYNTHETASES

Class I synthetases attach amino acids to the 2'-OH group of their cognate tRNA. Class I includes the synthetases for the aliphatic and sulfur containing amino acids (Lue, Ile, Val, Cys and Met; Class Ia), monomeric synthetases for charged amino acids (Glu, Gln and Arg; Class Ib) and some aromatic amino acids (Trp and Tyr; Class Ic). Although subtle differences exist between the subclasses, all Class I synthetases are defined by a characteristic N-terminal domain which contains the active site for adenylation. This active site includes the Rossman nucleotide binding-fold, an alternating pattern of parallel β-strands and overlaying α-helices. The active site also contains two class I signature sequences, the HIGH and the KMSKS motifs, which interact with the ATP molecule (reviewed in Arnez and Moras, 1997).

The N-terminal domain of the class I synthetase is believed to be the more ancient portion of the protein since it alone is often sufficient to catalyze the synthetase reactions (Schimmel and Ribas De Pouplana, 2000). Charging of an

acceptor arm mini-helix is frequently possible with an isolated N-terminal domain of a class I synthetase (reviewed in Ribas de Pouplana and Schimmel, 2001).

Individual class I synthetases often contain insertions in the N-terminal domain which provide functional attenuation or additional features. One common insertion, connective peptide 1 (CP1), has been shown to confer species-specific acceptor arm recognition (Auld and Schimmel, 1995; Wakasugi et al., 1998). In addition, this insertion sometimes functions in tRNA editing, as seen for the *Thermus. thermophilus* ValRS (Lin et al., 1996).

In contrast to the N-terminal domain, the C-terminal domains of class I synthetases are structurally diverse, although they are functionally homologous. No common motifs are seen, although all C-terminal domains function in RNA binding, specifically by interacting with the D-anticodon stacked helical region of the substrate tRNA (Auld and Schimmel, 1996). This interaction is not always crucial for synthetase activity, as some C-terminal deletion mutants retain synthetase activity (see above). As yet, an isolated C-terminal domain binding to an isolated D-anticodon mini-helix has not been demonstrated. However, some synthetases do require the C-terminus for activity and removal often causes a reduction in the kinetic rate constants for the enzyme. The C-terminal domains of class I aaRSs often perform functions in addition to RNA binding. For example, the C-terminal domain of the mammalian TyrRS has cytokine activity following proteolytic processing (Wakasugi et al., 1998).

Tyrosyl-tRNA synthetases, such as the CYT-18 protein, have the highly conserved class I N-terminal domain described above. Additionally, four conserved regions, or clusters, of amino acid residues implicated in tRNA binding have been identified in TyrRS, two in each of the N-terminal and C-terminal domains (Bedouelle, 1990; Nair et al., 1997). The N-terminal clusters interact with the acceptor arm of the tRNA while the two clusters of amino acid residues in the C-terminal domain are involved in binding the D-anticodon arm of the tRNA (Bedouelle, 1990).

1.6 OVERVIEW OF THE DISSERTATION RESEARCH

This research in this dissertation has focused on three closely related subjects: 1) Examining the role of CYT-18 in promoting the splicing of group I introns; 2) Identification of regions of contact between the CYT-18 protein and its intron RNA substrates; and 3) Testing the hypothesis that CYT-18 has adapted to function in splicing by recognizing a conserved tRNA-like structure within the core of the group I intron.

This dissertation is written in four chapters. Following this introductory chapter, I will describe the materials and methods used for the dissertation research. In Chapter 3, I will describe research on the role of CYT-18 in promoting the formation of the catalytically active RNA structure required for group I intron splicing. The major result of this research is demonstrating CYT-18's ability to suppress structural mutations that impair the self-splicing of the bacteriophage T4 *td*

intron. Chapter 4 describes a directed hydroxyl-radical cleavage approach to define regions of interaction between the CYT-18 protein and its group I intron substrates. By substituting cysteine residues for amino acids in RNA binding regions, portions of the protein could be localized to specific regions within the intron RNA substrates. Furthermore, in collaboration with Dr. Stephen Cusack and coworkers of the EMBL Outstation in Grenoble, France, molecular modeling revealed that the intron RNAs likely have a tRNA-like disposition when bound to the CYT-18 protein.

Figure 1.1: Group I intron splicing mechanism and structure.

Outline of the two sequential transesterification reactions of group I intron splicing, initiated by an exogenous guanosine.

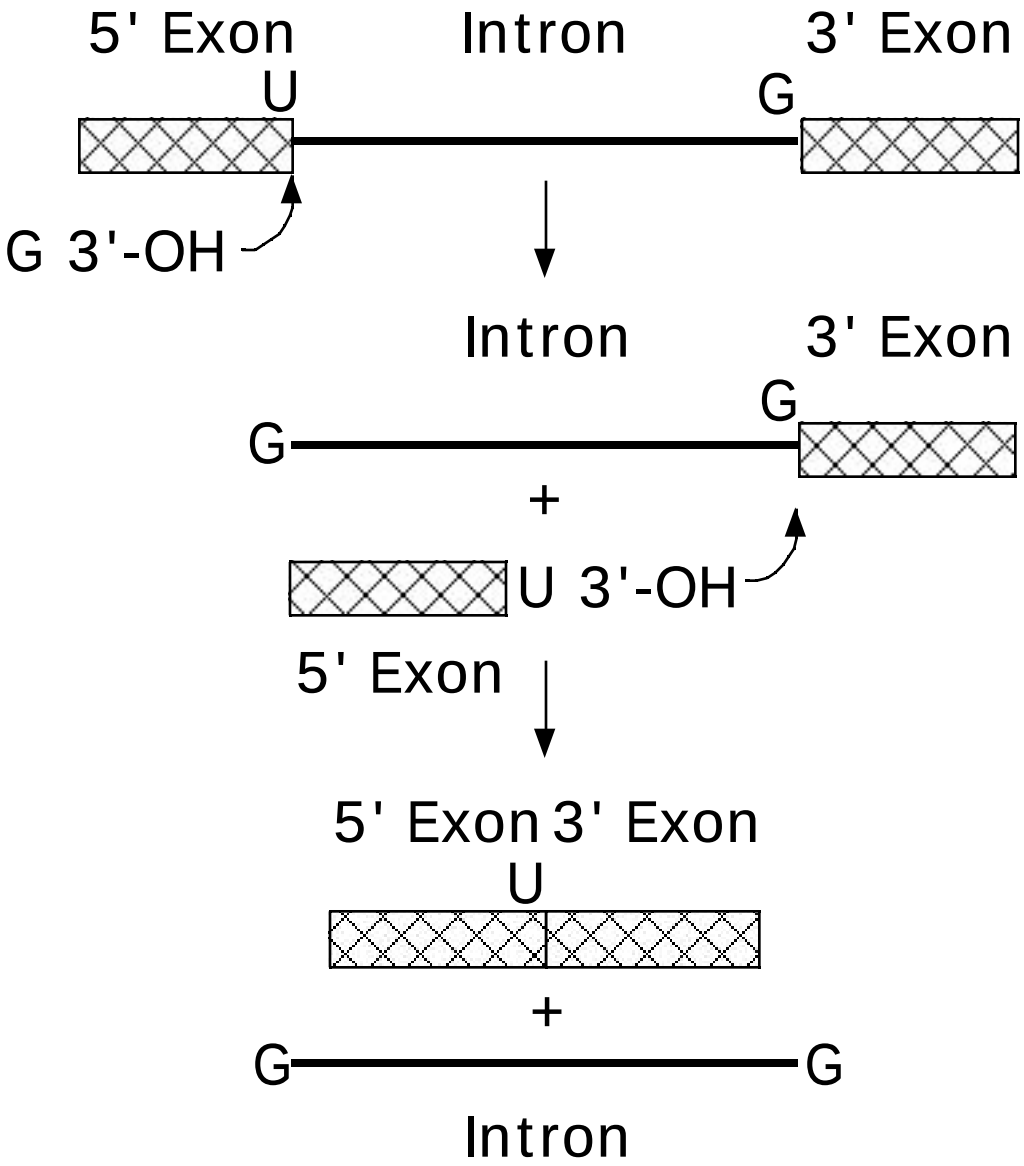


Figure 1.2: Predicted secondary structures of group I introns.

Figure 1.2A and B show the predicted secondary structure of the *N. crassa ND1* and the *Tetrahymena* LSU introns respectively. The introns are drawn in the format of Cech et al., 1994), and numbered according to Belfort et al. (1987). Intron pairing regions are numbered according to Burke et al. (1987). The P4-P6 domain is shown in green and the P3-P9 domain is shown in blue.

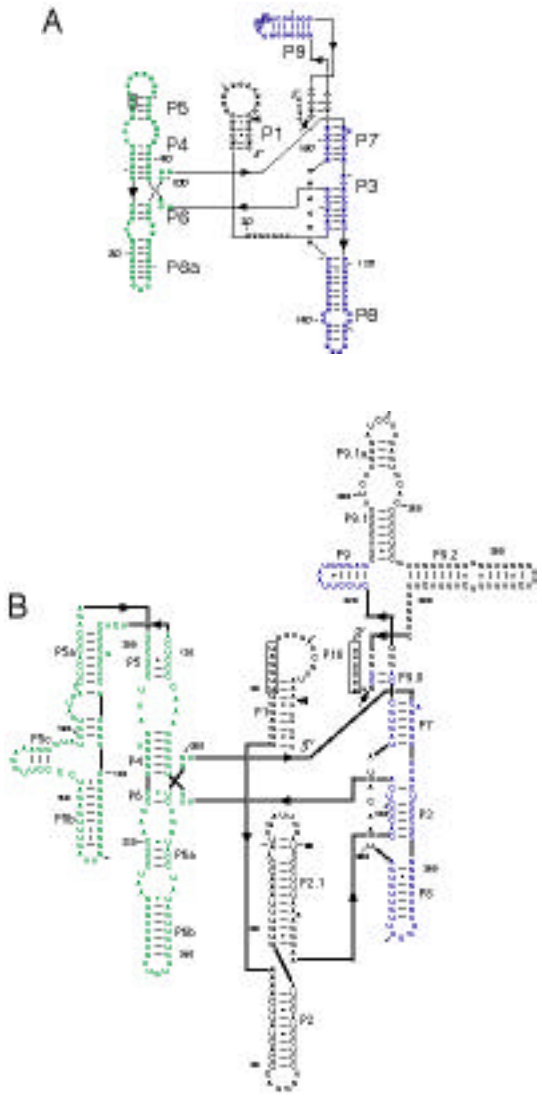


Figure 1.3: Tertiary structure models of group I introns.

Tertiary structure models of the *N. crassa ND1* intron (A; Caprara et al., 1996b) and the *Tetrahymena* LSU intron (B; Golden et al., 1998) are shown. The P4-P6 domain is shown in green and the P3-P9 domain is shown in blue. Base-paired regions are demarcated.

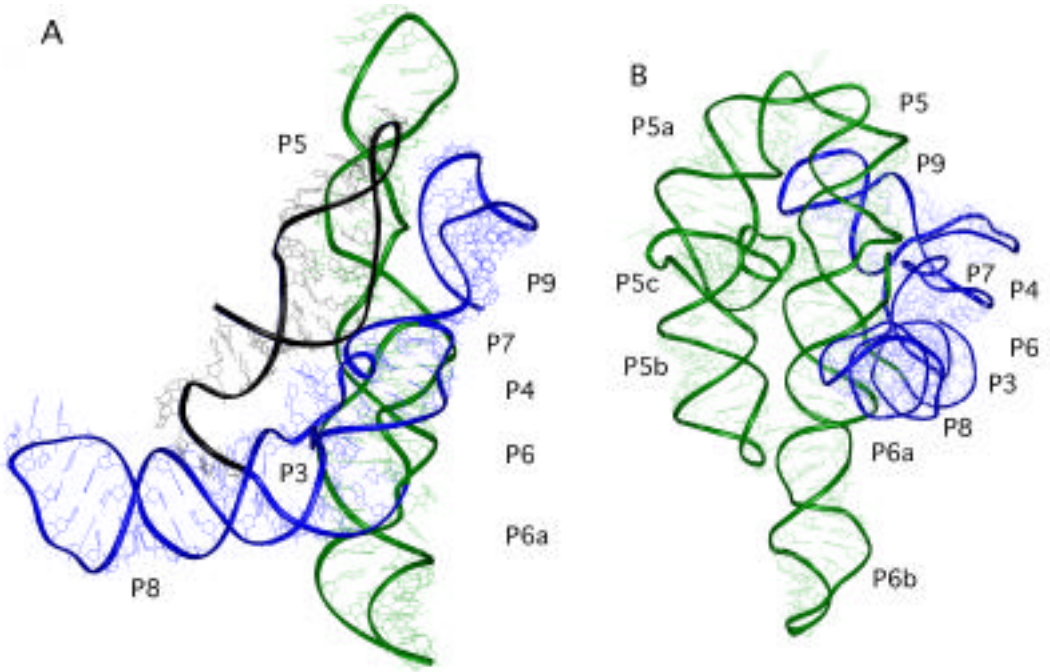
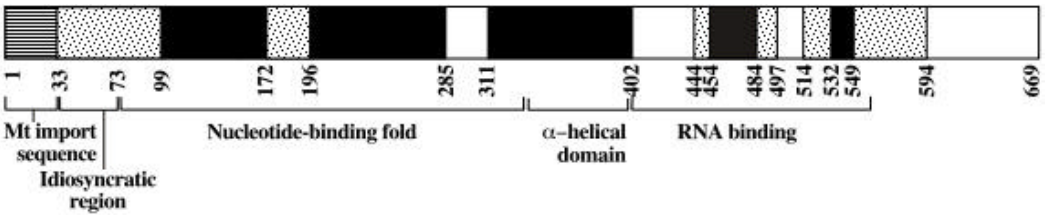


Figure 1.4: Map of the CYT-18 protein.

Black boxes indicate regions strongly conserved among bacterial TyrRSs, and stippled boxes indicate regions conserved only between the *N. crassa* and *P. anserina* mt TyrRSs, which function in group I intron splicing (Mohr et al., 2001). The boundaries of the nucleotide-binding fold, α-helical intermediate domain, and C-terminal RNA-binding domain based on homology to other TyrRSs are shown below (Mohr et al., 2001).



Chapter 2: Materials and Methods

2.1 MATERIALS AND METHODS FOR STUDYING THE ABILITY OF CYT-18 PROTEIN TO SUPPRESS GROUP I INTRON MUTATIONS *IN VIVO*

2.1.1 *E. coli* strains and growth media

E. coli strains were C600 *thyA::kan*^{r}, used for *td* plating assays (Bell-Pedersen et al., 1991); DH5αF', used for plasmid amplification and cloning; and BL21 (DE3; pLysS); Novagen, Madison, WI), used for expression of the CYT-18 protein. Bacteria were grown on LB or minimal media (MM) supplemented with 0.1% Norit-A-treated Casamino acids and 0.2% (w/v) glucose (Belfort et al., 1983). Thymine was added at 50 mg/l. Antibiotics were added at the following concentrations: ampicillin, 100 mg/l; kanamycin, 40mg/l; chloramphenicol, 25 mg/l; and trimethoprim, 20mg/l.

2.1.2 Construction of *td* intron mutants

All *td* intron constructs were based on pTZ*td*ΔP6-2, which contains a 265-nt derivative of the *td* intron cloned in pTZ18U (Galloway Salvo et al., 1990). The 265-nt intron has a deletion in L6a (positions 99-849, see Figure 3.1), which removes the intron ORF and introduces a *Sma*I site. pTZ*td*1302 was derived from pTZ*td*ΔP6-2 by filling in the *Eco*RI site in the polylinker with Klenow polymerase. Plasmid pTZ*td*1303, the parent construct for the P7[5'] and P7[3'] mutants, was derived from

pTZ*td*1302 by oligonucleotide-directed mutagenesis (Kunkel et al., 1987) to introduce an *Eco*RI site in L9.1 by changing U976 to G. All of the above *td* intron constructs gave the same phenotypes as pTZ*td*ΔP6-2 in *td* plating assays.

td introns containing random nucleotides in specific regions were generated by PCR mutagenesis (Cormack, 1992). To construct the P7[5'] and P7[3'] mutants, PCR was carried out using pTZ*td*1302 as the template. Primers CAM2 (5' CTG CCC GGG TTC TAC ATA AAT GCC TAN NNN CTA TCC CTT) and GM20 (5' CGG AAT TCT ATC CAG C) for the P7[5'] mutants, and primers CAM1 (5' CGG AAT TCT ATC CAG CTG CAT GTC ACC ATG CAG AGC NNN CTA TAT CTC) and the M13 forward primer for the P7[3'] mutants, where N refers to a randomized nucleotide during primer synthesis. Mutants having single nucleotide substitutions in G871 were constructed in the same way as P7[5'] mutants, but using primer CAM3 (5' CTG CCC GGG TTC TAC ATA AAT GCC TAA CHA CTA TCC CTT, where H is A, C or T) instead of CAM2. The PCR products were purified in 1% (w/v) agarose gels, digested with *Eco*RI and *Sma*I and cloned into the corresponding sites of pTZ*td*1303, which had been similarly digested and purified.

2.1.3 Colony assays of CYT-18-dependent splicing of mutant *td* introns

Libraries of pTZ plasmids containing mutant *td* introns were created as described above and transformed into the *E. coli* strain DH5αF'. Individual plasmids were isolated from randomly chosen colonies and transformed intro the *E. coli* C600 *thyA::kan*^{r} strain, either alone or in combination with pA550, which expresses the

wild-type CYT-18 protein (Mohr et al., 1992). Cells were grown overnight at 37ºC in LB medium containing the appropriate antibiotics (kanamycin to select for the *thyA*$^-$ strain, ampicillin to select for the pTZ*td* plasmid, and chloramphenicol when necessary to select for pA550). For the plating assays, 2 µl of cells were spotted on plates containing minimal media (MM), minimal media plus thymine (MMT), or minimal media plus thymine and trimethroprim (TTM), in combination with the appropriate antibiotics. Splicing phenotypes were characterized after overnight incubation at 37ºC and are based on analysis of four to sixteen colonies in at least two independent experiments (Mohr et al., 1992). Phenotypes are defined as follows: +++, maximal splicing, growth on MM comparable to MMT and no growth on TTM; ++, moderate splicing, growth on MM comparable to MMT and incomplete inhibition on TTM; +, weak splicing, growth on MM less than or equal to that on MMT, and growth on TTM comparable to MMT; -, no splicing, no growth on MM and growth on TTM comparable to MMT. After scoring the phenotype, mutations in the targeted regions were identified by dideoxy sequencing (Sanger et al., 1977), using primer NBS2 (5' GAC GCA ATA TTA AAC GGT). The mutant *td* introns were then sequenced completely to insure that no adventitious mutations had been introduced during mutagenesis.

2.1.4 Synthesis of DNA templates for *in vitro* transcription of RNAs for RNA structure mapping and RNA splicing

The *in vitro* transcripts used in the RNA structure mapping experiments begin at intron position 8, and end 5-nt upstream of the 3' splice site (position A1012; Figure 3.1). The transcripts used have an extra 5' G residue from the T7 promoter and 5 extra nucleotides (5'AATTT) derived from the PCR primer at their 3' end. DNA templates used to synthesize these transcripts were derived from the pTZ*td* plasmids by PCR with primers TDSXM-5' (GGT ACC TAA TAC GAC TCA CTA TAG GGC CTG AGT ATA AGG TGA C), which contains a sequence corresponding to introns position 8 to 26 and adds a phage T7 RNA polymerase promoter, and TDSXM-3' (5' TTT AAA TGT TCA GAT AAG GTC GTT AA), which contains a sequence complementary to intron positions 991 to 1012. PCR was carried out in 100 µl of reaction media containing 1 µM of each primer and 5 units of *Taq* DNA polymerase (Life Technologies, New York, New York) for 25 to 30 cycles: one minute 94ºC denaturation, one minute 55ºC annealing, one minute 72ºC extension with an initial denaturation of 5 minutes at 94ºC and a final extension of 10 minutes at 72ºC. The PCR products were purified in a 1% (w/v) agarose gel, extracted with phenol/cholorform/isoamyl alcohol (25:24:1 by vol.; phenol-CIA), ethanol precipitated and then used directly for *in vitro* transcription.

DNA templates used to synthesize *in vitro* transcripts for RNA splicing reactions were derived from the pTZ*td* plasmids by PCR using primers TDK5' (5'

CGG GAT CCC GTA ATA CGA CTC ACT ATA GGG ATC AAC GCT CAG), which contains a sequence corresponding to positions –33 to –15 in the 5' exon and introduces a phage T7 RNA promoter, and GM24 (5' GCT CTA GAC TTA GCT ACA ATA TGA AC), which contains a sequence complementary to positions 34 to 51 in the 3' exon and introduces an *Xba*I site. The PCR products were digested with *Xba*I, purified in a 1% (w/v) agarose gel, extracted with phenol-CIA, ethanol precipitated and then used directly for *in vitro* transcription.

2.1.5 *In vitro* transcription reactions for RNA used in RNA structure mapping

In vitro transcription reactions were carried out for two hours at 37ºC in 100 µl of reaction media containing 500 units of T7 RNA polymerase (BRL-GIBCO), 5 µg of DNA template, 1 mM of each NTP, 40 mM Tris-HCl (pH 8.0), 25 mM NaCl, 8 mM $MgCl_2$, 2mM spermidine-$(HCl)_3$, 5 mM dithiothreitol (DTT) and 1 µl of ribonuclease inhibitor (Amersham, Arlington Heights, IL). After transcription, the DNA was digested with 50 units of DNase I (Pharmacia, Piscataway, NJ) for 20 minutes at 37ºC. Transcripts were then extracted with phenol-CIA, centrifuged through a spun column containing Sephadex G-50 (Sigma, St. Louis, MO), and ethanol precipitated. To prepare 5' end-labeled transcripts for Fe(II)-EDTA cleavage experiments, *in vitro* transcription was carried out in the presence of 1 mM guanosine to generate RNAs having a 5'OH group. These RNA's were 5' end-labeled directly with phage T4 polynucleotide kinase (New England Biolabs, Beverly, MA) and [γ-^{32}P]ATP (3000 Ci/mmol; NEN-DuPont, Boston, MA) (Sambrook et al., 1989;

Caprara et al., 1996a). ^{32}P labeled transcripts were purified by electrophoresis in a 6% polyacrylamide/8 M urea gels. After using autoradiography to visualize the RNA bands, the bands were sliced out of the gels and the RNA was isolated by electroelution, ethanol precipitated, dissolved in distilled water, phenol-CIA extracted and centrifuged through a Sephadex G-50 spun column.

2.1.6 *In vitro* transcription reactions for RNA used for *in vitro* splicing

In vitro transcription reactions were carried out for two hours at 37ºC in 100 µl of reaction media containing 500 units of T7 RNA polymerase, 2 µg of DNA template, 30 mM Tris-HCl (pH 7.9), 7.5 mM DTT, 4.5 mM $MgCl_2$, 1.5 mM spermidine-$(HCl)_3$, 1 µl of human placental ribonuclease inhibitor, 100 to 150 µCi of [α-^{32}P]UTP (3000 Ci/mmol; NEN-DuPont), 1.25 mM UTP, 2.5 mM ATP and GTP, and 3.75 mM CTP. For those introns with high self-splicing activity, the CTP concentration was increased to 5.0 or 7.5 mM to chelate free magnesium ions, thereby suppressing self-splicing during transcription (Galloway Salvo et al., 1990). Following transcription the DNA template was digested with DNase I, and the *in vitro* transcripts were purified by electrophoresis in 6% polyacrylamide/8 M urea gels, as described above. The resulting transcripts contain a 33-nt 5' exon, the 265-nt intron, and a 51-nt 3' exon.

2.1.7 Fe(II)-EDTA analysis

For Fe(II)-EDTA analysis, 5'-end-labeled *in vitro* transcripts (~ 400,000 cpm, 6 pM) were dissolved in 7 µl of reaction media containing 20 mM HEPES (pH 7.5), 100 mM KCl, and 0, 3 or 8 mM $MgCl_2$. The RNA was renatured by incubation for 20 minutes at 55ºC, followed by 20 minutes at 37ºC. Then, 1 µl each of 10 mM $(NH_4)_2Fe(SO_4)_2$, 20 mM EDTA, and 100 mM DTT were added sequentially to the side of the tube and then rapidly mixed with the RNA by brief centrifugation. The reactions were incubated for 90 minutes at 37ºC and terminated by adding 1 µl of 100 mM thiourea and 11 µl of 2X gel loading buffer (10 M urea, 45 mM Tris-borate (pH 7.6), 11 mM EDTA, 0.1% (w/v) bromophenol blue and 0.1% (w/v) xylene cyanol. Fe(II)-EDTA cleavage products were analyzed in 9% polyacrylamide/8 M urea gels, alongside ladders obtained from the same 5' end-labeled RNAs by partial alkaline hydrolysis or partial digestion with RNase T_1 or U_2 (Pharmacia). Gels were run for different distances to resolve different regions of the intron RNA and quantified using a Molecular Dynamics Phosphorimager 445SI. Lanes were normalized for loading differences based on the radioactivity in three to five bands that were accessible to Fe(II)-EDTA cleavage under all assay conditions (*cf*. Laggerbauer et al., 1994). Protected regions were identified by visual inspection and quantified by phosphorimager analysis. A position was considered protected if the ratio of the normalized radioactivity at 0 mM magnesium to that of 3 or 8 mM magnesium was ≥ 1.25 (*cf.* Pan, 1995). The regions of protection generally have two- to fourfold less

radioactivity than regions of cleavage, with the greatest differences being about tenfold.

2.1.8 DMS and kethoxal modification

DMS and kethoxal modification were in 100 µl of reaction media containing 20 nM *in vitro* transcript, 20 mM HEPES (pH 7.5), 100 mM KCl, plus 3 or 8 mM $MgCl_2$ or in 20 mM HEPES (pH 7.5), 1 mM EDTA (denaturing conditions). After renaturing the RNAs by incubation in reaction medium for ten minutes at 55ºC and ten minutes at 37ºC, modification reactions were initiated by adding 1 µl of kethoxal (7.5 mg/ml in 80% ethanol; ICN Biomedical, Costa Mesa, CA) or 1µl of DMS (diluted 1:6 in ethanol; Aldrich, Milwaukee, WI). The reactions were incubated for four minutes (kethoxal) or three minutes (DMS) at 37ºC or for 30 seconds at 55ºC (denaturing conditions), then terminated by adding 10 µg of *E. coli* tRNA and EDTA to 10 mM. For kethoxal reactions, $K_2B_4O_7$ was added to a final concentration of 25 mM, to stabilize modified G residues. The modified RNAs were ethanol precipitated twice, which was found to be crucial for efficient reverse transcription, and then dissolved in 10 µl of 50 mM Tris-HCl (pH 8.3), 75 mM KCl, plus 25 mM $K_2B_4O_7$ for kethoxal-modified RNAs. Sites of modification were detected as stops to reverse transcription one base before the modified nucleotide residue (Caprara et al., 1996a). Reverse transcription reactions were carried out with M-MLV reverse transcriptase (Life Technologies) using 5' end-labeled primer TDSXM-3', which is complementary to a sequence in P9.2. The reverse transcription products were

analyzed in 6% polyacrylamide/8 M urea gels, which were autoradiographed and quantified by phosphorimager analysis. A position was considered protected if the normalized radioactivity at 3 or 8 mM magnesium was ≤ 50% of that under denaturing conditions (*cf.* Caprara et al., 1996a).

2.1.9 *In vitro* splicing reactions and kinetic analysis

In vitro splicing reactions were carried out in 150 to 200 µl of reaction media containing 20 nM precursor ^{32}P-RNA (100 to 130 Ci/mmol), 20 mM Tris-HCl (pH 7.5), 100 mM KCl, and 3 or 8 mM $MgCl_2$ plus 100 nM CYT-18 protein dimer, where indicated. The CYT-18 protein was synthesized from the expression plasmid pEX560 in *E. coli* BL21(DE3;pLysS) (Kittle et al., 1991) and purified by the PEI-precipitation procedure (Saldanha et al., 1995). Prior to splicing, the precursor RNA was renatured by heating to 94ºC for one minute and then cooling to 37ºC over a period of 2 minutes. The RNA was diluted to 20 nM in 20 mM Tris-HCl (pH 7.5), 3 mM $MgCl_2$, 100 mM KCl and incubated on ice for at least 30 minutes, then added to the final reaction medium, and incubated for five minutes at 37ºC. k_{cat} and K_m^{GTP} were determined from time-courses of splicing reactions at different GTP concentrations, as described by Williamson et al. (1987). Splicing reactions were initiated by the addition of GTP to a final concentration of 250 nM to 5 mM and incubated at 37ºC. The GTP added was pre-equilibrated with an equal concentration of magnesium to prevent the GTP from chelating magnesium in the reaction media. At each time-point, a 20 µl portion was removed, and the reaction was terminated by adding 1 µl of

500 mM EDTA, then placed on solid CO_2. After the series was completed, the products were phenol-CIA extracted, resolved in 6% polyacrylamide/8 M urea gels, and quantified by phosphorimager analysis. At least 6 intermediate time-points were taken for each reaction, with a final 8 hours time-point representing the complete reactions. Controls for the slowest reactions (*i.e*. those at the lowest GTP concentrations) showed that no further splicing occurred during overnight incubation. Reactions of longer than one hour were covered with mineral oil or periodically spun down to minimize the effects of evaporation. The fraction of active precursor remaining at each time point (F) was determined as follows: $[products]_\infty$-$[products]_t$/ $[products]_\infty$ (Bass and Cech, 1986), where $[products]_\infty$ is the amount that had reacted at the 8 hours time-point, and $[products]_t$ is the amount that had reacted at time *t*. k_{obs} at each GTP concentration was determined from the slope of a plot of lnF vs. *t.* A plot of k_{obs} vs. $[GTP]/k_{obs}$ gives a straight line defined by the equation $k_{obs}=K_m^{GTP}(k_{obs}/[GTP]) + k_{cat}$. K_m^{GTP} was then determined from the slope of the line and k_{cat} from the *y*-intercept (Williamson et al., 1987).

2.2 MATERIALS AND METHODS FOR DIRECTED HYDROXYL-RADICAL CLEAVAGE

2.2.1 *E. coli* strains and growth media

E. coli strain DH5αF' was used for cloning and HMS174(DE3;pLysS) (Novagen) for protein expression of wild-type and mutant CYT-18 proteins. Bacteria were grown in

LB media supplemented with ampicillin (100 mg/l) or chloramphenicol (25 mg/l) to maintain plasmids.

2.2.2 Construction of CYT-18 protein mutants

Wild-type CYT-18 protein was expressed from plasmid pEX560, which contains a 638-amino acid CYT-18 ORF, lacking the mt targeting sequence, cloned downstream of the phage T7 promoter in pET3a (Mohr et al., 1992). CYT-18 proteins lacking one or both endogenous cysteine residues, or containing single-cysteine substitutions at different positions, were constructed by generating two partially overlapping PCR products, which were then ligated together into the parent plasmid to replace the desired region of the CYT-18 ORF. The 5'- and 3'-PCR products each contain a unique restriction site for cloning and an overlap region, which contains both the desired mutation and a common restriction site, introduced via silent mutations, for joining the fragments. The regions generated by PCR were sequenced fully to insure no adventitious mutations were present.

2.2.2 Recombinant plasmids used for RNA transcription

Recombinant plasmids pLSUcam and pND10 were used to synthesize *in vitro* transcripts containing the *N. crassa* mt LSU and *ND1* introns, respectively. pLSUcam was derived from pBD5a (Guo and Lambowitz, 1992) by replacing the phage T3 promoter with the phage T7 promoter via PCR mutagenesis with primers P1/P3top (5'- CGG AAG CTT AAT ACG ACT CAC TAT AGG AAA AGC TAC

GCT AGG G) and P7/P9bot (5'- CGG GAT CCT TTG TAC ACT CTT GCG CAA A). Primer P1/P3top corresponds to exon positions -19 to -2, from the 5' splice site, and contains the T7 promoter and a *Hind*III site. Primer P7/P9bot corresponds to exon positions +11 to +30 the 3' splice site and a *Bam*HI site. The resulting 483-nt PCR product was gel-purified on a 1% (w/v) agarose gel and cloned between the *Bam*HI and *Hind*III sites of pUC19 (NEB). Digestion of pLSUcam with *Bam*HI and transcription with phage T7 RNA polymerase yields a 440-nt *in vitro* transcript with a 19-nt 5' exon, the 388-nt intron, and a 33-nt 3' exon.

Plasmid pND10 was derived from pND1m (Wallweber et al., 1997) by PCR mutagenesis with primers 5'ND1Xal (5' CGG GAT CCT TAG TAG CCT CCT GAT GAG TCC GTG AGG ACG AAA CGG TAC CCG GTA CCG TCG AGG CTA CTA ATA TGA T) and 3'ND1Xal (5'GGA ATT CGG TCC CAT TCG CCA TGC CGA AAG CAT GTT GCC CAG GCG GCG CCA GCG AGG AGG CTG GGA CCA TGC CGC CCT GAT TTC ATT TTT TTT A). 5'ND1Xal corresponds to positions -6 to +11, appends a hammerhead ribozyme to the 5' end of the *ND1* intron sequence, and contains a *Bam*HI site for cloning. Primer ND1Xal3' corresponds to positions +187 to +197, appends a hepatitis δ virus ribozyme at the 3' end of the *ND1* group I intron sequence and contains an *Eco*RI site for cloning. The resulting 314-nt PCR product was gel-purified on a 1% (w/v) agarose gel, and cloned between the *Bam*HI and *Eco*RI sites of pET-5-a (Novagen). Ribozymes were added to the 5' and 3' ends of the construct to produce a homogenous RNA population after transcription

(Ferre-D'Amare et al., 1998). Transcription of *Eco*RI-digested pND10 accompanied by ribozyme cleavage yields a 210-nt RNA containing a 6-nt 5' exon, the 197-nt *ND1* intron, and a 7-nt 3' exon.

2.2.3 Synthesis of intron-containing RNAs

In vitro transcripts containing ΔORF-derivatives of the *N. crassa* mt LSU and *ND1* introns and flanking exons were synthesized from pLSUcam and pND10, respectively. *In vitro* transcription and 5'-labeling with phage T4 polynucleotide kinase and gel purification of labeled RNAs were as described in section **2.1.6**. Prior to use, RNAs were renatured by incubating in splicing buffer (100 mM KCl, 5 mM $MgCl_2$, 25 mM Tris-HCl (pH 7.5), 10% glycerol (w/v)) for 20 min at 55°C and then cooling to room temperature (Caprara et al., 1996a).

2.2.4 Protein purification and biochemical assays

Wild-type and mutant CYT-18 proteins were purified by a modified polyethyleneimine-precipitation procedure (Saldanha et al., 1995). Dithiothreitol and EDTA were omitted from all solutions, as they compete or interfere with the Fe-EPD conjugation reaction. All solutions were treated with Chelex-100 (Sigma) to remove trace amounts of metals, which interfere with the Fe-EPD complex. Further, CYT-18 proteins were eluted from heparin-Sepharose columns using a step gradient. In the revised procedure, the loaded column was washed successively with 25 ml each of buffer (25 mM Tris-HCl (pH 7.5), 10% (w/v) glycerol) containing 25 mM KCl to

remove unbound proteins, and 250 ml KCl, to remove the *E. coli* TyrRS. Following these washes, the CYT-18 protein was eluted with 25 ml of buffer containing 500 mM KCl. 1-ml fractions were collected, and peak fractions were used directly. Tyrosyl-adenylation, TyrRS and RNA splicing assays were as described (Mohr et al., 2001). The RNA splicing assays used 10 nM CYT-18 protein and 20 nM ^{32}P-labeled RNA substrate. Wild-type and mutant CYT-18 preparations were >90% pure and contained 80-95% active protein, as judged by tyrosyl-adenylation assays. Protein concentrations were measured by the dye-binding method (Bradford, 1976), using a Bio-Rad protein assay kit (Bio-Rad, Hercules, CA) with bovine serum albumin as a standard. All CYT-18 concentrations refer to the homodimer, which is the active species for aminoacylation and RNA splicing (Saldanha et al., 1995).

2.2.5 Conjugation of Fe(II)-EPD to CYT-18 proteins

EPD-Fe-conjugated CYT-18 proteins were prepared by incubating 10 µM protein with 20 µM EPD-Fe (obtained from Dr. Robert Fox, The University of Texas, Medical Branch, Galveston, TX) in 100 µl of 0.5 M KCl, 25 mM Tris-HCl (pH 7.5), 10% glycerol (w/v) for 60 min at room temperature. Conjugation reaction mixes were then spotted onto a 0.45 µm Type VS filter (Whatman, Clifton, NJ), and dialyzed against 50 ml of 250 mM KCl, 25 mM Tris-HCl (pH 7.5), 10% glycerol (w/v) for 60 min at room temperature, to remove unincorporated EPD-Fe. The conjugated proteins were stored at 4°C.

2.2.6 Cysteine accessibility assays

Accessibility of cysteine residues to modification was determined by using 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB; Ellman, 1959; Hall and Fox, 1999). 8 µl of a 1 M DTNB (Aldrich) solution was added to protein samples, giving a final volume of 800 µl and a final concentration of either 500 nM or 1 µM protein. Absorbances of the protein samples were recorded at 0, 5 min and 30 min, using a wavelength of 412 nm. After subtracting the background absorbance of a buffer solution containing no protein, the concentration of accessible thiols was calculated using $c = A/ \ b$, where A is the absorbance, is the extinction coefficient, and b is the path length of the cell, and $_{412} = 1.36 \times 10^4 \ M^{-1} \ cm^{-1}$. Accessibility was calculated as [accessible thiols] / [total thiols].

2.2.7 Directed hydroxyl radical probing

For hydroxyl radical cleavage, 5'-labeled RNAs (50 nM) were complexed with equimolar EPD-Fe-conjugated protein in splicing buffer (5 mM $MgCl_2$, 0.5 M KCl, 25 mM Tris-HCl (pH 7.5), 10% glycerol (w/v)) for 20 min at room temperature. Cleavage reactions were initiated by adding sodium ascorbate to 30 mM. The reactions were incubated for 60 min at room temperature and terminated by adding 1 µl of 10 M thiourea and 10 µg of *E. coli* tRNA carrier. Reactions were then extracted with phenol-CIA. The cleavage products were analyzed in denaturing 6 or 9% polyacrylamide/8 M urea gels, against RNA sequencing ladders generated by partial

alkaline hydrolysis of the same 5'-labeled RNAs or iodine cleavage of 5'-labeled 10% α-S-ATP- or α-S-GTP-substituted RNAs.

2.2.8 Three-dimensional graphic modeling

Graphic modeling was done using the Insight II software package (Accelrys, San Diego, CA). To construct the models, the previous group I intron RNA models were aligned with $tRNA^{Tyr}$ in the *Thermus thermophilus* cocrystal structure using the overlaps between the GC pair at P4 bp-3 and D-arm bp-2, P7 [5']-2,3 and variable arm 47 and 47:1, and L9-1 and the discriminator base (Caprara et al., 1996b).

Chapter 3: The CYT-18 protein suppresses structural defects in the two major helical domains of the group I intron core: Examination of the P7 region

3.1 INTRODUCTION

My goal for experiments in this chapter was to help establish the role of the CYT-18 protein in group I intron splicing and to identify potential regions of contact within the intron. At the time these experiments were performed, it had been established that the CYT-18 protein promoted the splicing of group I introns by stabilizing the core of the group I intron in the active conformation (Guo and Lambowitz, 1992; Mohr et al., 1992; Mohr et al., 1994). RNA structure mapping experiments had revealed that two CYT-18 dependent group I introns, the *N. crassa* mt LSU and *ND1* introns, formed most of the short-range helices of the conserved group I intron core secondary structure, but failed to form any recognizable tertiary structure in the absence of CYT-18 (Caprara et al., 1996a). Even in the presence of elevated levels of magnesium, which allows many group I introns to self-splice *in vitro*, the CYT-18 dependent group I introns remained largely unfolded. However, upon the addition of the CYT-18 protein, a pattern of protection consistent with folding of the intron into an active conformation was seen (Caprara et al., 1996a). These protections included those expected for folding of the intron core (Celander and Cech, 1991; Heuer et al., 1991) and additional protections that possibly reflect

protein contact. Notably, the majority of the potential CYT-18 contact sites seen were in the P4-P6 domain of the group I intron, but additional sites were seen within the P3-P9 domain (Caprara et al., 1996a). Consistent with these findings, a small RNA corresponding to the P4-P6 domain of the mt LSU intron bound the CYT-18 protein independently with high affinity (K_d = 130 pM), while addition of sequences from the P3-P9 domain strengthened binding about 6-fold (K_d = 22 pM) (Guo and Lambowitz, 1992; Saldanha et al., 1996). Additionally, *in vitro* selection experiments showed that critical nucleotides for CYT-18 binding are clustered around the junction of the P4-P6 stacked helix (Saldanha et al., 1996). Kinetic analysis showed that the binding of the CYT-18 protein to the *N. crassa* mt LSU intron occurs via an initial bimolecular step that is close to the diffusion limit (3.24×10^7 M^{-1} s^{-1}) followed by a slower step (0.54^{-1} s), presumably a conformational change (Saldanha et al., 1995). The above findings are consistent with a model in which the CYT-18 protein binds first to the P4-P6 domain of the unfolded intron, and then makes additional contacts to the P3-P9 domain, to stabilize the two domains into the catalytically active structure (Guo and Lambowitz, 1992; Caprara et al., 1996a).

Previously, collaboration between our laboratory and Dr. Marlene Belfort of The State University of New York, developed an *Escherichia coli* genetic assay for CYT-18-dependent group I intron splicing (Figure 3.1A, Mohr et al., 1992). This assay was based on the ability of the CYT-18 protein to suppress splicing-defective mutations in the normally self-splicing phage T4 *td* intron (Figure 3.2). In this assay,

splicing is detected by the ability of a plasmid-borne *td* gene, interrupted by the *td* intron, to provide thymidylate synthase (TS) to an *E. coli* strain containing a disruption in its endogenous *thyA* gene (Belfort et al., 1987). Plasmid pA550, which expresses the CYT-18 protein under control of the *lac* promoter, and a *td* plasmid bearing a mutant intron, were cotransformed into *E. coli thyA*::*kan*^{r} cells. If CYT-18 is able to suppress a specific mutation in the *td* intron, then the transformed cells are able to grow on minimal media lacking thymine (MM) and additionally, cannot grow on media containing trimethoprim (TTM). Failure of CYT-18 to suppress the mutations allows the cells to grow on media containing trimethoprim, but not minimal media lacking thymine. Simultaneous plating on media supplemented with thymine (TMM) is used as a control for growth comparisons (see Belfort et al., 1987). Inefficient splicing results in partial growth on MM and TTM.

In the initial development of this assay, a series of previously constructed *td* introns were used to test the ability of CYT-18 to suppress mutations that impaired the self-splicing ability of the intron (Mohr et al., 1992). It was shown that CYT-18 was able to suppress mutations found in the core of the group I intron, such as P4, P5 P6, P7, P8 and P9, but not outside the core, such as mutations in P1 that disrupt 5' splice site recognition (Mohr et al., 1992; Chandry and Belfort, 1987). *In vitro* experiments demonstrated that the CYT-18 protein was able to promote splicing of intron mutants at physiological magnesium concentrations (~3 mM, Lusk et al., 1968) where the wild-type *td* intron is efficiently self-splicing, but the mutants are impaired.

However, at elevated levels of magnesium (8 mM), several of the mutant *td* introns showed restored self-splicing activity, consistent with the mutations being structural in nature, or impaired in magnesium binding (Mohr et al., 1992). These results support the hypothesis that CYT-18 functions in group I intron splicing by stabilizing the catalytic core of the group I ribozyme into the catalytically active structure.

3.2 SPECIFIC AIMS

My goal was to extend the studies described above, which relied upon a previously developed mutant library, to examine a specific region of interest. In collaboration with two other members of the lab, we systematically examined three regions of interest within the core of the group I intron: the P7 region, and the two connecting strands J3/4 and J6/7. My studies focused on the P7 region of the intron.

3.3 EXPERIMENTAL STRATEGY

The basic approach was to construct libraries of *td* intron mutants by randomizing nucleotides in the regions of interest of the intron core, in my case the 5' and 3' strands of P7. Then, the *E. coli* genetic assay could be used to rapidly identify splicing defects that could or could not be suppressed by CYT-18. Libraries of pTZ*td* plasmids carrying random mutations of the P7 region of the *td* intron were transformed into the *E. coli thyA*$^-$ strain in the presence or absence of pA550, which expresses the wild-type CYT-18 protein. Individual transformants were then plated on MM, MMT and TTM plates. A representative plating assays is shown in Figure 3.1B. Cells carrying splicing defective *td* introns lack functional TS and are unable to

grow on MM, but grow on MMT and TTM plates. Restoration of splicing by CYT-18 results in synthesis of TS, which enables cells to grow on MM, but not on TTM. Partial suppression is indicated by intermediate growth phenotypes. After determination of the phenotypes, the mutant introns were sequenced to identify the mutations. Base-specific probes were used in chemical modification assays to assess the effect of mutations on the P7 region. RNA structure mapping with Fe(II)-EDTA was used to determine the extent of the structural defects in representative mutants, and *in vitro splicing* was used to determine the degree of suppression by CYT-18.

3.4 THE P7 REGION

The P7 helix is the binding site for the guanosine cofactor used to initiate the splicing reactions (see Figures 1.1 and 3.2). P7 is a long-range base-pairing interaction formed between P7[5'] and P7[3']. The G-C base-pair at P7 bp-2 is conserved in virtually all group I introns, and interacts with the guanosine cofactor (Michel et al., 1989). Biochemical and molecular modeling studies suggest that the guanosine cofactor forms a nucleotide triple in the deep groove of the G-C base pair, possibly with additional H-bonds to flanking nucleotides that require it to be rotated slightly out of the plane of the G-C base pair (Michel et al., 1989; Yarus et al., 1991). The bulged nucleotide residue adjoining the G-C base-pair is always either a C or an A in naturally occurring group I introns (Michel and Westhof, 1990; Cannone et al., 2002). In the *Tetrahymena* LSU intron, the bulged A has been modeled to form a nucleotide triple in the shallow groove of P7 bp-1, where it would help form the

structure of the guanosine binding site, but not directly contact the guanosine cofactor (Yarus et al., 1991). Additionally, lead cleavage studies have revealed a potential metal-binding site within the P7 region of the intron (Streicher et al., 1993).

3.5 METHOD OF ANALYSIS

Two sets of P7 mutants were analyzed, one randomized the first four nucleotides of P7[5'] (A869-A872) and the second randomized the last three nucleotides of P7[3'] (U945-U947) (Tables 3.1 and 3.2 respectively). Also included were separately constructed single nucleotide substitutions of the universally conserved G residue of the guanosine binding site (G871), and two previously described mutations in the bulged C of P7[5'] (5'-AuGA and 5' AgGA; Mohr et al., 1992, where the lower case letters indicate mutations relative to the wild-type sequences).

3.6 RESULTS OF THE PLATING ASSAY

Only three of the 60 P7 mutants analyzed spliced in the absence of CYT-18: P7[5'] 5'-AuGA, which has a substitution in the bulged nucleotide residue, and P7[5'] 5'-ACGg and P7[3'] 3'-UCc, both of which affect P7 bp-3 (P7[3'] mutants are listed in 3' to 5' orientation to convey base-pairing potential with P7[5'], see Tables 3.1 and 3.2). CYT-18 was able to suppress 17 of the remaining 57 mutants. CYT-18 was able to suppress many mutations that disrupt one base-pair in P7, but only two mutants that disrupted 2 base-pairs (P7[3'] 3'-Uaa and 3'-cCc), and none that disrupt all three base-pairs. CYT-18 was unable to suppress any of 12 mutations that

replaced the universally conserved G (G871) residue, including all three single nucleotide substitutions. However, CYT-18 was able to suppress mutations at all other positions, including substitution of any nucleotide for the conserved C residue that base-pairs with the conserved G residue (C946, P7[3'] 3'-UnU, where n represents any nucleotide). Among the P7 bp-1 mutations, CYT-18 suppressed only those with a purine in the 5' strand and a pyrimidine in the 3' strand. In addition, replacement of the A-U base-pair at P7 bp-1 with G-U in the mutant P7[5'- 5'-gaGA inhibited splicing while P7[5'] 5'-AaGA spliced efficiently, as shown by Schroeder et al., 1991. These results suggest that the *td* intron has some steric or chemical preference for the nucleotide residues at P7 bp-1 beyond their ability to base-pair.

3.7 RESULTS OF *IN VITRO* CHEMICAL MODIFICATION

To directly assess the effect of the mutations on the base-pairing of the P7 region, four representative mutants were analyzed by chemical modification (Figures 3.3 and 3.4). *In vitro* transcripts corresponding to the intron's core regions were modified with DMS or kethoxal in the reaction medium containing 0, 3 or 8 mM Mg^{2+}. These magnesium concentrations were chosen to reflect denaturing conditions, *in vivo* conditions, and conditions which restore *in vitro* splicing to the *td* mutants (See Materials and Methods). The *in vitro* transcripts lack the 5' and 3' splice sites, which prevents splicing activity that could result in a heterogeneous RNA population. After exposure to the modifying reagent, sites of modification were mapped by reverse transcription using a primer complementary to a sequence in P9.2. DMS

modifies unpaired A and C residues, whereas kethoxal modifies unpaired G residues (Ehresmann et al., 1987). Representative gels are shown in Figures 3.3 (DMS) and 3.4 (kethoxal). The wild-type intron was also analyzed simultaneously for comparison purposes (see Materials and Methods).

The four mutant introns that were chosen for structural examination were: P7[3'] 3'-cCc, P7[3'] 3'-Uaa, P7[5'] 5'-AuGg and P7[5'] 5'-ggGa. The modification patterns for the mutant introns at 3 and 8 mM Mg^{2+} had most of the protections corresponding to the short-range helices of the conserved group I intron secondary structure (*e.g.* P4, P5. P6a, P8 and P9), but showed increased modification in P7 (positions A869-A875, U943-U947), the region of the mutations. Additionally, increased modification was seen in the P3 region (G910-A915), another long-range base-pairing interaction also found in the P3-P9 domain. The impaired formation of P7 and P3 was particularly evident at 3 mM Mg^{2+}. Also at 3 mM Mg^{2+}, all of the P7 mutants showed increased modification in J6/6a (A81-A83) and P12[5'] (G898), and most showed increased modification in P6[3'] (G864, C865), J8/7 (G938, A941) and J7.1/7.2 (A890). G871, the universally conserved G residue at the guanosine binding-site, was protected at 3 mM Mg^{2+} in the wild-type and P7[5'] mutant introns, but was modified in the P7[3'] mutants (position indicated by arrowhead in Figure 3.4). The extent of modification was reduced in many cases at 8 mM Mg^{2+}, however, the P7, P3 and J8/7 regions still showed increased modification relative to the wild type.

3.8 RESULTS OF Fe(II)-EDTA ANALYSIS

The effect of some selected mutations on the structure of the intron's catalytic core was investigated by Fe(II)-EDTA structure mapping. Hydroxyl-radicals generated by Fe(II)-EDTA cleave the phosphodiester backbone of RNAs in regions exposed to solvent. Because the cleavages are largely insensitive to sequence or secondary structure, this reagent has been used extensively to assess tertiary folding of group I introns (Celander and Cech, 1991; Heuer et al., 1991). Fe(II)-EDTA assays were done at the same magnesium concentrations described in Section **3.7**. The products of the Fe(II)-EDTA reactions were analyzed in a series of denaturing 9% polacrylamide gels run for different distances, and the cleavage patterns were quantified by phosphorimager analysis. A representative gel of the Fe(II)-EDTA analysis is shown in Figure 3.5. Control reactions of mock-treated RNAs were done at 0 and 8 mM Mg^{2+} to assess for differences in background hydrolysis that might occur differentially at higher magnesium levels. The wild-type intron was analyzed simultaneously for comparison and to confirm that the wild-type intron showed regions of protection expected for tertiary structure formation (*cf*., Heuer et al., 1991). It should be noted that regions of protection are more clearly defined at 8 mM Mg^{2+} than at 3 mM.

At 3 mM Mg^{2+}, all four mutants examined showed cleavage patterns indicative of severe structural disruption. Both P7[3'] mutations, showed loss of protection throughout the intron. Mutant P7[5'] 5'AuGg showed only limited

protection in J6/6a (U80 and A82) and J7.2/3 (G905-A907). Mutant P7[5'] 5'-ggGA showed only limited protection in P7.2[5'] (G891-U894) and J7.2/3 (G905-A907). At 8 mM Mg^{2+}, all four mutants still showed loss of protection in P3[3'] (U912-A913), P5[5']-L5 (U58, C59 and U62), and J6/a/6-P7[5'] (A862-A875). Mutant P7[3'] 3'-Uaa also showed loss of protection in J5/4 and J6/6a (A68-U72 and U80-U84 respectively).

3.9 RESULTS OF KINETIC ANALYSIS

To determine the extent of suppression by CYT-18, I compared kinetic parameters for *in vitro* splicing of wild-type and representative mutant introns, including those analyzed structurally with chemical probes and Fe(II)-EDTA. Splicing reactions were carried out at different GTP concentrations in reaction medium containing 3 mM Mg^{2+} in the presence of CYT-18 or 8 mM Mg^{2+} in the absence of CYT-18 (see Materials and Methods). k_{cat} and K_m^{GTP} were determined in these experiments. The results are summarized in Tables 3.3 and 3.4.

From a qualitative point of view, the behavior of the mutant introns was as expected based upon the *in vivo* phenotypes. None of the isolated P7 mutants showed any detectable self-splicing in reaction medium containing 3 mM Mg^{2+}. As expected from the ability of the CYT-18 protein to suppress the splicing defects *in vivo*, the purified CYT-18 protein restored splicing of the mutant introns at 3 mM Mg^{2+}, but had little effect on the rate of the wild-type intron. The splicing of the mutant introns was also stimulated at 8 mM Mg^{2+}, which presumably compensates for RNA

structural defects in the absence of CYT-18. From a quantitative point of view, neither CYT-18 nor 8 mM Mg^{2+} restored splicing of the mutant introns to wild-type levels *in vitro,* and two mutant introns could not be spliced to completion (Tables 3.3, 3.4), suggesting that a substantial portion of these introns remain misfolded. These findings are consistent with the chemical probing and Fe(II)-EDTA assays, which showed that none of the mutant introns fold into a fully wild-type structure *in vitro*, even at high magnesium. In the presence of the CYT-18 protein, the P7 mutants had k_{cat} values that were only two-to-tenfold lower than the wild-type intron (pTZ*td*, Table 3.3), but had greatly elevated K_m^{GTP} values (100 to ~375 fold greater than wild-type), as might be expected for mutations at the guanosine-binding site. The increase in K_m^{GTP} was most pronounced in P7[3'] 3'-Uaa, which disrupts G871/C946 and P7 and -3, and least pronounced in P7[3'] 3'-Uca, which disrupts P7 bp-3. This is in good agreement with the relative rates of splicing in the presence of CYT-18 *in vivo* (see section **3.6**). Generally similar results were obtained for all P7 mutants at 8 mM Mg^{2+}, with k_{cat} values approaching that of the wild-type intron and K_m^{GTP} values remaining elevated. The more complete restoration of k_{cat} relative to K_m^{GTP} could reflect that CYT-18 or 8 mM Mg^{2+} restore most of the active site structure, but that specific nucleotides and/or local structure in P7 are required for optimal interaction with GTP.

3.10 RELATED RESULTS

The *in vivo* plating assay described in this chapter has also been used by others to examine the ability of the CYT-18 protein to suppress mutations in other key regions of the group I intron catalytic core. As these results are relevant to the discussion, I will present an overview of those studies.

Using strategies similar to those outlined above, the ability of the CYT-18 protein to suppress mutations of the J3/4 and J6/7 regions was examined. These connecting sequences are part of the triple-helical scaffold structure of the catalytic core, forming base triples with nucleotide residues in P6 and P4, respectively (Michel and Westhof, 1990). In the *td* intron, both J3/4 and J6/7 are three bases long with the sequences UAA and CUA respectively (Figure 3.2).

The importance of the J3/4 region was underscored by the fact that no self-splicing mutations were isolated. All single nucleotide substations, (nAA, UnA, and UAn) were deficient in self-splicing activity. The finding that any single nucleotide substitution at J3/4-1 inhibits self-splicing suggests that it plays an important role. In the crystal structure of the *Tetrahymena* intron, the residue at J3/4-1 was potentially able to interact with nucleotide residues in P6 and P3 (Golden et al., 1998). CYT-18 was able to suppress all single nucleotide substitutions in this region, and also suppressed a number of double mutants expected to disrupt both base-triple interactions, as well as one mutant in which all three positions had been changed

(Myers et al., 1996). Fe(II)-EDTA analysis revealed that the J3/4 mutations showed loss of protection primarily in the P4-P6 domain (Myers et al., 1996).

J6/7 is also part of the triple-helical-scaffold structure (Michel and Westhof, 1990). J6/7-1 and –2 form nucleotide triples in the deep groove with the P4 bp-2 and –3 respectively (Michel et al., 1990; Michel and Westhof, 1990). The A at J6/7-3 is universally conserved in group I introns and has been proposed to interact with the P3-P9 domain to help establish the correct relative orientation of the two major helical domains of the group I intron catalytic core (Michel and Westhof, 1990). Mutational analysis in combination with the *in vivo* genetic assay revealed that all single nucleotide substitutions at the base-triple-forming positions retained self-splicing activity. All double mutations that retained the universally conserved A residue at J6/7-3 (A368) lost self-splicing activity, but were suppressible by CYT-18 (Myers et al., 1996). CYT-18 was also able to suppress several mutants with alterations at A368, including one triple-mutation (Myers et al., 1996). Fe(II)-EDTA analysis of representative J6/7 mutants showed that the mutations caused loss of protection throughout the entire intron (ggA mutant) or around the area of mutation (CUg and aaA mutants). The loss of protections seen in the latter class of mutants could possibly reflect misalignment of all, or part of, the P3-P9 domain.

More recently, the genetic assay was used to examine the ability of CYT-18 to suppress mutations in the P4-P6 region of the *td* intron. This region has been identified as a potential binding site for CYT-18 by *in vitro* selection and RNA

footprinting experiments (Saldanha et al., 1996; Caprara et al., 1996ab). As seen for other regions, CYT-18 was able to suppress mutations in this region that impaired the self-splicing activity of the *td* introns (Chen et al., 2000). Mutations at the P4-P6 junction that disrupt base-pairing, base-stacking or base-triple interactions presumably lead to an altered geometry of the region. Thermal denaturation studies and Fe(II)-EDTA analysis indicated that this altered geometry in turn leads to grossly impaired tertiary-structure formation (Chen et al., 2000).

3.11 DISCUSSION

My results show that the CYT-18 protein can suppress splicing defects resulting from mutations in the P7 region of the group I intron core. Mutations that disrupt base-pairing in this region, as well as substitutions to two highly conserved positions, the bulged C residue at position 870, and C946, the base-pairing partner to G871, were suppressible by CYT-18. Notably, CYT-18 was unable to suppress mutations that changed the universally conserved G-residue at the guanosine-binding site (G871). All of the mutations suppressed by CYT-18 *in vivo* were also suppressible by elevated levels of magnesium *in vitro*. Taken together, these results suggest that the CYT-18-suppressible mutations affect group I intron structure rather than components directly involved in catalysis (*cf*. Burke et al., 1986). The findings that the phylogenetically conserved nucleotides can be replaced with any other nucleotide indicates that either that these conserved nucleotides do not play an

essential role in catalysis and/or that their function can be served by other nucleotides at lower efficiency.

Although G871 appears indispensable for splicing activity in our assays, some activity has been reported for the corresponding G to A mutation in a *Tetrahymena* LSU intron construct inserted in the *E. coli* large rRNA (Zhang et al., 1995). The splicing rate of this mutant intron in *E. coli* was estimated to be 200-fold less than that of the corresponding wild-type construct. The apparent difference in our results could reflect either that replacement of the conserved G-residue has a greater effect in the *td* intron than in the *Tetrahymena* intron or that such mutations reduce the rate of splicing below the limit of detection in the *td* plating assay.

The finding that CYT-18 suppresses mutations that impair base-pairing of P7 likely reflects its direct interaction with the P3-P9 domain of the catalytic core. In the *Tetrahymena* intron, the P4-P6 domain folds independently (Laggerbauer et al., 1994). Assuming that the *td* intron follows the same ordered folding pathway as the *Tetrahymena* LSU intron, the initial assembly of the P4-P6 domain should not be affected by mutations in P7, which forms late in the folding pathway (Zarrinkar and Williamson, 1994; Sclavi et al., 1998). Significantly, RNA structure mapping indicates that the P7 mutations also impair the formation of the P3 element, thereby grossly affecting the folding and alignment of the P3-P9 domain. This is particularly evident in the Fe(II)-EDTA structure mapping experiments that show that the introns are largely unfolded. Many of the observed protections of the group I intron arise

from the wrapping of the P3-P9 domain around the P4-P6. With P3 and P7 unable to form, the P4-P6 domain is largely accessible to cleavage. The ability of CYT-18 to suppress such defects may reflect that it binds to each of the two major helical domains of the catalytic core and stabilizes them in the correct relative orientation to form the intron's active site. Alternatively, CYT-18 may independently assist the assembly of each of the domains, and the assembled domains may then interact without further mediation by the protein. The finding that the P7 mutations also affect P3 provides evidence that the formation of these two long-range pairings is interdependent in the *td* intron. A similar interdependence was found in the *Tetrahymena* LSU intron by analyzing the effect of P3 and P7 mutations on the kinetics and extent of folding monitored by accessibility of DNA oligonucleotides and RNase H (Zarrinkar and Williamson, 1994; Zarrinkar and Williamson, 1996). It should also be noted that the assembly of the P3-P9 domain in the *td* intron might be less dependent upon the P4-P6 domain than in the *Tetrahymena* LSU intron. This is because the P3-P9 domain in the *td* intron may be stabilized by the peripheral elements P7.1/P7.2, P9.1/P9.2 and/or P12, which have been shown to stabilize the catalytic core of the bacteriophage T4 *sunY* intron (Doudna and Szostak, 1989; Michel et al., 1992). These additional peripheral elements may also contribute to the observed Fe(II)-EDTA protections in the P3-P9 domain of the *td* intron.

Kinetic analysis of the *in vitro* splicing reactions of several P7 mutants showed that CYT-18 or 8 mM Mg^{2+} restore k_{cat} to within tenfold of the wild-type

value, but K_m^{GTP} values in many cases remain 100-to 400-fold higher than those for the wild-type intron. These findings suggest that the active core structure is largely restored in these mutants, but that the substituted nucleotide residues still affect guanosine binding, either because they interact directly with the guanosine cofactor or affect the local structure of P7. In principle, this experimental system could be used to obtain further information about the interaction of specific nucleotide residues with the guanosine cofactor.

Considered with similar data for other regions generated using the *td* plating system, the nature of the mutations suppressed by CYT-18 supports a model in which CYT-18 binds to both the P4-P6 and P3-P9 domains of group I introns, and stabilizes them in the correct relative orientation to form the catalytic core. In principle, the CYT-18 protein may act either by increasing the rate of formation of correctly folded RNA structures, or by binding to correctly folded RNA structures, thereby slowing their disassembly. The latter would be analogous to the tertiary structure capture mechanism found for the protein-assisted splicing of the yeast *cob*-I5 intron by CBP2 (Weeks and Cech, 1996). Pre-steady state kinetic analysis suggests that the rate-limiting step in the interaction of the CYT-18 protein with the *N, crassa* mt LSU intron is a conformational change that occurs after CYT-18 binding (Saldanha et al., 1995). Thus, CYT-18 appears to play an active role in promoting tertiary structure formation in this intron.

The CYT-18 dependent group I introns in *N. crassa* mitochondria were presumably self-splicing initially and then became protein-dependent as a result of structural mutations that impair splicing (Lambowitz and Perlman, 1990; Lambowitz et al., 1999; Mohr et al., 1992). The finding that CYT-18 can compensate for a variety of structural changes in group I introns raises the possibility that CYT-18 dependent introns could evolve toward minimal size, using the protein as a scaffold to position the essential RNA structures for catalytic activity. Nevertheless, the CYT-18 dependent *N. crassa* mt introns retain the conserved group I intron secondary structures and phylogenetically conserved nucleotides at all key positions. The retention of these conserved structural features may reflect that they are required for optimal rates of splicing or that the introns have become CYT-18 dependent relatively recently in evolution, so that the RNAs have not yet had sufficient time to evolve beyond the original structure (Mohr et al., 1992).

3.12 SUMMARY

The results of this section show that CYT-18 is able to suppress mutations in the P7 region of the *td* intron that impair its base-pairing. CYT-18 was unable to suppress mutations to the universally conserved G residue (G871), which binds the guanosine cofactor required for splicing. However, CYT-18 was able to suppress mutations in its base-pairing partner, C987, and mutations in the highly conserved bulged A residue (A870). Chemical base-probing with DMS and kethoxal revealed that the mutations impaired the base-pairing of P7 and P3. Solution structure

mapping with Fe(II)-EDTA showed that the mutations largely impaired the overall tertiary structure of the introns at physiological magnesium concentrations. Quantitatively, kinetic analysis showed that the mutants had largely unaffected k_{cat} values, while K_m^{GTP} values were greatly increased. Qualitatively, neither CYT-18 nor elevated levels of magnesium restored splicing activity to wild-type levels. This work established the use of the *td* plating assay as a useful method in examining the ability of CYT-18 to suppress mutations in key regions of the core of the *td* intron. The plating assay has since been used to examine the ability of CYT-18 to suppress mutations in the J3/4, J6/7, P4, P6 and L9-P5 regions of the *td* intron (Myers et al., 1996; Chen et al., 2001; X. Chen and A.M. Lambowitz, unpublished). Taken together these results support the hypothesis that CYT-18 promotes group I intron splicing by binding both major helical domains of the catalytic core and stabilizing them in the correct relative orientation to form the intron's active site.

Figure 3.1: *In vivo* plating assay for representative P7 mutants.

(A) Strategy for *in vivo* plating assay. An *E. coli* strain having a disruption in the *thyA* gene, which encodes thymidylate synthase, was cotransformed with a plasmid containing the bacteriophage T4 *td* gene (pTZ*td*) and a plasmid expressing the CYT-18 protein (pA550). Transformants were selected on LB plates supplemented with antibiotics and replated on MMT, MM and TTM plates as described in section **2.1.3**. The bottom indicates splicing phenotypes corresponding to different patterns of growth on the three media.
(B) *In vivo* splicing of wild-type and representative *td* P7 intron mutants in the presence or absence of the CYT-18 protein. *E. coli* C600 *thyA::kan*^{r} containing wild-type or mutant pTZ*td* plasmids plus pA550 (+ CYT-18) or the vector pACYC184 (Chang and Cohen, 1978; - CYT-18) were plated on MM, MMT or TTM plates containing kanamycin, chloramphenicol and ampicillin and grown overnight at 37ºC. Splicing phenotypes are as defined in Table 3.1 and in section **2.1.3**. The wild-type (top) and mutant sequences (below) are shown to the left, with mutant nucleotide residues indicated in lowercase letters.

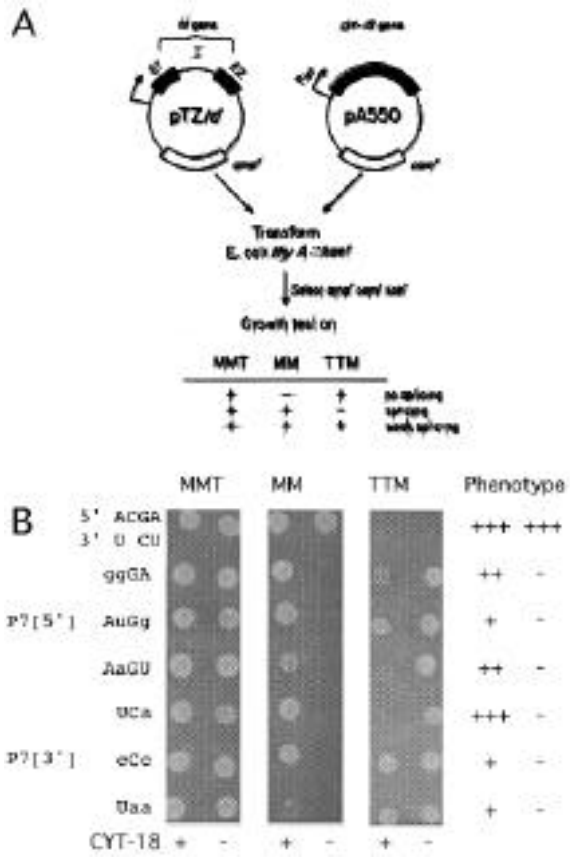


Figure 3.2: Secondary structure model of the phage T4 *td* intron.

The predicted secondary structure of the 265-nt *td* intron derivative, pTZ*td*1304. Nucleotide residues in the intron and exons are indicated in uppercase and lowercase letters respectively. Arrows indicate 5' and 3' splice sites. Thin, connecting lines between boxes indicate tertiary interactions. ΔORF indicates the deleted region containing the intron ORF. Nucleotide residues changed to create restriction sites are circled. The shaded area indicates nucleotide residues in P7 that were mutagenized for analysis.

Because of the University of Texas at Austin thesis requirements, this figure is too small to be useful. I suggest checking the original references.

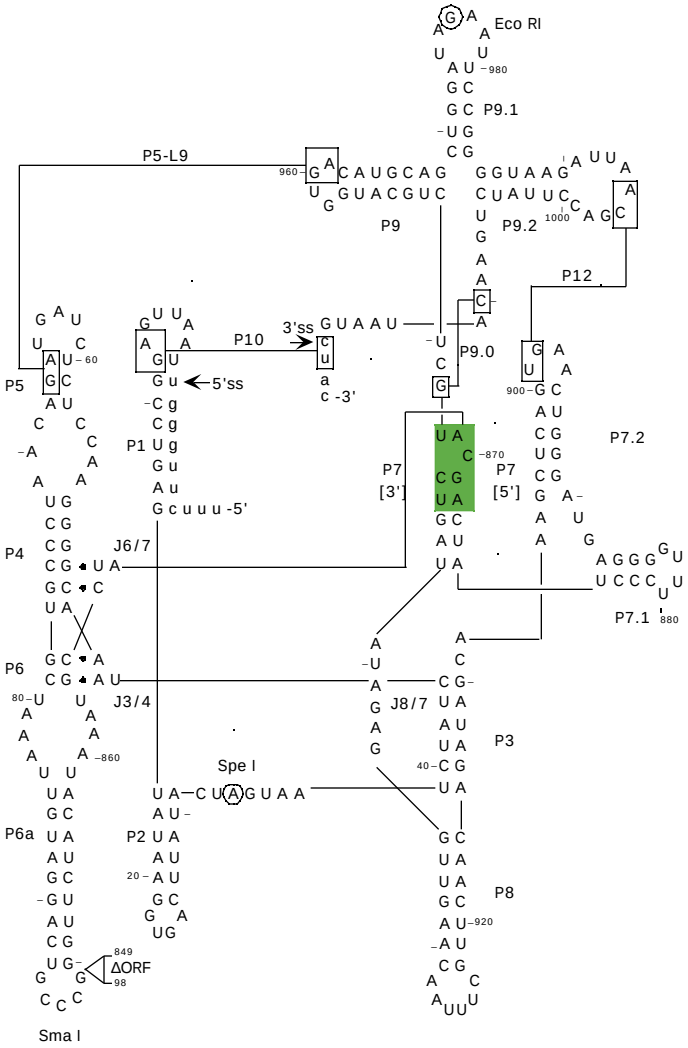


Figure 3.3: DMS modification patterns of wild-type and representative P7 mutant *td* introns.

In vitro transcripts containing the wild-type and mutant *td* introns were incubated with (+) or without (-) DMS in reaction media containing 0, 3 or 8 mM Mg^{2+} as indicated. The modified RNAs were reverse transcribed with M-MLV reverse transcriptase, using 5' end-labeled primer TDSXM-3', which is complementary to a sequence in P9.2. The products were resolved in a 6% polyacrylamide/8 M urea gel, which were autoradiographed and quantified by phosphorimager analysis. The regions analyzed extended from A70 to C941. Nucleotide positions that were sites of strong reverse transcription stops and/or regions of gel compression (C95-C97, C877-C879) were excluded from the analysis. Wild-type (left) and mutant sequences (right) are shown above the lanes, with mutant nucleotide residues indicated in lowercase letters. Dideoxy-sequencing ladder (T, G, C, A) obtained from the wild-type plasmid (pTZ*td*1303) using the same 5' end labeled primer were run in parallel lanes. Regions of the intron are demarcated to the right.

Because of the University of Texas at Austin thesis requirements, this figure is too small to be useful. I suggest checking the original references.

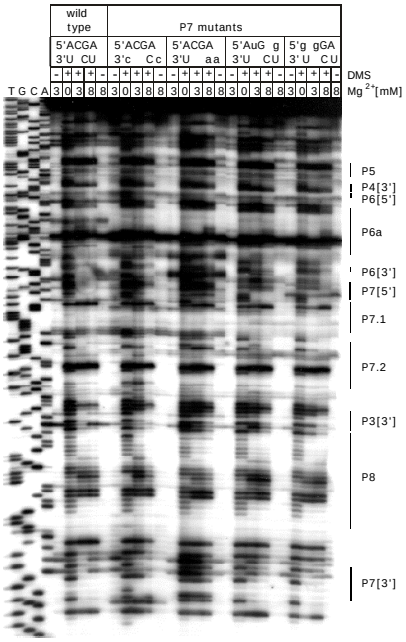


Figure 3.4: Kethoxal modification patterns of wild-type and representative P7 mutant *td* introns.

In vitro transcripts containing the wild-type and mutant *td* introns were incubated with (+) or without (-) kethoxal in reaction media containing 0, 3 or 8 mM Mg^{2+} as indicated. The modified RNAs were reverse transcribed with M-MLV reverse transcriptase, using 5' end-labeled primer TDSXM-3', which is complementary to a sequence in P9.2. The products were resolved in a 6% polyacrylamide/8 M urea gel, which was autoradiographed and quantified by phosphorimager analysis. The regions analyzed extended from A70 to C941. Nucleotide positions that were sites of strong reverse transcription stops and/or regions of gel compression (G98-G851, G891-G893) were excluded from the analysis. Wild-type (left) and mutant sequences (right) are shown above the lanes, with mutant nucleotide residues indicated in lowercase letters. Dideoxy-sequencing ladder (C, T, A, G) obtained from the wild-type plasmid (pTZ*td*1303) using the same 5' end labeled primer were run in parallel lanes. Regions of the intron are demarcated to the right. The arrowhead to the right indicates the position of G871.
Because of the University of Texas at Austin thesis requirements, this figure is too small to be useful. I suggest checking the original references.

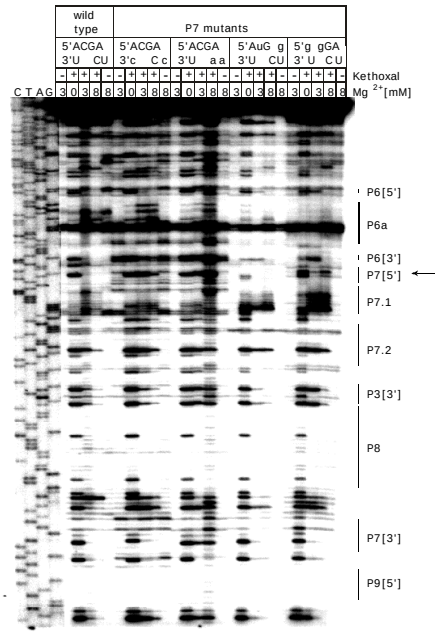


Figure 3.5: Fe(II)-EDTA cleavage patterns of the wild-type and representative P7 mutant *td* introns.

The 5' end labeled *in vitro* transcripts containing wild-type or mutant *td* introns were incubated with (+) or without (-) Fe(II)-EDTA in reaction media containing 0, 3 or 8 mM Mg^{2+} , as indicated. Cleavage products were analyzed in a 9% polyacrylamide/8 M urea gel, which was autoradiographed and quantified by phosphorimager analysis. The region analyzed extended from U43 to A913. Nucleotide positions cleaved in the absence of Fe(II)-EDTA were not included in the analysis. Wild-type (left) and mutant sequences (right) are shown above the lanes, with mutant nucleotide residues indicated in lowercase letters. Ladders: OH, partial alkaline hydrolysis; G, partial hydrolysis with RNase T_1; A, partial hydrolysis with RNase U_2. Regions of the intron are demarcated to the right. The arrow indicates P7[5'], the site of the mutations.

Because of the University of Texas at Austin thesis requirements, this figure is too small to be useful. I suggest checking the original references.

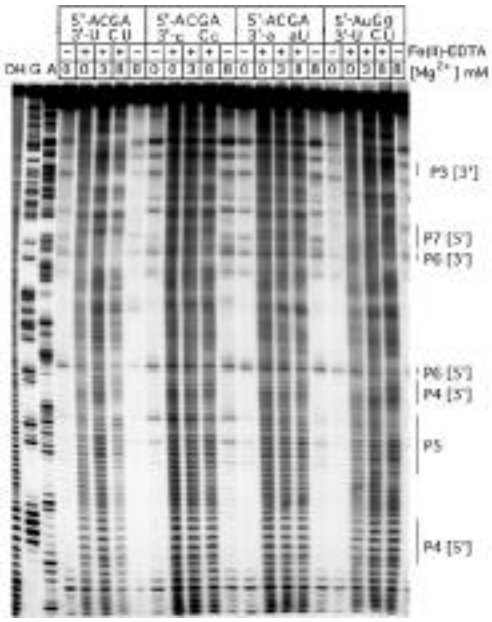


TABLE 3.1: *In vivo* splicing phenotypes of P7[5'] mutants.

Self-Splicing or CYT-18 Suppressible[a]		
Mutant	Splicing Phenotype[c] -CYT-18	+CYT-18
wt 5'-ACGA 3'-U CU	+++	+++
5'-ACGg 3'-U CU	++	+++
5'-AuGA[d] 3'-U CU	++	+++
5'-AgGA[d] 3'-U CU	-	++
5'-gaGA 3'-U CU	-	++
5'-ggGA 3'-U CU	-	++
5'-AaGu 3'-U CU	-	++
5'-AgGg 3'-U CU	-	++
5'-AuGg 3'-U CU	-	+
5'-gaGg 3'-U CU	-	++
5'-ggGg 3'-U CU	-	++

Non-Splicing[b]
Mutant
5' ACaA ACcA ACuA caGA gCGc cuGg uaGc ugGc ugGg Aaag ugaA cCag caug cgau cgcg cgug ggag gguc

[a] P7 sequences of the wild-type (top) and mutant introns (below) that were self-splicing or could be suppressed by coexpression of the CYT-18 protein. Lowercase letters indicate mutant nucleotide residues.

[b] P7[5'] sequences of mutant introns that did not splice *in vivo* in either the presence or absence of CYT-18.

[c] Splicing phenotypes in the presence or absence of CYT-18 scored by plating assays, as described in section **2.3.1**.

[d] Data from Mohr *et al*. (1992).

TABLE 3.2: *In vivo* splicing phenotypes of P7[3'] mutants.

Self-Splicing or CYT-18 Suppressible[a]		
Mutant	Splicing Phenotype[c]	
	-CYT-18	+CYT-18
wt 5'-ACGA 3'-U CU	+++	+++
5'-ACGA 3'-U Cc	++	+++
5'-ACGA 3'-U Ca	-	+++
5'-ACGA 3'-U Cg	-	++
5'-ACGA 3'-U aU	-	++
5'-ACGA 3'-U gU	-	+
5'-ACGA 3'-U uU	-	++
5'-ACGA 3'-c CU	-	+++
5'-ACGA 3'-U aa	-	+
5'-ACGA 3'-U uc	-	+
5'-ACGA 3'-c Cc	-	+

Non-Splicing[b]
Mutant
3' aCU Uga aCc aCg cCa gCc cCg aaU gaU cgU ggU aua auc cac cgc cua cuc gaa

[a] P7 sequences of the wild-type (top) and mutant introns (below) that were self-splicing or could be suppressed by coexpression of the CYT-18 protein. Lowercase letters indicate mutant nucleotide residues.

[b] P7[5'] sequences of mutant introns that did not splice *in vivo* in either the presence or absence of CYT-18.

[c] Splicing phenotypes in the presence or absence of CYT-18 scored by plating assays, as described in section **2.3.1**.

[d] Data from Mohr *et al*. (1992).

TABLE 3.3: Steady-state kinetic parameters of wild-type and selected mutant *td* introns[a] at 3 mM Mg^{2+} plus CYT-18.

Mutant	*In vivo* Phenotype + CYT-18	3 mM Mg^{2+} plus CYT-18			
		K_m^{GTP} (μM)	k_{cat} ($s^{-1}x10^{-3}$)	k_{cat}/K_m^{GTP} ($s^{-1}/\mu M x10^{-7}$)	Spliced[b] (%)
wild type					
pTZ*td* 1303[c]	+++	1.6	5.1	32,000	99
P7 mutants					
5'-AuGg 3'-U CU	+	340	1.5	44	52
5'-AaGu 3'-U CU	++	270	0.5	19	57
5'-ggGA 3'-U CU	++	360	2.4	67	86
5'-ACGA 3'-U Ca	+++	160	1.0	63	70
5'-ACGA 3'-c Cc	+	250	0.5	20	90
5'-ACGA 3'-U aa	+	600	0.8	13	25

[a] Summary of k_{cat} (s^{-1}), K_m^{GTP} (μM), catalytic efficiency (k_{cat}/K_m^{GTP}) in reaction medium containing 3 mM Mg^{2+} in the presence of CYT-18.

[b] Percent of intron spliced after 8 h at saturating GTP concentrations.

[c] Parent construct for the P7 mutants.

TABLE 3.4: Steady-state kinetic parameters of wild-type and selected mutant *td* introns[a] at 8 mM Mg^{2+} in the absence of CYT-18.

Mutant	*In vivo* Phenotype	8 mM Mg^{2+} (self-splicing)			
		K_m^{GTP} (μM)	k_{cat} (s^{-1}x10^{-3})	k_{cat}/K_m^{GTP} (s^{-1}/μMx10^{-7})	Spliced[b] (%)
wild types					
pTZ*td* 1303[d]	+++	2.2	4.2	19,000	96
P7 mutants					
5'-AuGg 3'-U CU	+	100	1.2	120	88
5'-AaGu 3'-U CU	++	85	2.4	280	85
5'-ggGA 3'-U CU	++	200	3.8	190	91
5'-ACGA 3'-U Ca	+++	28	2.0	710	95
5'-ACGA 3'-c Cc	+	200	1.7	85	96
5'-ACGA 3'-U aa	+	407	2.0	49	84

[a] Summary of k_{cat} (s^{-1}), K_m^{GTP} (μM), catalytic efficiency (k_{cat}/K_m^{GTP}) in reaction medium containing 8 mM Mg^{2+} in the absence of CYT-18.

[b] Percent of intron spliced after 8 h at saturating GTP concentrations.

[c] Parent construct for the P7 mutants.

Chapter 4: tRNA-like recognition of group I introns by CYT-18: Directed hydroxyl-radical cleavage studies using single-cysteine CYT-18 proteins

4.1 INTRODUCTION

The *Neurospora crassa* mitochondrial tyrosyl-tRNA synthetase, or CYT-18 protein, functions in both $tRNA^{Tyr}$ aminoacylation and group I intron splicing (Akins and Lambowitz, 1987). The splicing function reflects that CYT-18 recognizes conserved structural features of group I intron and promotes the formation of the catalytically active RNA structure (Guo and Lambowitz, 1992; Mohr et al., 1992; Caprara et al., 1996ab). Group I introns, like tRNAs, have minimal sequence conservation, but share a conserved three-dimensional structure consisting of two double-helical domains (Michel and Westhof, 1990; Golden et al., 1998). Biochemical and genetic studies led to a model in which CYT-18 binds first to the P4-P6 domain to promote its assembly, and then makes additional contacts with the P3-P9 domain to stabilize the two domains in the correct relative orientation to form the intron's active site (Caprara et al., 1996ab; Saldanha et al., 1996; Myers et al., 1996; Chen et al., 2001).

Class I aaRSs, like TyrRSs, consist of a structurally conserved N-terminal nucleotide-binding fold domain, which catalyzes amino acid activation and tRNA charging, followed by an aaRS-specific C-terminal RNA-binding domain, which interacts with the anticodon arm of the tRNA (Arnez and Moras, 1997). The

nucleotide-binding fold interacts with the tRNA's acceptor stem and typically contains an insertion, denoted connective peptide 1 (CP1), which contributes to determining acceptor-helix identity (Rould et al., 1989; Wakasugi et al., 1998). In TyrRSs, the nucleotide-binding fold is followed by an α-helical intermediate domain linked via a flexible hinge to a C-terminal RNA-binding domain, containing a region homologous to ribosomal protein S-4 (Gaillard and Bedouelle, 2001). In agreement with previous models, a cocrystal structure of the *Thermus thermophilus* TyrRS complexed with $tRNA^{Tyr}$ showed that the tRNA binds across the surface of the two subunits of the α2-homodimer, with the tRNA's acceptor stem interacting with the nucleotide-binding fold of one subunit and the anticodon arm interacting with the C-terminal domain of the other subunit (A. Yaremchuk, I. Kriklivyi, M. Tukalo & S. Cusack, submitted). The X-ray crystal structure also confirmed that TyrRS is unique among class I aaRSs in approaching the acceptor stem from the major groove side, as is characteristic of class II aaRSs (Bedouelle, 1990).

The CYT-18 protein is homologous to bacterial TyrRSs (Figure 5.1) and also functions as an α2-homodimer, with each dimer binding one group I intron RNA (Saldanha et al., 1995). Despite these structural similarities, however, only the *N. crassa* mt TyrRS and that of the related fungus *Podospora anserina* function in group I intron splicing, while the *Escherichia coli* and yeast mt TyrRSs do not function in splicing (Kämper et al., 1992). Thus, the two fungal mt TyrRSs must have some adaptation of the canonical structure that confers the ability to promote splicing

activity. Mutational analysis showed that regions of CYT-18 required for splicing overlap those that function in tRNA binding (Kittle et al., 1991). An apparent exception was an idiosyncratic N-terminal extension, whose deletion specifically inhibited group I intron splicing (Cherniack et al., 1990), but recent findings suggest that this region functions indirectly by stabilizing the structure of another region that contacts the intron RNA directly (Mohr et al., 2001). Otherwise, most mutations in different protein regions have parallel effects on splicing and TyrRS activity and a group I intron RNA was found to be a competitive inhibitor of aminoacylation, providing direct evidence that the two RNAs compete for the same or an overlapping binding site (Kittle et al., 1991; Guo and Lambowitz, 1992). Together, these findings suggest that CYT-18 forms similar but not identical complexes with its $tRNA^{Tyr}$ and group I intron RNA substrates.

Several different approaches have been used to try and identify regions of interaction between the CYT-18 protein and its RNA substrates. These include deletion and linker-insertion analysis of the CYT-18 protein, mentioned above (Kittle et al., 1991). Deletion analysis of the mt LSU intron showed that the CYT-18 protein required an intact catalytic core to promote splicing (Guo et al., 1991). Nitrocellulose-filter binding and deletion analysis of the mt LSU intron defined P4-P6 as the smallest RNA sequence with high affinity for CYT-18 (K_d = 130 pM), while addition of sequences from the P3-P9 domain increased binding 6-fold (K_d = 22 pM) (Guo and Lambowitz, 1992, Saldanha et al., 1996). Iterative *in vitro* selection experiments with

the high affinity RNA molecule showed that critical nucleotides for CYT-18 binding were clustered around the junction of the P4-P6 stacked helix (Saldanha et al., 1996). The *td in vivo* plating system identified amino acids critical for splicing activity within the N-terminal domain of CYT-18 (Mohr et al., 2001). Finally, RNA footprinting experiments with the *N. crassa* mt LSU and *ND1* introns showed that CYT-18 protects regions of the catalytic core opposite the active site cleft (Caprara et al., 1996a). Most of the protection sites were in the P4-P6 domain, but additional sites were seen within the P3-P9 domain (Caprara et al., 1996a). These experiments support a model in which CYT-18 functions in group I splicing by binding to both of the major helical domains of the intron core, and stabilizes them in the correct conformation such that the active site is properly formed.

The findings that the CYT-18 protein is an aminoacyl-tRNA synthetase and uses the same or overlapping regions to bind both the tRNA and the group I intron suggested that the CYT-18 protein may have adapted to function in splicing by utilizing its pre-existing RNA binding domains. Guo and Lambowitz (1992) noted a high degree of similarity between the mt $tRNA^{Tyr}$ and core sequences of the mt LSU intron, particularly in the P7/variable arm regions, where 10 out of 14 nucleotides are identical. This is interesting as the *Neurospora* mt $tRNA^{Tyr}$ is a class II tRNA, characterized by long variable arms (Rich and RajBhandary, 1976). Michel and Westhof (1990) noted that the stacked helix formed by P6 and P4 is structurally

analogous to the tRNA helix formed by the stacking of the anticodon and D-arms, and is also stabilized by base-triples.

Caprara et al. (1996b) compared the CYT-18 binding sites in the *N. crassa* mt LSU and *ND1* introns with those in the *N. crassa* mt $tRNA^{Tyr}$. Using three-dimensional modeling, based on chemical modification and RNA-footprinting, an extended overlap of the phosphodiester backbones between the tRNA and the group I intron core was seen. In this overlap, the P4-P6 region of the intron core superimposed upon the D-anticodon stacked arms of the tRNA. Additionally, the P7 region partially overlapped the variable arm, and the 5' and 3' strands of the P9 region largely parallel the acceptor arm strands, with the discriminator base (A73) directly overlapping L9-1. These observations suggest that the CYT-18 protein may have been pre-disposed to function in group I intron splicing by recognizing conserved structural features in the group I intron core that resemble those found in tRNAs.

4.2 SPECIFIC AIMS

My goal was to define regions of proximity between the CYT-18 protein and its group I intron RNA substrates. Further, I wished to test the hypothesis that CYT-18 had adapted to function in splicing by using its pre-existing RNA binding domains. To meet these goals, I chose to study the CYT-18 intron complexes using directed hydroxyl-radical cleavage assays. In these assays, specific cysteine

substitutions were made in four clusters of consered amino acids involved in tRNA binding to determine if they were in proximity to the intron RNA.

4.3 EXPERIMENTAL STRATEGY

The experimental strategy was to construct a series of CYT-18 protein mutants having single cysteine substitutions at different positions for conjugation of EPD-Fe (Figure 4.1; Hall and Fox, 1999). In the presence of a reducing agent, the EPD-Fe conjugate produces short-lived hydroxyl-radicals that cleave the phosphodiester backbone of nearby RNA. The conjugated protein is bound to 5'-labeled RNA, and the locations of the cleavage sites are determined by gel electrophoresis, yielding information about the proximity of specific sites in the protein and RNA. Because the Fe is tethered by a 14 Å flexible linker and generates hydroxyl radicals with a 10 Å cleavage radius, cleavages are expected within ~24 Å of the conjugated cysteine residue (Hall and Fox, 1999). Previously, this method has proven useful in examining protein conformation (Ermacora et al., 1994, Ermacora et al., 1996), and protein-DNA interactions (Mazzarelli et al., 1993; Lavoie et al., 1996). This method has also been used to investigate RNA-protein interactions in RNase P (Biswas et al., 2000). A similar reagent, Fe-BABE has been used by Noller and co-workers to study protein-RNA interactions in the ribosome (Heilek et al., 1995; Heilek and Noller, 1996ab).

4.4 CONSTRUCTION OF A CYSTEINELESS CYT-18 PROTEIN

The CYT-18 protein is a favorable one for the incorporation of EPD-Fe because it contains only two endogenous cysteine residues. To obtain a cysteineless CYT-18 protein as a starting point for further modifications, we replaced these endogenous cysteines with the corresponding amino acids from the *P. anserina* mt TyrRS, which also functions in group I intron splicing (Kämper et al., 1992). Biochemical assays showed that the mutant proteins lacking one (C309P) or both (C309P:C494A) endogenous cysteines have wild-type tyrosyl-adenylation and TyrRS activity and that their EPD-Fe derivatives have wild-type splicing activity with both the *N. crassa* mt LSU and *ND1* introns (not shown).

Figure 4.2 shows hydroxyl radical cleavage assays for the wild-type and mutant proteins complexed with 5'-labeled *N. crassa* mt LSU intron RNA. The EPD-Fe derivatives of the wild-type and C309P proteins both gave specific cleavages, which were not seen with the protein lacking the two endogenous cysteines. The specific cleavages observed in Figure 4.2 map to three regions of the mt LSU intron: J6/6a-P6[3'](235-237), P6a[5'](148-150), and P5[5'](116-117), with the cleavages in J6/6a-P6 by far the strongest. In control reactions, the specific cleavages were abolished in the presence of 0.5 M KCl, which dissociates CYT-18/intron RNA complexes (lanes 4). Additionally, cleavages were competed by a two-fold excess of unmodified wild-type protein (lanes 3), and were not competed by a five-fold excess of non-specific RNA (lanes 5). The finding that the same cleavages were obtained

with the wild-type and C309P proteins indicates that all the cleavages result from EPD-Fe modification at C494, which is located in the C-terminal RNA-binding domain. The percentage of input RNA cleaved ranges from 4%, for the cleavage in the J6/6a-P6[3'] region, to 0.4% for the cleavage in P5[5'].

4.5 DIRECTED HYDROXY RADICAL CLEAVAGE ASSAYS

The CYT-18 protein lacking both endogenous cysteine residues was used as a starting point for constructing a series of single-cysteine substitutions in CYT-18. The cysteines were engineered in or near four regions of CYT-18 that contain clusters of conserved amino acid residues involved in tRNA binding in bacterial TyrRSs (clusters 1-4; see Figure 4.1; Bedouelle,1990; Nair et al., 1997). Clusters 1 and 2 in the nucleotide-binding fold region interact with the acceptor stem of $tRNA^{Tyr}$, and clusters 3 and 4 in the C-terminal domain interact with the variable and anticodon arms. The endogenous cysteine, C494, lies in cluster 3. Typically, we replaced non-conserved amino acids neighboring conserved basic amino acids identified as being involved in tRNA binding. In some cases, however, we replaced the conserved amino acid itself, the assumption being that loss of a single phosphate-backbone contact would not substantially impair the very tight binding of CYT-18 to the intron RNA (apparent K_d measured kinetically = <0.3 pM; Saldanha et al., 1996). In all, we constructed a total of 20 single-cysteine substitutions of which 17 gave active protein and 8 gave specific cleavages (see Figure 4.1). All the active proteins have wild-type tyrosyl-adenylation activity, and all but one retain 80-100% wild-type splicing

activity with both the LSU and *ND1* introns after conjugation with EPD-Fe (not shown). The single exception, EPD-Fe R540C, had 20-50% wild-type splicing activity with both introns. Representative hydroxyl radical cleavage assays are shown in Figures 4.3 and 4.4, and the locations of the cleavages in the intron RNAs are summarized in Figure 4.5.

Clusters 1 and 2 are expected to interact with the group I intron structure P9/L9, which is the cognate to the tRNA's acceptor stem. Four cluster 1 and 2 mutants gave specific EPD-Fe cleavages: R213C and R214C near the end of CP1 in cluster 1, A287C in cluster 2, and E296C just downstream of cluster 2 in a 16 amino acid CYT-18-specific insertion. All four proteins cleaved both introns in P9/L9, as expected, and more weakly in P5, an adjacent structure that is a docking site for the L9 tetraloop. EPD-Fe R213C and A287C also cleaved a neighboring region of P4 in the *ND1* intron. Other cluster 1 and 2 mutants, EPD-Fe V215C, S226C, K231C, K233C, R288C, E289C, and R297C gave no specific cleavages, even though their cysteine residues were accessible for modification and their EPD-Fe derivatives were fully active for splicing both the LSU and *ND1* intron. Although EPD-Fe cleavages in the *N. crassa* mt $tRNA^{Tyr}$ could not be mapped precisely because of the relatively weak binding and small size of the RNA, additional experiments showed that all four cluster 1 and 2 mutants that cleaved the intron RNAs also cleaved the tRNA, and those that failed to cleave the intron RNAs also failed to cleave the tRNA (not shown). Thus, all of the tested clusters 1 and 2 positions are similarly disposed to

cleave or not cleave both the intron and tRNA substrates, and all four mutants that cleaved the intron did so in P9/L9, the cognate of the acceptor stem, and to a lesser extent in the adjacent structure P5.

Clusters 3 and 4, in the C-terminal RNA-binding domain, are expected to interact with the group I intron's P4/P6 stacked helices, the cognate of the tRNA's D-anticodon arm stacked helices, and with P7, the cognate of the variable arm. In addition to the endogenous cysteine C494 in the C309P mutant (Figure 4.2), three other cluster 3 and 4 mutants gave specific cleavages. EPD-Fe-G493C, which is adjacent to the endogenous C494 in cluster 3, cleaved the LSU intron at the same locations as C309P in J6/6a-P6[3'], P6a[5'], and P5[5'], as well as additional sites in P4[3'] and P2[3']. Both C309P and G493C cleaved the *ND1* intron in P8[3']-J8/7, P3[3'], J6/6a-P6[3'], and P6a[5']. The remaining two mutants, G497C in cluster 3 and R540C, the only active variant in cluster 4, cleaved the LSU intron but not the *ND1* intron, likely reflecting more extensive interaction of the C-terminal domain with the LSU intron. Both mutants cleaved the LSU intron around the junction of P4[3'] and P6[5'], and EPD-Fe R540C also cleaved the LSU intron in P6a[5'] and P6[3']. Thus, as expected, the cluster 3 and 4 mutants cleaved primarily in P4-P6 domain, the cognate of the D-anticodon arm stacked helices of $tRNA^{Tyr}$, with a few, mostly weaker cleavages in other regions.

4.6 STRUCTURE MODELLING

The positional information from the EPD-Fe cleavage experiments enabled us to orient the protein on the RNA for structure modeling. Figure 4.6A and B shows models of the complexes between CYT-18 and the mt LSU and *ND1* introns, respectively, based on the cocrystal structure of the *T. thermophilus* TyrRS/$tRNA^{Tyr}$ complex, which is shown in Figure 4.6C for comparison. The homologous regions of the *T. thermophilus* TyrRS and CYT-18 are expected to have similar three-dimensional structures, although CYT-18 has several small insertions and a longer C-terminal domain, which are not represented in the models. The models of the complexes were constructed without reference to the biochemical data by simply aligning the group I intron RNA with the $tRNA^{Tyr}$ in the *T. thermophilus* cocrystal structure using the previously described three-dimensional overlaps between group I introns and tRNAs (Caprara et al., 1996b; see section **2.2.8**). Auspiciously, the orientation of the tRNA's long variable arm in the X-ray crystal structure more closely parallels P7 than it did in the previous models, which were based on $tRNA^{Ser}$.

The models show that the group I introns, like $tRNA^{Tyr}$, bind across the surface of the two subunits of the homodimer, interacting with the N-terminal domain of one subunit and the C-terminal domain of the other. In agreement with the predicted group I intron/$tRNA^{Tyr}$ structural homologies, the C-terminal domain interacts primarily with the group I introns' P4-P6 stacked helices, the cognate of the tRNA's D-anticodon arm stacked helices, and the nucleotide-binding fold interacts

with P9/L9, the cognate of the tRNA's acceptor stem, as well as P5, which interacts with L9. As noted previously, the discriminator base and L9-1 are putative contact sites in both RNA substrates and overlap in the alignments. The group I introns of course lacks the 3' CCA, which extends into the TyrRS active site. The models and biochemical data suggest that there could also be interactions between the C-terminal domain and P7, the cognate of the long variable arm, and between the α-helical intermediate domain and J6/6a, the cognate of the anticodon loop. The extended C-terminal domain of CYT-18 may also interact with P2, P3, and P8, which do not have cognates in the tRNA.

Although the models are based entirely on the group I intron/$tRNA^{Tyr}$ alignment, there is nevertheless good agreement with biochemical data. Thus, 35/40 EPD-Fe cleavages (red and magenta regions) fall within 33 Å of the derivatized amino acid (yellow) and all strong cleavages fall within 28 Å. The only cleavages that cannot be explained by minor adjustments in the model are the weak cleavages from two cluster 3 positions (G493C and the endogenous C494 in C309P) in P5[5'] in the LSU intron. These cleavages are not observed in the more compact *ND1* intron and could reflect the conformational mobility of the flexibly hinged C-terminal domain in either subunit of the homodimer. If not, they require either a conformational change in the intron RNA or that the interacting C-terminal domain in the model be shifted up from the position in the cocrystal structure. At least some shift in the C-terminal domain appears necessary in any case to accommodate the

bulkier intron RNA in this region and to account for the lack of observed cleavages in P7, the cognate of the variable arm.

Also in agreement with the models, most of the phosphate-backbone protections mapped previously in iodine footprinting experiments (gray spheres) are in regions protected by the protein in the model. Importantly, the phosphate-backbone protections around the junction of the P4-P6 stacked helices are potentially superimposable with those at the junction of the D-anticodon arm stacked helices, with the protein facing the minor groove of P4 and the major groove of P6, as for the cognate structures in the tRNA (Caprara et al., 1996b, 2001). These findings strongly suggest a very similar disposition of the α-helical and C-terminal domains with respect to both the intron and tRNA substrates in this region. Except for two P5 protections in the *ND1* intron, the only unaccounted for phosphate protections in the models are in P8 in both introns and in P6a in the LSU intron and can be explained by additional interactions with CYT-18's larger C-terminal domain and/or some shift in the position of the C-terminal domain in the complex. We note that the biochemical data show more extensive C-terminal contacts with the LSU intron than with the *ND1* intron, consistent with the observation that the C-terminal domain plays a greater role in splicing the LSU intron than the *ND1* intron (Mohr et al., 2001).

4.7 DISCUSSION

Our results support the hypothesis that CYT-18 functions in splicing by recognizing highly conserved structural features of group I intron that resemble those

in tRNAs. The hydroxyl radical cleavage data identify putative interaction sites between the nucleotide-binding fold and the P9/P5 junction region, and between the C-terminal domain and the P4-P6 stacked helices. With these putative interacting regions fixed, the structural models based on the *T. thermophilus* TyrRS/tRNA^{Tyr} cocrystal structure indicate a decidedly tRNA-like recognition of the group I intron catalytic core, with the acceptor-arm cognate (P9/L9) interacting with the nucleotide-binding fold of one subunit, and the D-anticodon arm cognate (P4-P6) interacting with the C-terminal domain of the other subunit. The relatively large interface between the protein and intron RNA affords the potential for multiple contacts, as in tRNA binding, and explains how CYT-18 can suppress structural mutations throughout the group I intron catalytic core (Mohr et al., 1992; Myers et al., 1996; Chen et al., 2000). The ability of CYT-18 to function in splicing many different group I introns likely reflects that the tRNA-like structural features with which it interacts are highly conserved in group I introns, presumably because they are required for the catalytic activity of the intron RNAs. Within this framework of conserved interactions, additional intron-specific and non-tRNA-like interactions may additionally contribute to CYT-18-dependent splicing.

With respect to splicing mechanism, our previous biochemical analysis suggested that CYT-18 interacts first with the P4-P6 domain to promote its assembly and then makes secondary contacts with the P3-P9 domain to bring the two major domains into the correct relative orientation to form the intron's active site (Caprara et

a., 1996ab). Regarding the first step, the models show that the region of the P4-P6 domain around the junction of the P4-P6 stacked helices interacts with a protein scaffold formed by the C-terminal and α-helical domains, which evolved to recognize the D-anticodon arm of the tRNA. The use of this protein scaffold readily explains how CYT-18 can induce assembly of the P4-P6 domain, as well as compensate for a variety of structural mutations that impair base-stacking, base-pairing, or base-triple interactions around the junction of the P4-P6 stacked helices (Myers et al., 1996, Chen et al., 2000). For the second step, the models suggest that CYT-18 could help establish the correct relative orientation of the two domains by stabilizing or substituting for critical interdomain interactions, including the P5/L9 tetraloop/receptor interaction, J3/4 and J6/7 interdomain contacts, and the minor groove interdigitation between P3 and J6/J6a (Michel and Westhof, 1990; Golden et al., 1998). Consistent with this hypothesis, genetic assays with the phage T4 *td* intron show that CYT-18 can suppress mutations at all positions in J3/4 and J6/7 (Myers et al., 1996), as well as mutations that disrupt the P5-L9 interaction (X. Chen and A.M.L., unpublished). The L9-P5 interaction is required for CYT-18 dependent splicing of the *Neurospora* mt LSU intron and deletion of this region from the *Tetrahymena* intron greatly destabilizes the intron (Guo et al., 1991; Caprara and Waring, 1994). Protection of P5 by P9 occurs early in the folding of the P3-P9 domain, suggesting that this interaction may also be important for the folding

pathway, in addition to stabilizing the active intron structure (Sclavi et al., 1998; Golden et al., 1998).

The models also explain previous findings that the N- and C-terminal domains of CYT-18 contribute differently to splicing different group I introns and in particular that deletion of the C-terminal RNA-binding domain abolished splicing and stable binding of the LSU intron, but left substantial splicing activity with the *ND1* intron and other group I introns (Mohr et al., 2001). The models show that this truncated protein can still interact extensively with the *ND1* intron via the nucleotide-binding fold and α-helical domains, with the latter responsible for many of the phosphate-backbone protections in P4-J6/6a. The greater interaction of the LSU intron with the C-terminal domain is reflected by the additional EPD-Fe cleavages from positions in clusters 3 and 4 and by the larger number of phosphate-backbone interactions extending farther down the P6 helix into P6a. We note that even though the C-terminal truncation did not abolish splicing of the *ND1* intron, it increased k_{off} by about fivefold, suggesting that the C-terminal domain makes a non-critical contribution to binding (Mohr et al., 2001). Further, group I introns that are not ordinarily dependent on the C-terminal domain can readily acquire such dependence by mutations in certain regions of the catalytic core (X. Chen and A.M.L., unpublished). Together, these findings suggest that all group I introns form fundamentally similar complexes involving the N- and C-terminal domains of CYT-

18 and that dependence on the C-terminal domain reflects primarily the nature of the specific structural deficiencies that impair self-splicing in different introns.

We suggested previously that group I introns were initially self-splicing and became dependent on cellular proteins as a result of mutations that impaired self-splicing (Lambowitz and Perlman,1990; Lambowitz et al., 1999). Although it is not yet clear to what extent other aaRSs function in group I intron splicing, it is possible that TyrRS has structural features that make it uniquely suited for such adaptation. Such features may include the ability to recognize a tRNA with a long variable arm in a specific orientation that matches P7, and the flexibly hinged C-terminal domain, which may facilitate the binding of some non-tRNA substrates. The latter may include not only group I introns, but also the 3'-terminal tRNA-like structures of RNA viruses (Mans et al., 1991). In addition, CYT-18 contains several idiosyncratic insertions and an extended C-terminal domain, which may contribute to splicing via unique interactions with group I introns.

The yeast mt LeuRS, which functions in splicing the closely related mt group I introns bI4 and aI4, may also recognize tRNA-like structural features of these introns (Herbert et al., 1988; Houman et al., 2000). Further, the ability to splice these introns is inherent in other mt or bacterial LeuRSs, which can fully complement a yeast null mutant lacking the mt LeuRS (Houman et al., 2000). Unlike CYT-18, which functions in splicing many different group I introns, the LeuRS functions only in splicing the two closely related yeast mt DNA introns and does so by acting in

concert with a maturase encoded by one of the introns. Thus, the LeuRS must recognize either idiosyncratic features of bI4 and aI4 or a subset of the conserved tRNA-like features in a manner not sufficient to promote splicing in the absence of the maturase.

Finally, the alignment with the crystal structure further supports the conclusion that group I introns have structural similarities to tRNAs. As discussed previously, these structural similarities could be coincidental, or they could reflect an evolutionary relationship between group I introns and tRNAs. One possibility is that group I introns evolved in the RNA world and gave rise to tRNAs during the evolution of protein synthesis (Guo and Lambowitz, 1992). Another possibility is that a tRNA or tRNA-like structure evolved into a group I intron by acquiring catalytic activity, which enabled it to propagate as a mobile element by reverse splicing into other RNA sites (Caprara et al., 1996b). Both possibilities are remarkable.

4.7 SUMMARY

In these experiments, I used directed hydroxyl radical cleavage assays to show that the nucleotide-binding fold and C-terminal domains of CYT-18 interact with the expected group I intron cognates of the aminoacyl-acceptor stem and D-anticodon arms, respectively. Further, three-dimensional graphic modeling, supported by biochemical data, shows that conserved regions of group I introns can be superimposed over interacting regions of the tRNA in a *Thermus thermophilus*

TyrRS/tRNA^{Tyr} cocrystal structure. These results support the hypothesis that CYT-18 and other aminoacyl-tRNA synthetases interact with group I introns by recognizing conserved tRNA-like structural features of the intron RNAs.

Figure 4.1: Map of the CYT-18 protein and location of single-cysteine replacements.

Map of the CYT-18 protein showing the location of single cysteine substitutions for incorporation of EPD-Fe. Black boxes indicate regions strongly conserved among bacterial TyrRSs, and stippled boxes indicate regions conserved only between the *N. crassa* and *P. anserina* mt TyrRSs, which function in group I intron splicing (Mohr et al., 2001). CYT-18 variants having single cysteine substitutions for conjugation of EPD-Fe at different positions are shown above. Asterisks indicate variant proteins that's EPD-Fe-derivatives gave specific cleavages in the intron RNAs, and brackets indicate proteins that were largely insoluble. The remaining proteins were soluble and active, but did not give specific cleavages. The boundaries of the nucleotide-binding fold, α-helical intermediate domain, and C-terminal RNA-binding domain based on homology to other TyrRSs are shown below (Mohr et al., 2001). Black bars indicate boundaries of CP1 and clusters 1-4, which contain conserved amino acid residues involved in tRNA binding in bacterial TyrRSs, with their predicted $tRNA^{Tyr}$ interaction sites indicated below (Bedouelle, 1990; Nair et al., 1997).

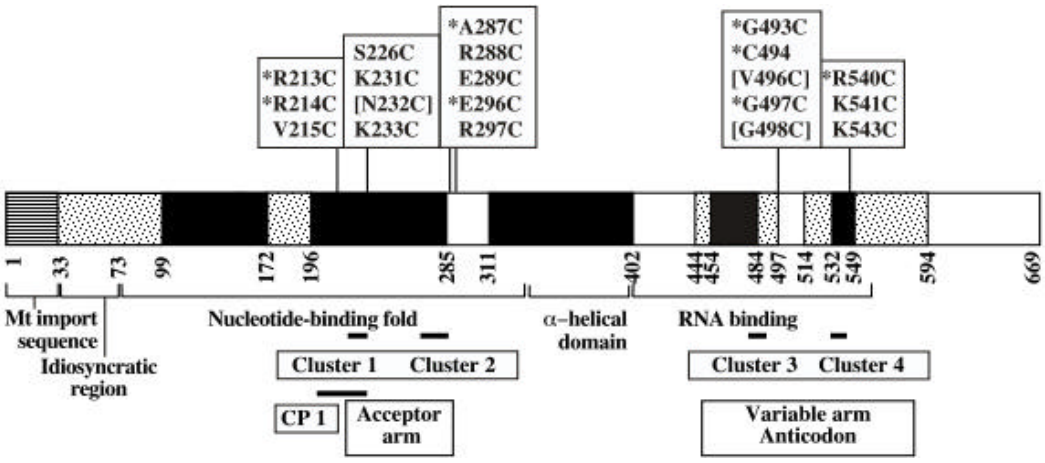


Figure 4.2: Directed hydroxyl radical cleavage assay.

Directed hydroxyl radical cleavage of the *N. crassa* mt LSU intron with EPD-Fe derivatives of wild-type and variant CYT-18 proteins lacking one or both endogenous cysteine residues. A 5'-labeled *in vitro* transcript containing the *N. crassa* mt LSU intron was complexed with the indicated EPD-Fe-conjugated CYT-18 proteins, and cleavage sites were mapped in a denaturing 6% polyacrylamide/8 M urea gel. Lanes: OH^-, alkaline hydrolysis ladder; G and A, RNA sequencing ladders; lanes 1, EPD-Fe modified protein incubated with RNA in the absence of ascorbate to initiate hydroxyl radical cleavage; lanes 2, cleavage reactions after addition of ascorbate; lanes 3, cleavage reactions in the presence of a two-fold molar excess of unmodified wild-type CYT-18 protein; lanes 4, cleavage reactions in 0.5 M KCl; lanes 5, cleavage reactions in the presence of a five-fold molar excess of non-specific competitor RNA (pTRSE15/*Nsi*I; Chen and Lambowitz, 1997). Regions of the mt LSU intron are shown to the left, and the cleavages sites are indicated to the right.

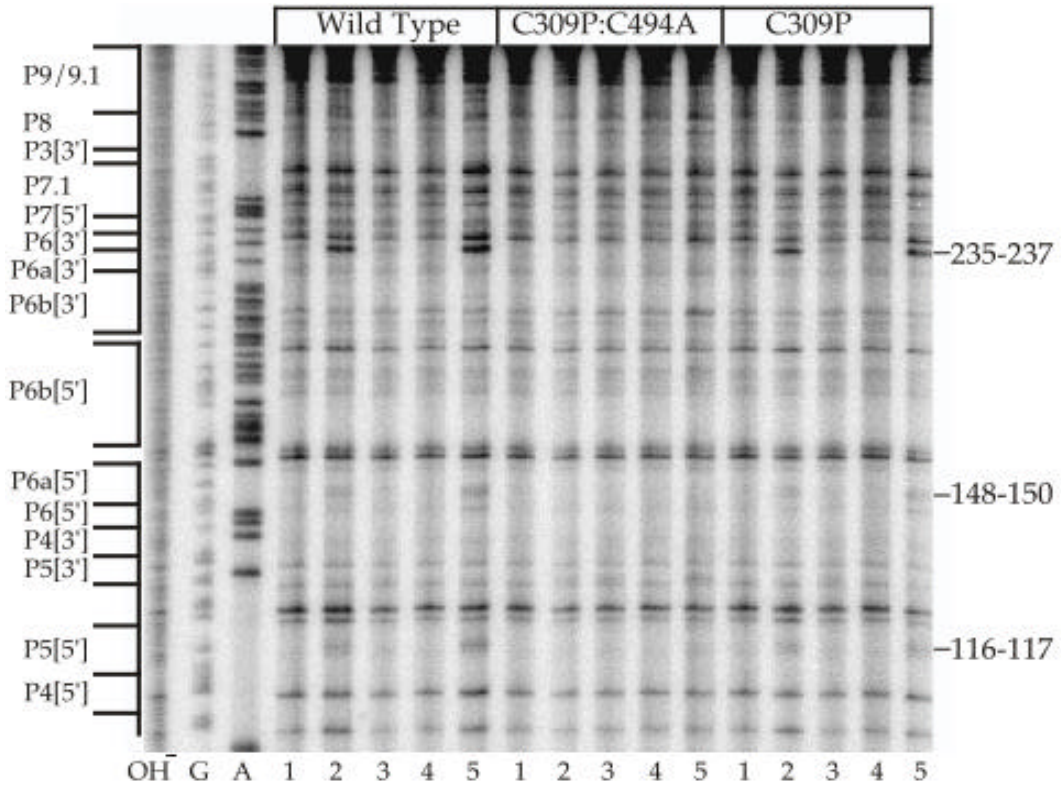


Figure 4.3: Directed hydroxyl radical probing of the *N. crassa* mt LSU intron.

Directed hydroxyl radical probing of the *N. crassa* mt LSU intron with variant CYT-18 proteins containing EPD-Fe conjugated to different positions. Reactions and gel lanes are as in Figure 4.2. Mutant proteins are indicated above, and cleavage sites are indicated to the right.

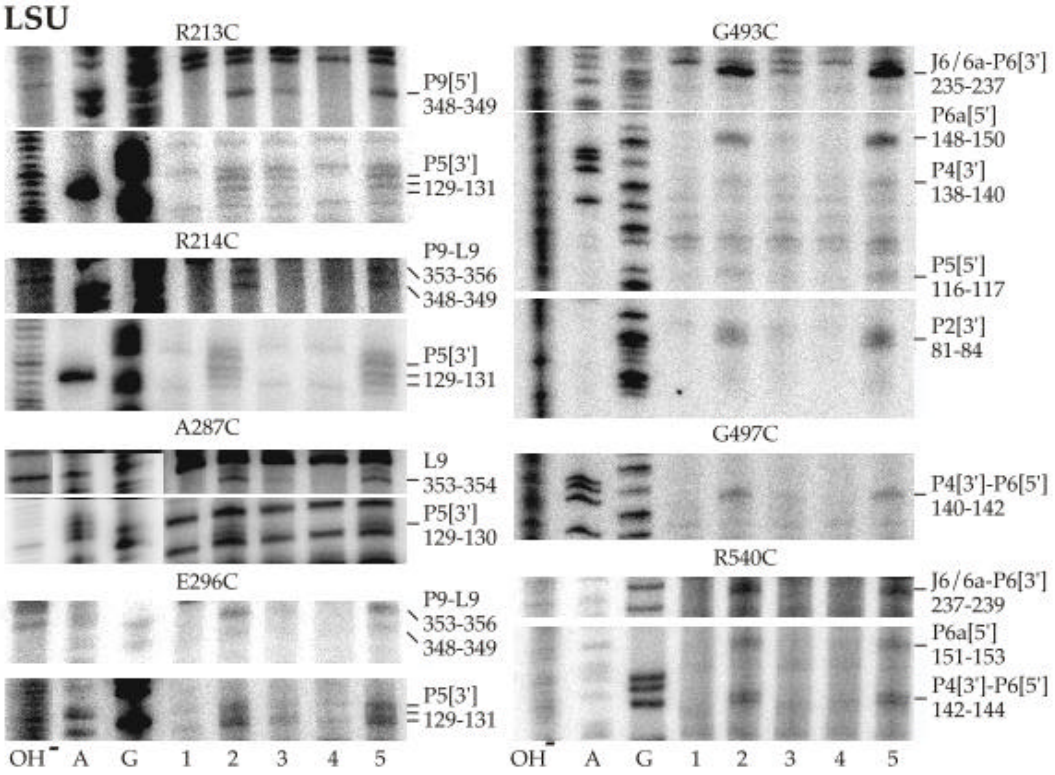


Figure 4.4: Directed hydroxyl radical probing of the *N. crassa ND1* intron.

Directed hydroxyl radical probing of the *N. crassa ND1* intron with variant CYT-18 proteins containing EPD-Fe conjugated to different positions. Reactions and gel lanes are as in Figure 4.2. Mutant proteins are indicated above, and cleavage sites are indicated to the right.

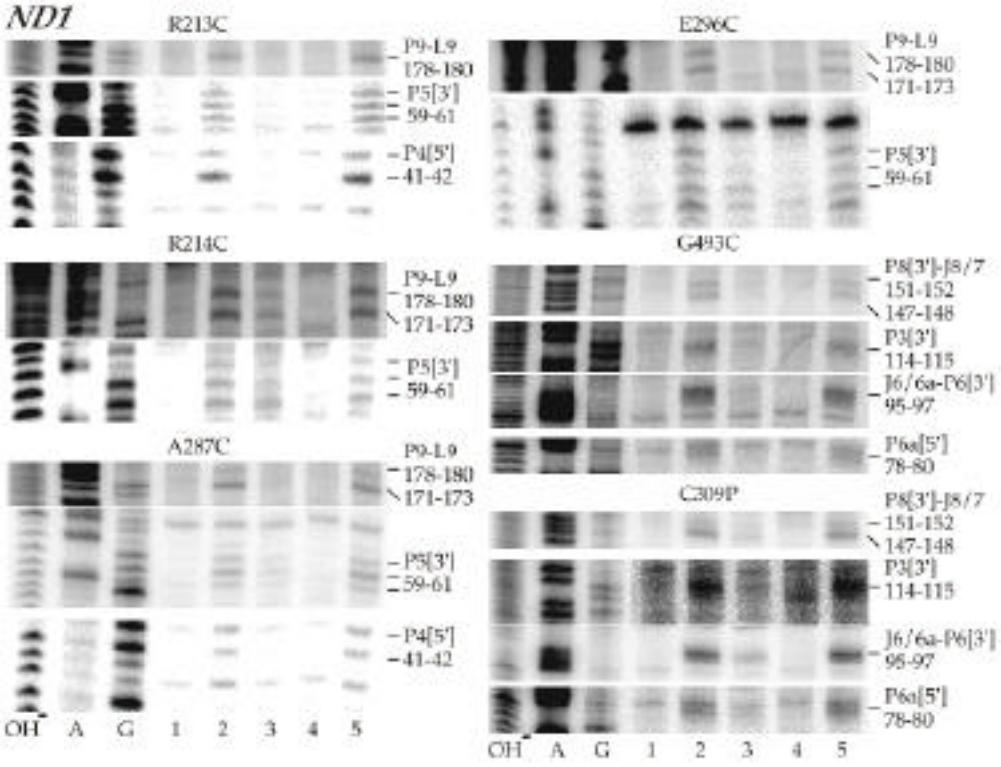


Figure 4.5: Summary of directed hydroxyl radical cleavages in the *N. crassa* mt LSU and *ND1* introns.

(A) and (B) show the predicted secondary structures of the 388-nt mt LSU and 197-nt *ND1* introns used in this study, respectively. Cleavage from EPD-Fe conjugated to different positions in the N- and C-terminal domains of CYT-18 are indicated in red and magenta, respectively. Thin, connecting lines between boxes indicate tertiary interactions. Arrows indicate splice sites. (C) shows the secondary structure of *N. crassa* mt $tRNA^{Tyr}$ drawn in a similar orientation for comparison.
Because of the University of Texas at Austin thesis requirements, this figure is too small to be useful. I suggest checking the original references.

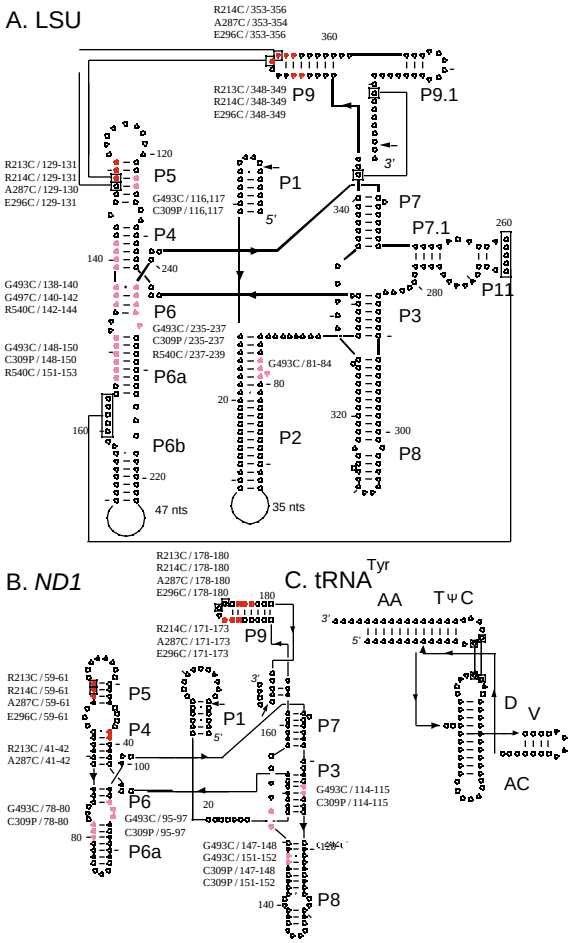
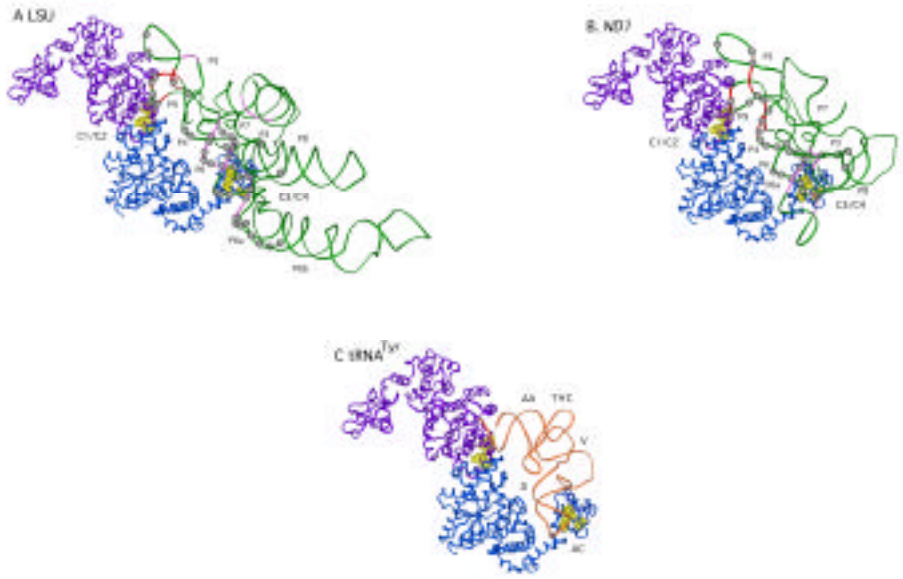


Figure 4.6: Models of the *N. crassa* mt LSU and *ND1* introns complexed with the *T. thermophilus* TyrRS.

The models are shown in (A) and (B), with a drawing of the *T. thermophilus* TyrRS/tRNA^{Tyr} cocrystal structure shown in (C) for comparison (A. Yaremchuk, I. Kriklivyi, M. Tukalo, and S. Cusack, submitted). The two TyrRS subunits are blue and purple, and the RNA phosphodiester backbone is green for the introns and orange for the tRNA. EPD-Fe-conjugated amino acid residues that gave specific cleavages in the intron RNAs are highlighted in yellow, and the cleaved regions of the phosphodiester backbone are shown in red and magenta, respectively. E296C, a residue in a CYT-18-specific insertion, is represented by yellow highlighting of the closest amino acid residue (S290). Gray balls indicate putative protein-protection sites on the phosphodiester backbone detected in iodine footprinting experiments (Caprara et al., 1996ab). Protection sites likely due to protein-induced RNA structural changes in P3[5'] in both introns and J4/5, P7[3'], P7.1, and P6b in the LSU intron are not shown. Abbreviations: C1/C2, clusters 1 and 2; C3/C4, clusters 3 and 4.

Because of the University of Texas at Austin thesis requirements, this figure is too small to be useful. I suggest checking the original references.



Bibliography

Akins, R.A. and Lambowitz, A.M. 1987. A protein required for splicing group I introns in Neurospora mitochondria is mitochondrial tyrosyl-tRNA synthetase or a derivative thereof. Cell *50*(3), 331-345.

Arnez, J.G. and Moras, D. 1997. Structural and functional considerations of the aminoacylation reaction. Trends Biochem. Sci. *22*(6), 211-216.

Auld, D.S. and Schimmel, P. 1995. Switching recognition of two tRNA synthetases with an amino acid swap in a designed peptide. Science *267*(5206), 1994-1996.

Auld, D.S. and Schimmel, P. 1996. Single sequence of a helix-loop peptide confers functional anticodon recognition on two tRNA synthetases. EMBO J *15*(5), 1142-1148.

Bass, B.L. and Cech, T.R. 1986. Ribozyme inhibitors: deoxyguanosine and dideoxyguanosine are competitive inhibitors of self-splicing of the Tetrahymena ribosomal ribonucleic acid precursor. Biochemistry *25*(16), 4473-4477.

Bedouelle, H. 1990. Recognition of tRNA(Tyr) by tyrosyl-tRNA synthetase. Biochimie *72*(8), 589-598.

Belfort, M., Moelleken, A., Maley, G.F., and Maley, F. 1983. Purification and properties of T4 phage thymidylate synthetase produced by the cloned gene in an amplification vector. J. Biol. Chem. *258*(3), 2045-2051.

Belfort, M., Chandry, P.S., and Pedersen-Lane, J. 1987. Genetic delineation of functional components of the group I intron in the phage T4 *td* gene. Cold Spring Harb. Symp. Quant. Biol. *52*, 181-192.

Bell-Pedersen, D., Galloway Salvo, J.L., and Belfort, M. 1991. A transcription terminator in the thymidylate synthase (*thyA*) structural gene of Escherichia coli and construction of a viable *thyA*::Km^r deletion. J. Bacteriol. *173*(3), 1193-1200.

Biswas, R., Ledman, D.W., Fox, R.O., Altman, S., and Gopalan, V. 2000. Mapping RNA-protein interactions in ribonuclease P from Escherichia coli using disulfide-linked EDTA-Fe. J. Mol. Biol. *296*(1), 19-31.

Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. *72*, 248-254.

Burke, J.M., Irvine, K.D., Kaneko, K.J., Kerker, B.J., Oettgen, A.B., Tierney, W.M., Williamson, C.L., Zaug, A.J., and Cech, T.R. 1986. Role of conserved sequence elements 9L and 2 in self-splicing of the Tetrahymena ribosomal RNA precursor. Cell *45*(2), 167-176.

Burke, J.M., Belfort, M., Cech, T.R., Davies, R.W., Schweyen, R.J., Shub, D.A., Szostak, J.W., and Tabak, H.F. 1987. Structural conventions for group I introns. Nucleic Acids Res. *15*(18), 7217-7221.

Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J.R., D'Souza, L.M., Du, Y., Feng, B., Lin, N., Madabusi, L.V., Muller, K.M., Pande, N., Shang, Z., Yu, N., and Gutell, R.R. 2002. The Comparative RNA Web (CRW) Site: An Online Database of Comparative Sequence and Structure Information for Ribosomal, Intron, and other RNAs. BioMed Central Bioinformatics *3*(2), 2-66.

Caprara, M.G. and Waring, R.B. 1994. Deletion of P9 and stem-loop structures downstream from the catalytic core affects both 5' and 3' splicing activities in a group-I intron. Gene *143*(1), 29-37.

Caprara, M.G., Mohr, G., and Lambowitz, A.M. 1996. A tyrosyl-tRNA synthetase protein induces tertiary folding of the group I intron catalytic core. J. Mol. Biol. *257*(3), 512-531.

Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.R., and Doudna, J.A. 1996. Crystal structure of a group I ribozyme domain: principles of RNA packing. Science *273*(5282), 1678-1685.

Cech, T.R., Zaug, A.J., and Grabowski, P.J. 1981. In vitro splicing of the ribosomal RNA precursor of Tetrahymena: involvement of a guanosine nucleotide in the excision of the intervening sequence. Cell *27*(3 Pt 2), 487-496.

Cech, T.R. 1985. Self-splicing RNA: implications for evolution. Int. Rev. Cytol. *93*, 3-22.

Cech, T.R., Damberger, S.H., and Gutell, R.R. 1994. Representation of the secondary and tertiary structure of group I introns. Nat. Struct. Biol. *1*(5), 273-280.

Celander, D.W. and Cech, T.R. 1991. Visualizing the higher order folding of a catalytic RNA molecule. Science *251*(4992), 401-407.

Chandry, P.S. and Belfort, M. 1987. Activation of a cryptic 5' splice site in the upstream exon of the phage T4 *td* transcript: exon context, missplicing, and mRNA deletion in a fidelity mutant. Genes Dev. *1*(9), 1028-1037.

Chang, A.C. and Cohen, S.N. 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. *134*(3), 1141-1156.

Chen, B. and Lambowitz, A.M. 1997. De novo and DNA primer-mediated initiation of cDNA synthesis by the mauriceville retroplasmid reverse transcriptase involve recognition of a 3' CCA sequence. J. Mol. Biol. *271*(3), 311-332.

Chen, X., Gutell, R.R., and Lambowitz, A.M. 2000. Function of tyrosyl-tRNA synthetase in splicing group I introns: an induced-fit model for binding to the P4-P6 domain based on analysis of mutations at the junction of the P4-P6 stacked helices. J. Mol. Biol. *301*(2), 265-283.

Cherniack, A.D., Garriga, G., Kittle, J.D., Jr., Akins, R.A., and Lambowitz, A.M. 1990. Function of Neurospora mitochondrial tyrosyl-tRNA synthetase in RNA splicing requires an idiosyncratic domain not found in other synthetases. Cell *62*(4), 745-755.

Cormack, B., Site-directed mutagenesis by the polymerase chain reaction., in *Short Protocols in Molecular Biology*, F.M. Ausubel, et al., Editors. 1992, John Wiley and Sons: New York. p. 8.18-18.25.

Doudna, J.A. and Szostak, J.W. 1989. Miniribozymes, small derivatives of the sunY intron, are catalytically active. Mol. Cell Biol. *9*(12), 5480-5483.

Dujon, B. 1989. Group I introns as mobile genetic elements: facts and mechanistic speculations--a review. Gene *82*(1), 91-114.

Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.P., and Ehresmann, B. 1987. Probing the structure of RNAs in solution. Nucleic Acids Res. *15*(22), 9109-9128.

Ellman, G.L. 1959. Tissue Sulfhydryl Groups. Archives of Biochemistry and Biophysics *82*, 70-77.

Ermacora, M.R., Ledman, D.W., Hellinga, H.W., Hsu, G.W., and Fox, R.O. 1994. Mapping staphylococcal nuclease conformation using an EDTA-Fe derivative attached to genetically engineered cysteine residues. Biochemistry *33*(46), 13625-13641.

Ermacora, M.R., Ledman, D.W., and Fox, R.O. 1996. Mapping the structure of a non-native state of staphylococcal nuclease. Nat. Struct. Biol. *3*(1), 59-66.

Ferre-D'Amare, A.R., Zhou, K., and Doudna, J.A. 1998. A general module for RNA crystallization. J. Mol. Biol. *279*(3), 621-631.

Gaillard, C. and Bedouelle, H. 2001. An essential residue in the flexible peptide linking the two idiosynchratic domains of bacterial tyrosyl-tRNA synthetases. Biochemistry *40*(24), 7192-7199.

Galloway Salvo, J.L., Coetzee, T., and Belfort, M. 1990. Deletion-tolerance and trans-splicing of the bacteriophage T4 *td* intron. Analysis of the P6-L6a region. J. Mol. Biol. *211*(3), 537-549.

Gampel, A., Nishikimi, M., and Tzagoloff, A. 1989. CBP2 protein promotes in vitro excision of a yeast mitochondrial group I intron. Mol. Cell Biol. *9*(12), 5424-5433.

Golden, B.L., Gooding, A.R., Podell, E.R., and Cech, T.R. 1998. A preorganized active site in the crystal structure of the Tetrahymena ribozyme. Science *282*(5387), 259-264.

Guo, Q. and Lambowitz, A.M. 1992. A tyrosyl-tRNA synthetase binds specifically to the group I intron catalytic core. Genes Dev. *6*(8), 1357-1372.

Guo, Q.B., Akins, R.A., Garriga, G., and Lambowitz, A.M. 1991. Structural analysis of the Neurospora mitochondrial large rRNA intron and construction of a mini-intron that shows protein-dependent splicing. J. Biol. Chem. *266*(3), 1809-1819.

Hall, K.B. and Fox, R.O. 1999. Directed cleavage of RNA with protein-tethered EDTA-Fe. Methods *18*(1), 78-84.

Heilek, G.M., Marusak, R., Meares, C.F., and Noller, H.F. 1995. Directed hydroxyl radical probing of 16S rRNA using Fe(II) tethered to ribosomal protein S4. Proc. Natl. Acad. Sci. USA *92*(4), 1113-1116.

Heilek, G.M. and Noller, H.F. 1996. Directed hydroxyl radical probing of the rRNA neighborhood of ribosomal protein S13 using tethered Fe(II). RNA *2*(6), 597-602.

Henke, R.M., Butow, R.A., and Perlman, P.S. 1995. Maturase and endonuclease functions depend on separate conserved domains of the bifunctional protein encoded by the group I intron aI4 alpha of yeast mitochondrial DNA. EMBO J. *14*(20), 5094-5099.

Herbert, C.J., Labouesse, M., Dujardin, G., and Slonimski, P.P. 1988. The NAM2 proteins from S. cerevisiae and S. douglasii are mitochondrial leucyl-tRNA synthetases, and are involved in mRNA splicing. EMBO J. *7*(2), 473-483.

Heuer, T.S., Chandry, P.S., Belfort, M., Celander, D.W., and Cech, T.R. 1991. Folding of group I introns from bacteriophage T4 involves internalization of the catalytic core. Proc. Natl. Acad. Sci. USA *88*(24), 11105-11109.

Hill, J., McGraw, P., and Tzagoloff, A. 1985. A mutation in yeast mitochondrial DNA results in a precise excision of the terminal intron of the cytochrome b gene. J. Biol. Chem. *260*(6), 3235-3238.

Houman, F., Rho, S.B., Zhang, J., Shen, X., Wang, C.C., Schimmel, P., and Martinis, S.A. 2000. A prokaryote and human tRNA synthetase provide an essential RNA splicing function in yeast mitochondria. Proc. Natl. Acad. Sci. USA *97*(25), 13743-13748.

Ikawa, Y., Shiraishi, H., and Inoue, T. 2000. Minimal catalytic domain of a group I self-splicing intron RNA. Nat. Struct. Biol. *7*(11), 1032-1035.

Kämper, U., Kück, U., Cherniack, A.D., and Lambowitz, A.M. 1992. The mitochondrial tyrosyl-tRNA synthetase of *Podospora anserina* is a bifunctional enzyme active in protein synthesis and RNA splicing. Mol. Cell Biol. *12*(2), 499-511.

Kittle, J.D., Jr., Mohr, G., Gianelos, J.A., Wang, H., and Lambowitz, A.M. 1991. The Neurospora mitochondrial tyrosyl-tRNA synthetase is sufficient for group I intron splicing in vitro and uses the carboxy-terminal tRNA- binding domain along with other regions. Genes Dev. *5*(6), 1009-1021.

Kunkel, T.A., Roberts, J.D., and Zakour, R.A. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. *154*, 367-382.

Labouesse, M. 1990. The yeast mitochondrial leucyl-tRNA synthetase is a splicing factor for the excision of several group I introns. Mol. Gen. Genet. *224*(2), 209-221.

Laggerbauer, B., Murphy, F.L., and Cech, T.R. 1994. Two major tertiary folding transitions of the Tetrahymena catalytic RNA. EMBO J. *13*(11), 2669-2676.

Lambowitz, A.M. and Perlman, P.S. 1990. Involvement of aminoacyl-tRNA synthetases and other proteins in group I and group II intron splicing. Trends Biochem. Sci. *15*(11), 440-444.

Lambowitz, A.M., Caprara, M.G., Zimmerly, S., and Perlman, P.S., Group I and Group II Introns: Keys to the Future, Clues to the Past, in *The RNA World*, R.F. Gesteland, T.R. Cech, and J.F. Atkins, Editors. 1999, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New Yok. p. 451-485.

Lavoie, B.D., Shaw, G.S., Millner, A., and Chaconas, G. 1996. Anatomy of a flexer-DNA complex inside a higher-order transposition intermediate. Cell *85*(5), 761-771.

Li, G.Y., Becam, A.M., Slonimski, P.P., and Herbert, C.J. 1996. In vitro mutagenesis of the mitochondrial leucyl tRNA synthetase of Saccharomyces cerevisiae shows that the suppressor activity of the mutant proteins is related to the splicing function of the wild-type protein. Mol. Gen. Genet. *252*(6), 667-675.

Li, G.Y., Tian, G.L., Slonimski, P.P., and Herbert, C.J. 1996. The CBP2 gene from Saccharomyces douglasii is a functional homologue of the Saccharomyces cerevisiae gene and is essential for respiratory growth in the presence of a wild-type (intron-containing) mitochondrial genome. Mol. Gen. Genet. *250*(3), 316-322.

Lin, L., Hale, S.P., and Schimmel, P. 1996. Aminoacylation error correction. Nature *384*(6604), 33-34.

Lusk, J.E., Williams, R.J., and Kennedy, E.P. 1968. Magnesium and the growth of Escherichia coli. J. Biol. Chem. *243*(10), 2618-2624.

Mans, R.M., Pleij, C.W., and Bosch, L. 1991. tRNA-like structures. Structure, function and evolutionary significance. Eur. J. Biochem. *201*(2), 303-324.

Mazzarelli, J.M., Ermacora, M.R., Fox, R.O., and Grindley, N.D. 1993. Mapping interactions between the catalytic domain of resolvase and its DNA substrate using cysteine-coupled EDTA-iron. Biochemistry *32*(12), 2979-2986.

Michel, F., Hanna, M., Green, R., Bartel, D.P., and Szostak, J.W. 1989. The guanosine binding site of the Tetrahymena ribozyme. Nature *342*(6248), 391-395.

Michel, F., Ellington, A.D., Couture, S., and Szostak, J.W. 1990. Phylogenetic and genetic evidence for base-triples in the catalytic domain of group I introns. Nature *347*(6293), 578-580.

Michel, F. and Westhof, E. 1990. Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. J. Mol. Biol. *216*(3), 585-610.

Michel, F., Jaeger, L., Westhof, E., Kuras, R., Tihy, F., Xu, M.Q., and Shub, D.A. 1992. Activation of the catalytic core of a group I intron by a remote 3' splice junction. Genes Dev. *6*(8), 1373-1385.

Mohr, G. and Lambowitz, A.M. 1991. Integration of a group I intron into a ribosomal RNA sequence promoted by a tyrosyl-tRNA synthetase. Nature *354*(6349), 164-167.

Mohr, G., Zhang, A., Gianelos, J.A., Belfort, M., and Lambowitz, A.M. 1992. The neurospora CYT-18 protein suppresses defects in the phage T4 *td* intron by stabilizing the catalytically active structure of the intron core. Cell *69*(3), 483-494.

Mohr, G., Caprara, M.G., Guo, Q., and Lambowitz, A.M. 1994. A tyrosyl-tRNA synthetase can function similarly to an RNA structure in the Tetrahymena ribozyme. Nature *370*(6485), 147-150.

Mohr, G., Rennard, R., Cherniack, A.D., Stryker, J., and Lambowitz, A.M. 2001. Function of the Neurospora crassa mitochondrial tyrosyl-tRNA synthetase in RNA splicing. Role of the idiosyncratic N-terminal extension and different modes of interaction with different group I introns. J. Mol. Biol. *307*(1), 75-92.

Murphy, F.L. and Cech, T.R. 1993. An independently folding domain of RNA tertiary structure within the Tetrahymena ribozyme. Biochemistry *32*(20), 5291-5300.

Myers, C.A., Wallweber, G.J., Rennard, R., Kemel, Y., Caprara, M.G., Mohr, G., and Lambowitz, A.M. 1996. A tyrosyl-tRNA synthetase suppresses structural defects in the two major helical domains of the group I intron catalytic core. J. Mol. Biol. *262*(2), 87-104.

Nair, S., Ribas de Pouplana, L., Houman, F., Avruch, A., Shen, X., and Schimmel, P. 1997. Species-specific tRNA recognition in relation to tRNA synthetase contact residues. J. Mol. Biol. *269*(1), 1-9.

Pan, T. 1995. Higher order folding and domain analysis of the ribozyme from Bacillus subtilis ribonuclease P. Biochemistry *34*(3), 902-909.

Ralston, C.Y., Sclavi, B., Sullivan, M., Deras, M.L., Woodson, S.A., Chance, M.R., and Brenowitz, M. 2000. Time-resolved synchrotron X-ray footprinting and its application to RNA folding. Methods Enzymol. *317*, 353-368.

Rho, S.B. and Martinis, S.A. 2000. The bI4 group I intron binds directly to both its protein splicing partners, a tRNA synthetase and maturase, to facilitate RNA splicing activity. RNA *6*(12), 1882-1894.

Ribas de Pouplana, L. and Schimmel, P. 2001. Operational RNA code for amino acids in relation to genetic code in evolution. J. Biol. Chem. *276*(10), 6881-6884.

Rich, A. and RajBhandary, U.L. 1976. Transfer RNA: molecular structure, sequence, and properties. Annu. Rev. Biochem. *45*, 805-860.

Roman, J. and Woodson, S.A. 1998. Integration of the Tetrahymena group I intron into bacterial rRNA by reverse splicing in vivo. Proc. Natl. Acad. Sci. USA *95*(5), 2134-2139.

Rould, M.A., Perona, J.J., Soll, D., and Steitz, T.A. 1989. Structure of E. coli glutaminyl-tRNA synthetase complexed with tRNA(Gln) and ATP at 2.8 A resolution. Science *246*(4934), 1135-1142.

Saldanha, R., Ellington, A., and Lambowitz, A.M. 1996. Analysis of the CYT-18 protein binding site at the junction of stacked helices in a group I intron RNA

by quantitative binding assays and in vitro selection. J. Mol. Biol. *261*(1), 23-42.

Saldanha, R.J., Patel, S.S., Surendran, R., Lee, J.C., and Lambowitz, A.M. 1995. Involvement of Neurospora mitochondrial tyrosyl-tRNA synthetase in RNA splicing. A new method for purifying the protein and characterization of physical and enzymatic properties pertinent to splicing. Biochemistry *34*(4), 1275-1287.

Sambrook, J., Fritssch, E.F., and Maniatis, T., *Molecular Cloning. A Labratory Manual*. 2nd ed. 1989, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sanger, F., Nicklen, S., and Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*(12), 5463-5467.

Schimmel, P. and Ribas De Pouplana, L. 2000. Footprints of aminoacyl-tRNA synthetases are everywhere. Trends Biochem. Sci. *25*(5), 207-209.

Schroeder, R., von Ahsen, U., and Belfort, M. 1991. Effects of mutations of the bulged nucleotide in the conserved P7 pairing element of the phage T4 td intron on ribozyme function. Biochemistry *30*(13), 3295-3303.

Sclavi, B., Sullivan, M., Chance, M.R., Brenowitz, M., and Woodson, S.A. 1998. RNA folding at millisecond intervals by synchrotron hydroxyl radical footprinting. Science *279*(5358), 1940-1943.

Streicher, B., von Ahsen, U., and Schroeder, R. 1993. Lead cleavage sites in the core structure of group I intron-RNA. Nucleic Acids Res. *21*(2), 311-317.

Thirumalai, D., Lee, N., Woodson, S.A., and Klimov, D. 2001. Early events in RNA folding. Annu. Rev. Phys. Chem. *52*, 751-762.

Wakasugi, K., Quinn, C.L., Tao, N., and Schimmel, P. 1998. Genetic code in evolution: switching species-specific aminoacylation with a peptide transplant. EMBO J. *17*(1), 297-305.

Wallweber, G.J., Mohr, S., Rennard, R., Caprara, M.G., and Lambowitz, A.M. 1997. Characterization of Neurospora mitochondrial group I introns reveals different CYT-18 dependent and independent splicing strategies and an alternative 3' splice site for an intron ORF. RNA *3*(2), 114-131.

Weeks, K.M. and Cech, T.R. 1995. Protein facilitation of group I intron splicing by assembly of the catalytic core and the 5' splice site domain. Cell *82*(2), 221-230.

Weeks, K.M. and Cech, T.R. 1995. Efficient protein-facilitated splicing of the yeast mitochondrial bI5 intron. Biochemistry *34*(23), 7728-7738.

Weeks, K.M. and Cech, T.R. 1996. Assembly of a ribonucleoprotein catalyst by tertiary structure capture. Science *271*(5247), 345-348.

Wernette, C.M., Saldahna, R., Perlman, P.S., and Butow, R.A. 1990. Purification of a site-specific endonuclease, I-Sce II, encoded by intron 4 alpha of the mitochondrial coxI gene of Saccharomyces cerevisiae. J. Biol. Chem. *265*(31), 18976-18982.

Williamson, C.L., Tierney, W.M., Kerker, B.J., and Burke, J.M. 1987. Site-directed mutagenesis of core sequence elements 9R', 9L, 9R, and 2 in self-splicing Tetrahymena pre-rRNA. J. Biol. Chem. *262*(30), 14672-14682.

Yarus, M., Illangesekare, M., and Christian, E. 1991. An axial binding site in the Tetrahymena precursor RNA. J. Mol. Biol. *222*(4), 995-1012.

Zarrinkar, P.P. and Williamson, J.R. 1994. Kinetic intermediates in RNA folding. Science *265*(5174), 918-924.

Zarrinkar, P.P. and Williamson, J.R. 1996. The kinetic folding pathway of the Tetrahymena ribozyme reveals possible similarities between RNA and protein folding. Nat. Struct. Biol. *3*(5), 432-438.

Zhang, F., Ramsay, E.S., and Woodson, S.A. 1995. In vivo facilitation of Tetrahymena group I intron splicing in Escherichia coli pre-ribosomal RNA. RNA *1*(3), 284-292.

Vita

Christopher Allan Myers was born in Gallipolis, Ohio, on March 23, 1968. After finishing high school at Gallia Academy High School, he attended the University of Rio Grande in Rio Grande, Ohio, where he majored in Math. After two years, he transferred to Ohio University in Athens, Ohio, where he received his Degree of Bachelor of Science in Forensic Chemistry, and Zoology in 1991. In 1991, he was awarded the Outstanding Graduate in Chemistry by the Chemistry Department and was honored by the American Chemical Society for academic success. He attended graduate school in the Department of Biochemistry, at the Ohio State University in Columbus, Ohio in 1991. In 1998, he was awarded his Degree of Master of Science in Biochemistry from Ohio State, before moving to the University of Texas at Austin, in Austin, Texas. During his graduate study, he published the following papers:

Myers, C. A., Wallweber, G. J., Rennard, R., Kemel, Y., Caprara, M. G., Mohr, G. and Lambowitz, A. M. 1996. A tyrosyl-tRNA synthetase suppresses structural defects in the two major helical domains of the group I intron catalytic core. J. Mol. Biol. *262*(2), 87-104.

Caprara, M. G., Myers, C. A. and Lambowitz, A. M. 2001. Interaction of the Neurospora crassa mitochondrial tyrosyl-tRNA synthetase (CYT-18 protein) with the group I intron P4-P6 domain. Thermodynamic analysis and the role of metal ions. J. Mol. Biol. *308*(2), 165-90.

Myers, C. A., Kuhla, B. and Lambowitz, A. M. 2002. tRNA-like recognition of group I introns by a tyrosyl-tRNA synthetase. Proc. Natl. Acad. Sci. U S A *99*(5), 2630-2635.

Permanent address: 610B Genard St., Austin, TX, 78751

The author typed this dissertation.