

Elucidation of the Kinetic Folding Pathway of a Group I Intron

by

Patrick P. Zarrinkar

B.S., Chemistry
University of South Carolina, 1990

Submitted to the Department of Chemistry in Partial Fulfillment of the Requirements for the
Degree of

DOCTOR OF PHILOSOPHY
IN BIOCHEMISTRY

at the

Massachusetts Institute of Technology

June, 1996

© Massachusetts Institute of Technology, 1996
All Rights Reserved

Signature of Author.....



Department of Chemistry
April 25, 1996

Certified by.....



James R. Williamson
Associate Professor of Chemistry
Thesis Supervisor

Accepted by.....

Dietmar Seyferth
Chair, Departmental Committee on Graduate Students

Science

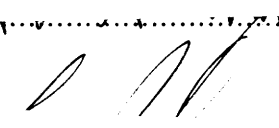
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

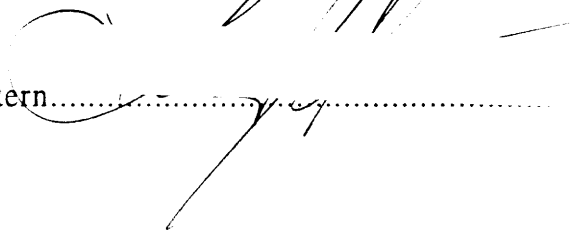
JUN 12 1996

This doctoral thesis has been examined by a committee of the Department of Chemistry as follows:

Professor John M. Essigmann.....

Chair

Professor James R. Williamson.....

Thesis Supervisor

Professor Lawrence J. Stern.....


Elucidation of the Kinetic Folding Pathway of a Group I Intron

by

Patrick P. Zarrinkar

Submitted to the Department of Chemistry in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biochemistry

ABSTRACT

The process by which large RNA molecules fold into complex higher order structures is a basic question in nucleic acid biochemistry. The *Tetrahymena* ribozyme has been used as a model system to study RNA folding, and a minimal kinetic folding pathway elucidated for this group I intron. A kinetic oligonucleotide hybridization assay was developed which exploits the differential accessibility of the RNA to binding by short, complementary oligodeoxynucleotides in the presence and absence of Mg^{2+} , and allows following the timecourse of folding in a sequence specific way. Folding of the ribozyme was shown to be a complex process involving multiple intermediates and both Mg^{2+} -dependent and Mg^{2+} -independent steps. A model for the kinetic folding pathway is proposed where the two main structural subdomains, called P4-P6 and P3-P7, form in a hierarchical manner, with P4-P6 forming first. The rate limiting step is a Mg^{2+} -independent rearrangement leading to the formation of the P3-P7 subdomain, and takes place on the minute timescale. Probing the molecular details of the observed kinetic steps by site directed mutagenesis revealed that the observed slow step involves the formation of specific tertiary interactions, including a triple helical scaffold connecting the two subdomains. The integrity of the fast folding P4-P6 subdomain is required for formation of P3-P7, confirming the hierarchical nature of the folding pathway. The P3-P7 subdomain forms in an interdependent manner, indicating that this structural subunit also represents a kinetic folding unit. The proposed folding mechanism reveals parallels between RNA and protein folding. In both classes of molecules, short range secondary structure forms fast and early, structural subdomains corresponding to kinetic folding units fold in a hierarchical manner, and tertiary interactions between subdomains form late during folding. An initial investigation of the folding properties of RNase P RNA revealed that the slow folding rate and the Mg^{2+} -independence of the slow step are conserved, and suggests that the folding pathway of the group I intron may be relevant to RNA folding in general. These studies provide a conceptual framework for RNA folding, and constitute a foundation for more detailed investigations.

Thesis Supervisor: James R. Williamson

Title: Associate Professor of Chemistry

ACKNOWLEDGMENTS

There are of course many people to thank without whom I either would not have made it this far at all, or without whom getting here would have been a lot more painful.

The greatest thanks must go to my advisor, Jamie Williamson. First of all, for not throwing me out of his office when the first student who wanted to work for him asked whether he had to do NMR if he joined the (then nonexistent) group. Fortunately for me, Jamie swallowed real hard and said “No, you do not have to do NMR if you join my lab”. So I started playing with these cool microorganisms called *Oxytricha* that could be seen devouring live algae under the microscope; but that’s another story, and as you can see from the title of this thesis, my adventures as a microbiologist did not last very long. Fortunately, there still was this RNA folding thing that seemed pretty interesting, too.... Much more importantly, though, I want to thank Jamie for being a great teacher and mentor who leads by example and whose philosophy of making ‘no small plans, for they have not the power to stir people’s hearts’ creates an exciting environment in which to do science. Rather than micromanaging experiments, Jamie has challenged me to find solutions to the problems and questions that inevitably arose along the way, and it has been in the process of trying to rise to these challenges that I feel I have learned the most. At the same time, he has always made time to give detailed advice when it was needed, and has always provided encouragement and perspective when things looked grim. Jamie’s concern, however, goes further than simply getting good results, and he has also always emphasized, and taught me about, the importance of being able to put those results in context and communicate them effectively. Thanks for everything, Jamie!

Next to Jamie, the person most responsible for my making it this far, and actually getting something accomplished, is Jody Puglisi. Jody provided a different outlook, and reminded me of the bigger picture, when I was ready to quit science. His endless patience in listening and giving advice probably saved me many months of failed experiments. Much of what I know about RNA is due to his readiness to engage in many long discussions and share his exhaustive knowledge of the subject. Just as importantly, though, Jody’s presence in the lab made it a much more fun and human place than it otherwise might have been. And who could ever forget that infamous trip to the Hong Kong...!

Many thanks also to all the current and former members of the Williamson lab for countless discussions, for advice on techniques and strategies, but most of all for contributing towards an atmosphere in the lab that, in spite of all the whining and carping sometimes displayed (of which I am as guilty as anyone), is very positive and fosters much cooperation. Thanks to Martha Rook and Dan Treiber, who are continuing the work on RNA folding and who have contributed much to help interpret many of the results described in the latter part of this thesis. To Robert Batey, Christopher Cilley, Alex Brodsky, John Battiste, Heidi Erlacher, Tom Tolbert, Hongyuan Mao, Jeff Orr, Feng Tao, Sharon Parker and Maki Inada. Thanks also to two undergraduate students I have had the opportunity to work with: Quinn Eastman, who helped get the folding project off the ground, and Jing Wang, who performed many of the RNase P studies. If Jing had not joined the lab there would be no chapter on RNase P.

Spending six years in a lab at MIT in freezing New England would have been far more painful without friends who always believed, and always provided some connection with the ‘real’ world. Major thanks to André LaCroix. Our annual ‘pilgrimages’ to the beaches and lowcountry of South Carolina have allowed me to keep my sanity up here.

Thanks to Ray Liles, Ellen Riddle and Joanna Kufel, who help remind me that there is a world outside of MIT and science, and to Magda Pop, who made 1 1/2 of the six years here much more bearable, and gave me much to think about.

Finally, I cannot forget my parents, without whom I would not even be in the U.S. today.

Table of Contents

ABSTRACT		3
ACKNOWLEDGMENTS		4
1. THE RNA FOLDING PROBLEM		8
The functional versatility of RNA		8
RNA folding: product and process		14
The <i>Tetrahymena</i> ribozyme as a model system		22
2. A KINETIC OLIGONUCLEOTIDE HYBRIDIZATION ASSAY TO STUDY RNA FOLDING		38
Concept		38
Feasibility		44
Controls		48
3. KINETIC INTERMEDIATES ON THE FOLDING PATHWAY OF THE <i>TETRAHYMENA</i> RIBOZYME		54
Identification of intermediates		54
A model for the kinetic folding mechanism		67
4. PROBING THE DETAILS OF THE FOLDING MECHANISM BY SITE-DIRECTED MUTAGENESIS		73
Interdependence of P3 and P7		73
Importance of the triple helical scaffold		80
Catalytic activity of mutant ribozymes		85
Role of the P9.1-P9.2 peripheral extension		86
Hierarchical relationship between the two subdomains		90
A refined model for the folding mechanism		94
5. FOLDING OF RNASE P RNA		100
RNase P as a second model system		100
Adapting the kinetic oligonucleotide hybridization assay		103
A complex folding pathway		106
Central features of RNA folding		117

6. PARALLELS TO PROTEIN FOLDING AND FUTURE DIRECTIONS	119
APPENDIX: MATERIALS AND METHODS	126
REFERENCES	135
BIOGRAPHICAL NOTE	152

1. THE RNA FOLDING PROBLEM

The functional versatility of RNA

The discovery that RNA can actively catalyze reactions in the absence of proteins (Cech *et al.*, 1981; Kruger *et al.*, 1982; Guerrier-Takada *et al.*, 1983; Guerrier-Takada and Altman, 1984) initiated a major shift in the understanding of RNA function. Where previously RNA was mainly viewed as a passive participant in processes related to information transfer from DNA to proteins, it is by now recognized to play a central and vital role in almost every step of gene expression and genome replication (Gesteland and Atkins, 1993) (Fig. 1). Although today RNA is neither the medium for storage of cellular genetic information nor its main final product, it does mediate all functions related to managing that information.

Gene Expression

To produce functional proteins, coding information must be extracted from the raw data stored in the genome and then translated into the amino acid code of proteins. This is accomplished by transcription of DNA genes into RNA and processing of the resulting primary transcripts to produce mature translatable messages. In recent years, evidence has been accumulating that two key steps involved in this process, splicing of non-coding introns from primary transcripts and translation, are at heart RNA catalyzed.

Indeed, it was the investigation of splicing that led to the discovery of RNA catalysis in the form of self-splicing group I introns. For both group I and group II intron splicing, it is the intron itself that serves as the catalyst (Cech *et al.*, 1981; Kruger *et al.*, 1982; Peebles *et al.*, 1986; van der Veen *et al.*, 1986). In contrast, spliceosomal splicing,

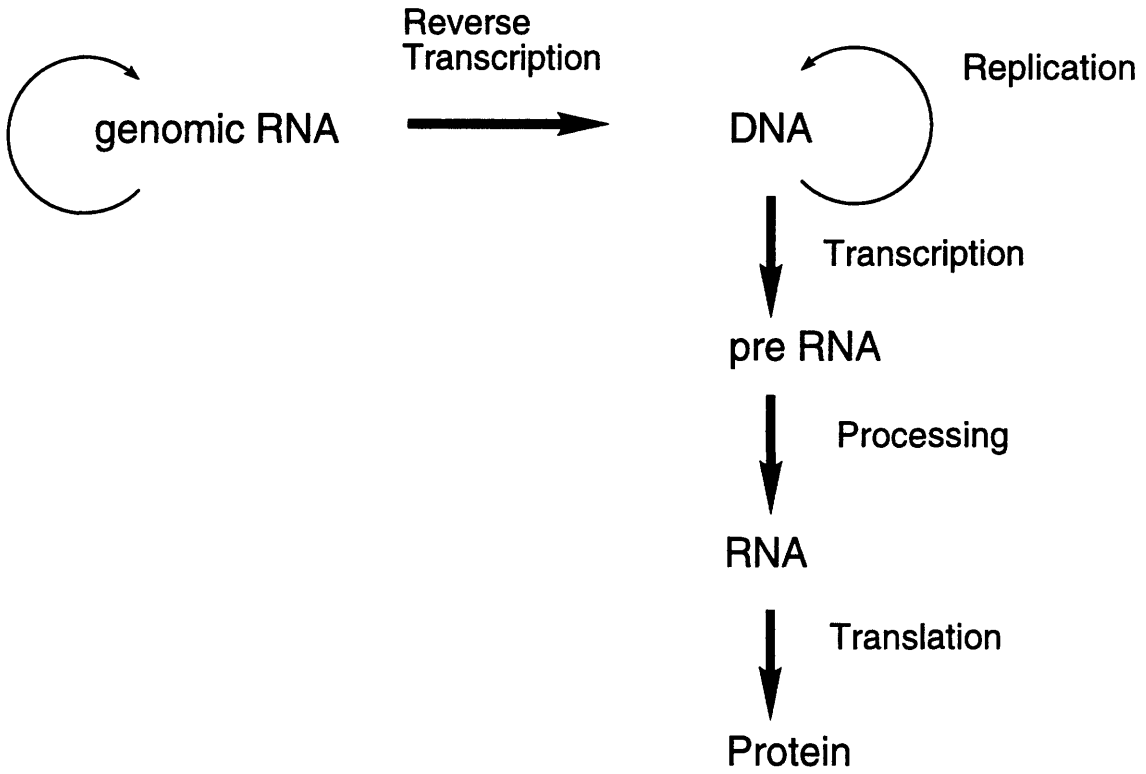


Figure 1. The central role of RNA in gene expression and genome replication.

which accounts for most splicing activity in contemporary organisms, is accomplished through the action of an external catalyst, the spliceosome, on the intron substrate (Moore *et al.*, 1993; Madhani and Guthrie, 1994). The spliceosome consists of five small nuclear RNAs (U1, U2, U4, U5 and U6 snRNAs) and an extensive and as yet incompletely defined set of proteins. It was the similar mechanisms of group II and spliceosomal splicing that first suggested that the RNA may be the catalytic component of the spliceosome (Cech, 1986; Peebles *et al.*, 1986; Schmelzer and Schweyen, 1986; van der Veen *et al.*, 1986; Sontheimer and Steitz, 1993; Padgett *et al.*, 1994). In addition to mechanistic parallels, it is now known that there must also be structural similarities between group II introns and spliceosomal RNAs, further supporting the hypothesis that spliceosomal splicing is RNA catalyzed (Madhani and Guthrie, 1994; Yu *et al.*, 1995).

The ribosome, like the spliceosome, is a large catalytic complex composed of RNA and proteins (Moore, 1993). It had long been assumed that the RNA merely forms a scaffold to organize the ribosomal proteins, and that peptidyl transfer, which is at the heart of the translational process, is a protein catalyzed reaction (Noller, 1991; Moore, 1993). Only once the concept of RNA catalysis was well established became the possibility apparent that it might in fact be the RNA that is the active component, and proteins the organizing scaffold (Noller, 1991; Moore, 1993; Noller, 1993). Experiments to test this hypothesis directly provided strong evidence that peptidyl transfer is RNA catalyzed (Noller *et al.*, 1992).

In addition to peptidyl transfer, RNA plays a key role in a number of other aspects of translation. The function of tRNA as an adapter molecule to connect the four base code of nucleic acids to the twenty amino acid code of proteins has long been recognized, and tRNA structure and function have been intensely investigated for over twenty years (Schimmel and Redfield, 1980; Giegé *et al.*, 1993; Söll, 1993). To produce mature tRNA

from primary pre-tRNA transcripts, a 5' leader sequence must be removed. This is accomplished by the ribonucleoprotein enzyme RNase P, another catalyst composed of RNA and protein (Cech, 1993; Altman *et al.*, 1995). In this case, it is well established that the reaction is RNA catalyzed, since for many prokaryotes the RNA alone, in the absence of protein, is active (Darr *et al.*, 1992; Cech, 1993; Altman *et al.*, 1995).

The primary RNA transcript itself, in addition to serving as substrate for processing and translation, can directly participate in the regulation of these processes. There is evidence for an interaction between the 5' and 3' terminal regions of mRNAs (Caponigro and Parker, 1995; Decker and Parker, 1995; Kuge and Richter, 1995), and this interaction is likely to play a role in regulating translation, often mediated by proteins bound to sequences in the 3' or 5' untranslated region of mRNAs (Decker and Parker, 1995; McCarthy and Kollmus, 1995), and mRNA degradation (Decker and Parker, 1995). Pre-mRNA structure has been implicated in the regulation of spliceosomal splicing (Goguel and Rosbash, 1993; Parker and Siliciano, 1993; Libri *et al.*, 1995), and translational events such as frameshifting (Brierley *et al.*, 1989) and selenocysteine incorporation (Ringquist *et al.*, 1993), as well, require structured regions in mRNAs.

While transcription is definitely protein catalyzed, and largely regulated by a complex network of protein:protein interactions, some regulatory events do involve the participation of RNA. In HIV, for example, interaction of the TAR RNA element and the Tat protein, mediated by the structure of the RNA, allows processive transcription of the viral RNA (Frankel, 1992). A similar mechanism of antitermination functions in the regulation of transcription during the infectious cycle of bacteriophage lambda (Greenblatt *et al.*, 1993).

Genome replication

In addition to the numerous roles played by RNA during gene expression, it is also actively involved in the replication of the genomes of eukaryotes and many viruses. To accurately replicate the ends of linear chromosomes in eukaryotes, the cell employs the ribonucleoprotein enzyme telomerase (Zakian, 1995). Telomerase is a specialized reverse transcriptase whose RNA component contains a sequence serving as template for the addition of telomeric sequences to the ends of chromosomes to prevent their progressive shortening through successive rounds of replication. Moreover, recent results provide evidence that the RNA can also directly modulate the activity of the enzyme (Gilley *et al.*, 1995).

Many plant viruses, viroids and virusoids and at least one human virus replicate their RNA genomes by a rolling circle mechanism. This involves the production of multiple, covalently linked copies of the genome which must be cleaved into single length units. Cleavage is accomplished by small, catalytic RNA motifs that are a part of the sequence. These motifs, each of which has been studied extensively *in vitro*, include the hammerhead, hairpin and hepatitis delta virus ribozymes (Symons, 1992; Lazinski and Taylor, 1995). Finally, retroviruses rely on structured tRNAs to prime the reverse transcription of genomic RNA into DNA, required for incorporation of viral sequences into the host genome (Litvak *et al.*, 1994).

Additional functions

The development of *in vitro* selection technology has allowed the isolation of RNAs with a wide range of ligand binding and catalytic capabilities (Gold *et al.*, 1993; Szostak and Ellington, 1993; Joyce, 1994). Most striking is the realization that RNAs with functions unlike those of natural RNAs can be produced. RNA sequences can be selected which bind many types of proteins, both those which do normally bind nucleic acids and those which do not, with very high affinity and specificity (Gold *et al.*, 1993; Eaton *et al.*,

1995; Gold *et al.*, 1995). RNAs able to specifically bind nucleotides, several amino acids, including the aliphatic valine, and numerous other small molecules and cofactors have been reported (Gold *et al.*, 1995). Derivatives of the *Tetrahymena* ribozyme (see below) that can catalyze phosphodiester bond cleavage in the presence of Ca^{2+} instead of Mg^{2+} (Lehman and Joyce, 1993), or cleave DNA and RNA substrates with similar efficiencies (Beaudry and Joyce, 1992), have been isolated. The DNA cleaving variants have also been demonstrated to cleave amide bonds effectively (Dai *et al.*, 1995), while the wild type ribozyme can catalyze aminoacyl ester bond cleavage at a low level (Piccirilli *et al.*, 1992). Finally, RNAs that have aminoacyl RNA synthetase (Illangasekare *et al.*, 1995), RNA ligase (Bartel and Szostak, 1993), kinase (Lorsch and Szostak, 1994) and alkylase (Wilson and Szostak, 1995) activity have been obtained from completely random sequences.

While the relevance of these activities to contemporary cellular functions may not be immediately obvious, there is at least one instance where the selection of an RNA with an ‘unnatural’ function foreshadowed the discovery of a cellular RNA that carried out an analogous reaction. Very recently it was demonstrated that incorporation of a specific group II intron into the DNA genome, a reaction that is responsible for the observed mobility of some group II introns, is in part RNA catalyzed (Zimmerly *et al.*, 1995). The intron cleaves one strand of the double stranded DNA, while a protein associated with the intron cleaves the second strand. DNA cleavage activity of RNA is therefore not a laboratory curiosity, as might have been initially thought, but forms part of the natural range of RNA functions. It may not be unreasonable to expect that other RNA functions identified by *in vitro* selection techniques may eventually be recognized as occurring in contemporary organisms.

In summary, the past fifteen years have shown that RNA, far from being a passive bystander in the cell, is in fact a driving force at the heart of almost every process connected

with information management. It has long been speculated that early in evolution life was based on RNA, and that in this 'RNA world' RNA acted both as catalyst and carrier of genetic information (Gesteland and Atkins, 1993). The demonstration that RNA can possess catalytic activity provided a strong boost in this theory. The most convincing evidence for this crucial role of RNA in evolution, however, is probably the very involvement even today of RNA in every step of processing genetic information. While the *in vivo* functions of RNA outlined above are almost always accomplished in conjunction with proteins, the RNA is in many cases the 'active ingredient' and in all cases required for proper function. It is unlikely that the full repertoire of RNA functions has yet been uncovered. Continued investigation of cellular processes will almost certainly reveal important new roles of RNA, and *in vitro* selection technology will continue to generate RNAs with interesting, and potentially useful, abilities.

RNA folding: product and process

RNA can fulfill the wide variety of functions described above because it has the ability to fold up into complex three dimensional structures that form active sites for catalysis and recognition sites for ligands. A thorough understanding of RNA function can only be gained by answering two questions. First, what are the active structures of RNA, and how do important ligands bind to these structures? Answers to this question will provide insight into the mechanisms of action of RNA and reveal how it interacts with other cellular components. Second, what is the process by which these structures are formed? The cell is a complex environment and proper formation of the active structure has to compete with formation of inactive structures, or misfolding, degradation of the RNA and formation of nonspecific complexes with cellular proteins. How does RNA circumvent these competing processes and what drives formation of the final structure? Together,

these two questions constitute an RNA folding problem that is analogous to the much studied protein folding problem (Herschlag, 1995; Pyle and Green, 1995).

Structure

Although structural information for RNA is still relatively scarce, significant progress has been made in recent years in the application of X-ray crystallography, nuclear magnetic resonance and biochemical techniques, and the library of known RNA structures is increasing slowly but steadily (Valegård *et al.*, 1994; Marino *et al.*, 1995; Nagai *et al.*, 1995; Pyle and Green, 1995; Shen *et al.*, 1995). Most of the high resolution structural information currently available has been obtained by studying functionally important fragments of larger RNAs either in their free form, or bound to domains or subdomains of protein ligands. Exceptions to this are the crystal structures of tRNAs and their complexes with aminoacyl-tRNA synthetases (Rould *et al.*, 1989; Rould *et al.*, 1991; Cavarelli *et al.*, 1993; Giegé *et al.*, 1993; Biou *et al.*, 1994) and elongation factor Tu (Nissen *et al.*, 1995), and the hammerhead ribozyme (Pley *et al.*, 1994; Scott *et al.*, 1995). It has become apparent that RNA plays a very active role in binding proteins and other ligands and, far from being passively recognized, generally undergoes significant conformational changes upon binding. Similarly, the structure of the hammerhead ribozyme revealed that the RNA must undergo a conformational change in going from the ground state to the transition state for catalysis.

Less is known about the structures of larger RNAs, such as group I and group II introns, RNase P, and ribosomal RNAs. While no high resolution structural information is available for these molecules to date, information obtained by a combination of phylogenetic comparisons and biochemical experiments has allowed the construction of models for the three dimensional architecture of group I introns (Michel and Westhof, 1990), RNase P RNA (Harris *et al.*, 1994; Westhof and Altman, 1994) and ribosomal

RNA (Stern *et al.*, 1988; Stern *et al.*, 1989; Malhotra and Harvey, 1994). While of relatively low resolution, these models do give a good indication of the relative positions of structural elements. A complementary strategy is the use of NMR and crystallography to obtain high resolution information about key parts of the overall structure (Chastain and Tinoco, 1992; Chastain and Tinoco, 1993; Wimberly *et al.*, 1993; Allain and Varani, 1995; Shen *et al.*, 1995; Szewczak and Moore, 1995; Varani, 1995). Large RNAs appear to be composed of distinct structural subdomains and in some cases, such as group I introns and RNase P, subdomains have been shown to be independently stable outside of the context of the entire molecule (Murphy and Cech, 1993; Pan, 1995). In at least one case it is also possible to assemble a functional molecule from separately synthesized subdomains (Murphy and Cech, 1993; Doudna and Cech, 1995; Pan, 1995). The details of how the subdomains interact with each other to form the final structure are still unclear.

Understanding these details of the higher order architecture, both in compact globular RNA structures like the large ribozymes and in more complex ribonucleoprotein complexes like the ribosome and the spliceosome, as well as in the interaction of distant parts of a long, presumably not as tightly structured, mRNA, presents one of the greatest challenges in achieving a complete appreciation of the role of RNA.

The folding process

The second question that must be asked to address the RNA folding problem is how a linear, polyanionic chain of nucleotides can fold into a compact, globular structure. What is the pathway by which large RNAs achieve their final, active conformation? The first step in learning how this is accomplished in the complex cellular environment is to decipher the basic rules of the folding process in a simplified, defined *in vitro* environment. To define a folding pathway, three issues must be addressed. First, are there any intermediates, and what is their relevance to productive folding? Second, if there are

intermediates, what is their structure? Third, what are the kinetics of folding? These issues are not easily addressed experimentally, and even less is known about RNA folding pathways than about RNA structure.

Prior to the work described in this thesis, the most extensive studies on RNA folding and dynamics were conducted on tRNA over twenty years ago. Optical spectroscopic techniques were used to follow changes in the structure in response to raising the temperature or altering ionic conditions. In some cases, these optically detected changes were correlated with alterations in NMR spectra to determine which regions of the RNA were involved (Crothers *et al.*, 1974). Consistent with observations of hairpin and duplex formation (Pörschke and Eigen, 1971; Gralla and Crothers, 1973), secondary structure was found to form very rapidly (in milliseconds or faster) and at moderate monovalent ion concentrations. At least some of the tertiary structure in tRNA apparently can form in high concentrations of Na⁺ (> 100-200 mM) in the absence of Mg²⁺ (Cole *et al.*, 1972). Mg²⁺ does, however, greatly stabilize tertiary interactions, and may do so by binding to several specific, high affinity sites (with dissociation constants in the micromolar range) which are only present when the tertiary structure is formed (Stein and Crothers, 1976a; Stein and Crothers, 1976b). Optical changes observed upon the addition of Mg²⁺ to a high salt form of the RNA also indicate that in the absence of Mg²⁺, tRNA is not in a fully native conformation (Cole *et al.*, 1972). The tertiary structure can only form if the secondary structure is already present, and kinetic studies showed that formation of the native structure occurs in several discrete steps (Cole *et al.*, 1972; Yang and Crothers, 1972; Stein and Crothers, 1976a). The rate limiting step when folding is induced by addition of Mg²⁺ at intermediate monovalent ion concentrations (170 mM Na⁺) is formation of tertiary structure (Cole *et al.*, 1972; Stein and Crothers, 1976a), and takes place in milliseconds. If native structure formation was initiated by addition of Mg²⁺ to a low salt

form of tRNA (in 12 mM Na⁺), the rate of folding was observed to be quite slow, with a timeconstant of 80 to 1900 seconds (Cole *et al.*, 1972). This slow step has a large activation energy, which led to the conclusion that the low salt form may include non-native interactions which must be disrupted to form the native structure. Similar results were obtained in a fluorescence study of Mg²⁺-induced tRNA folding at low Na⁺ concentration (10-15 mM) (Lynch and Schimmel, 1974). A sequential folding pathway including several intermediates and both Mg²⁺-dependent and Mg²⁺-independent steps was proposed. Folding occurred on the second timescale at 37°C, and the rate limiting step had a large activation energy (30 kcal/mol) (Lynch and Schimmel, 1974).

The importance of Mg²⁺ for the formation of higher order tertiary structure in RNA was confirmed by more recent studies of Mg²⁺-induced folding of group I introns (Latham and Cech, 1989; Celander and Cech, 1991; Heuer *et al.*, 1991; Weeks and Cech, 1995), RNase P RNA (Pan, 1995) and the hammerhead ribozyme (Bassi *et al.*, 1995). In all cases significant conformational changes were observed that could not be effected by monovalent ions alone. In the large group I intron and RNase P RNAs regions of phylogenetically conserved core segments became protected from Fe(II)-EDTA cleavage in the presence of Mg²⁺, suggesting formation of a globular conformation with a defined 'inside' and 'outside', similar to that found in globular proteins (Latham and Cech, 1989; Celander and Cech, 1991; Heuer *et al.*, 1991; Pan, 1995; Weeks and Cech, 1995). Although these studies were all performed at equilibrium, and did not examine the kinetics of the observed changes, they did produce evidence for folding intermediates at intermediate Mg²⁺ concentrations. In both the *Tetrahymena* group I intron (Celander and Cech, 1991) and a group I intron from yeast (Weeks and Cech, 1996) higher order structure forms in only a defined subdomain at Mg²⁺ concentrations below those required for full native structure formation.

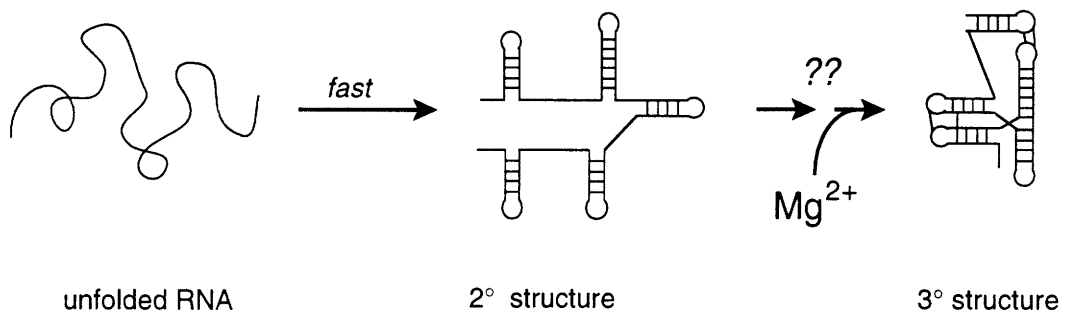


Figure 2. Minimal folding pathway for RNA.

The results from these studies allow the construction of a minimal model for the folding pathway of highly structured RNAs (Fig. 2) (Westhof and Michel, 1992). Folding appears to be a multistep process where simple secondary structure elements, such as hairpin stems, form rapidly and early, and do not require the presence of divalent metals. Formation of higher order tertiary structure is likely to require Mg^{2+} , and to take place on slower timescales. It would not be unexpected if all tertiary structure, particularly in larger RNAs, did not form in a single transition, but in several discrete steps. Folding may also be hierarchical, where formation of secondary structure and early forming tertiary interactions is required for later folding steps. The work described in this thesis addresses the kinetics and the possible involvement of intermediates in the Mg^{2+} -induced formation of higher order structure.

An issue that has received increasing attention recently is the role of proteins during RNA folding. There are two ways in which proteins can influence folding (Herschlag, 1995). One way is by binding to and stabilizing the active structure of the RNA. Most functional RNAs *in vivo* act in the form of ribonucleoprotein complexes, and the RNA alone may not be able to fold into its active conformation under physiological conditions. This has been demonstrated, for example, for ribosomal RNA (Moazed *et al.*, 1986), some group I introns (Weeks and Cech, 1995), and RNase P RNA (Tranguch *et al.*, 1994; Pan, 1995). In these cases, folding is really an assembly process to produce the RNP. RNP formation, as evidenced by ribosome assembly, can be a complex, stepwise hierarchical process (Traub and Nomura, 1969; Held and Nomura, 1973; Sieber and Nierhaus, 1978; Powers *et al.*, 1993). Rearrangements induced by the association of primary binding proteins with rRNA create recognition sites for secondary binding proteins, and, *in vitro*, some of these rearrangements can be very slow. The exact roles of RNA and protein during this process remain to be defined. Studies of simple model systems where the

active structure of a group I intron requires the presence of one specific protein, have given conflicting results. For the group I intron bI5, the fifth intron in cytochrome b pre-mRNA in yeast mitochondria, which requires the CBP-2 protein for full activity, assembly kinetics were found to be governed by folding of the RNA alone, with the rate limiting step being formation of the RNA core (Weeks and Cech, 1996). The protein could only bind to the RNA once the core structure was formed. For a group I intron in mitochondrial large rRNA in *Neurospora*, which requires a tyrosyl-tRNA synthetase (the CYT-18 protein) for activity, the rate limiting step was reported to be protein binding itself, suggesting a more direct involvement of the protein in RNA folding (Saldanha *et al.*, 1995). It may be that different RNPs assemble by fundamentally different mechanisms, or, alternatively, these observations may only reflect differences of degree, with the underlying mechanism being largely conserved. In some instances, the protein of an RNP can, at least in part, be substituted by high Mg^{2+} concentrations (Weeks and Cech, 1995), suggesting that in these cases the protein acts largely in shielding the negatively charged backbone to allow the close apposition of structural elements.

The second way in which proteins can influence RNA folding is by facilitating folding without remaining bound to the RNA in the final structure. Several proteins, including the ribosomal protein S12 (Coetzee *et al.*, 1994) and the *E. coli* protein stpA (Zhang *et al.*, 1995a), have been identified which can promote formation of the active structure of RNAs, but can be removed after folding without loss of RNA activity. These proteins have been shown to possess helicase activity, and it has been speculated that they promote formation of the active structure by preventing misfolding (Coetzee *et al.*, 1994; Herschlag, 1995; Zhang *et al.*, 1995a). Furthermore, the HIV nucleocapsid protein NC17 and hnRNP proteins can stimulate hammerhead ribozyme activity (Tsuchihashi *et al.*, 1993; Bertrand and Rossi, 1994; Herschlag *et al.*, 1994), and are proposed to do so by

promoting the association and dissociation of substrate. This observation is consistent with the RNA annealing activity of hnRNP proteins (Portman and Dreyfuss, 1994). While it seems reasonable that *in vivo* folding even of RNAs which do not require protein cofactors for activity *in vitro* may involve proteins, the *in vivo* relevance of any of these observations remains to be established.

The *Tetrahymena* ribozyme as a model system

The *Tetrahymena* ribozyme is derived from a group I intron found in pre-rRNA of the ciliated protozoan *Tetrahymena thermophila* (Cech and Rio, 1979). It was the first catalytic RNA to be identified (Cech *et al.*, 1981; Kruger *et al.*, 1982), and since its discovery has remained a paradigm system for the study of RNA structure and catalysis (Cech, 1993). The wealth of information available make it the best defined large RNA molecule, and an ideal model system in which to investigate the folding mechanism of large, highly structured RNAs.

Function

As a group I intron, the *Tetrahymena* ribozyme catalyzes its own excision from pre-rRNA in two successive transesterification reactions (Fig. 3A) (Cech *et al.*, 1981; McSwiggen and Cech, 1989). During the first step, the phosphate backbone is cleaved at the 5' splice site by attack of the 3' hydroxyl of a guanosine cofactor, resulting in attachment of the cofactor to the 5' end of the intron, and the liberation of the 3' end of the 5' exon (Zaug and Cech, 1982). Before the second step, the RNA undergoes a conformational change which leads to the last nucleotide of the intron, also a guanosine, being bound in the G-binding site previously occupied by the cofactor (Davies *et al.*, 1982; Michel *et al.*, 1989a; Burke *et al.*, 1990). The 3' hydroxyl of the 5' exon then attacks the phosphate at the 3' splice site, resulting in ligation of the two exons and release of the free

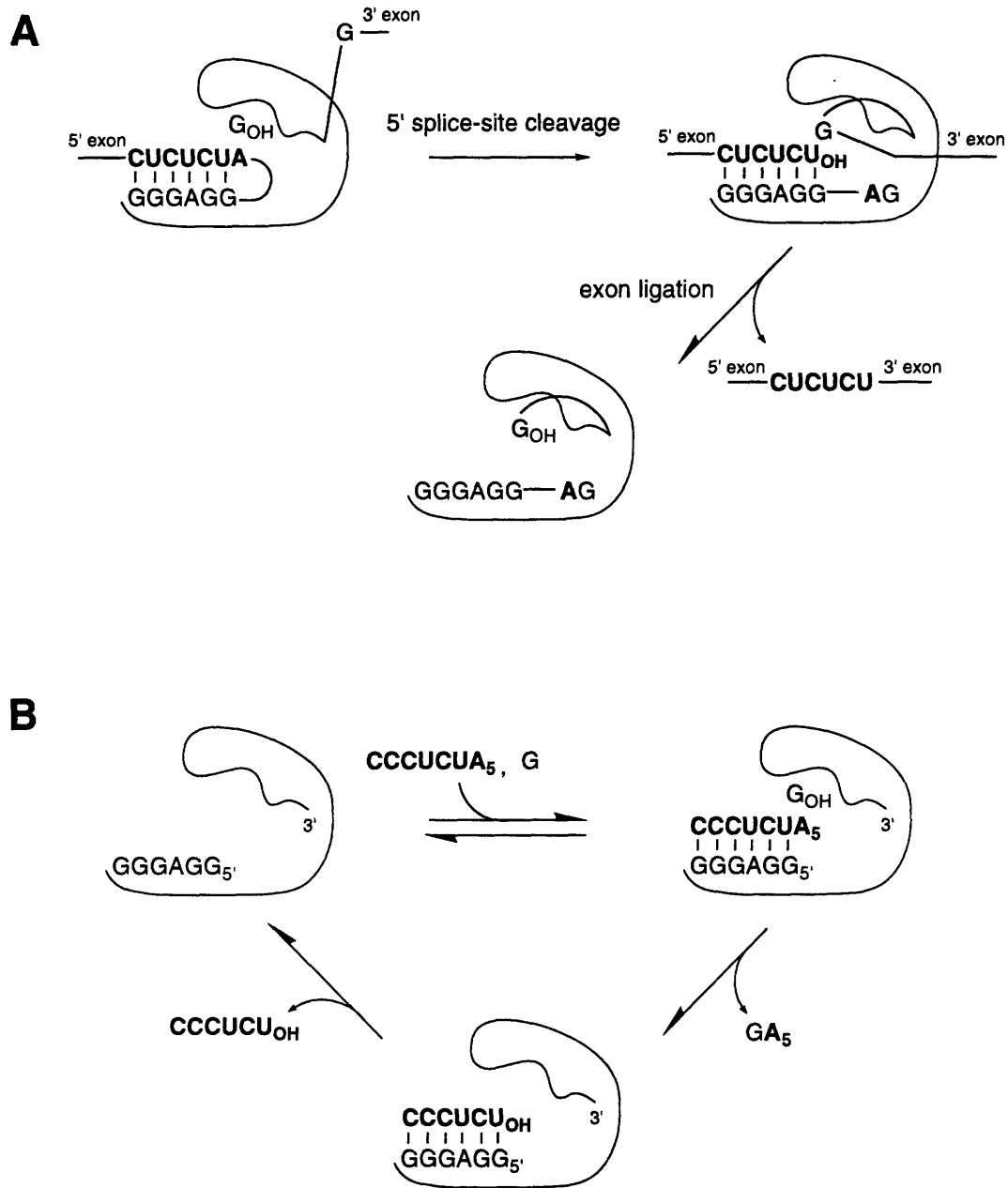


Figure 3. *Tetrahymena* ribozyme function. Adapted from Herschlag and Cech (1990a).
(A) self-splicing.
(B) endonuclease cleavage.

intron. This *cis*-splicing reaction takes place efficiently *in vitro* in the absence of any proteins (Kruger *et al.*, 1982). The only requirement is the presence of moderate concentrations of Mg^{2+} (full activity is observed in 2 mM $MgCl_2$) (Cech *et al.*, 1981; Celander and Cech, 1991). The 5' splice site is defined by a base pairing interaction between nucleotides immediately preceding the splice site and a short recognition sequence, termed the internal guide sequence (IGS) near the 5' end of the intron (Been and Cech, 1986; Waring *et al.*, 1986; Murphy and Cech, 1989). To convert this RNA into a true enzyme able to act on a substrate *in trans* with multiple turnover the intron was shortened at both the 5' and 3' ends (Zaug *et al.*, 1986; Zaug *et al.*, 1988). The deletion of five nucleotides from the 3' end removes the 3' splice site and prevents the G at the 3' splice site from participating in the reaction. Deletion of 21 nucleotides from the 5' end removes unnecessary sequences preceding the internal guide sequence. The resulting ribozyme (also called the L-21 Sca I *Tetrahymena* ribozyme due to its missing 21 nucleotides from the 5' end, and being transcribed from a plasmid template linearized with the restriction enzyme Sca I) can bind an RNA containing a sequence complementary to the IGS and, with the aid of a guanosine cofactor, cleave it at the site corresponding to the 5' splice site in a reaction analogous to the first step of self-splicing (Fig. 3B). The two products are released and the enzyme can undergo another round of catalysis. In the absence of the cofactor, hydroxide ion can act as the nucleophile, and efficient reaction is still observed (Herschlag and Cech, 1990a). The sequence of the IGS can be varied so that virtually any RNA can be targeted for cleavage (Murphy and Cech, 1989). The only real sequence requirement for efficient reaction is the presence of a U immediately preceding the cleavage site which can form a G•U wobble pair with the first nucleotide of the IGS, defining the site of cleavage (Barfod and Cech, 1989; Knitt *et al.*, 1994; Pyle *et al.*, 1994; Strobel and Cech, 1995).

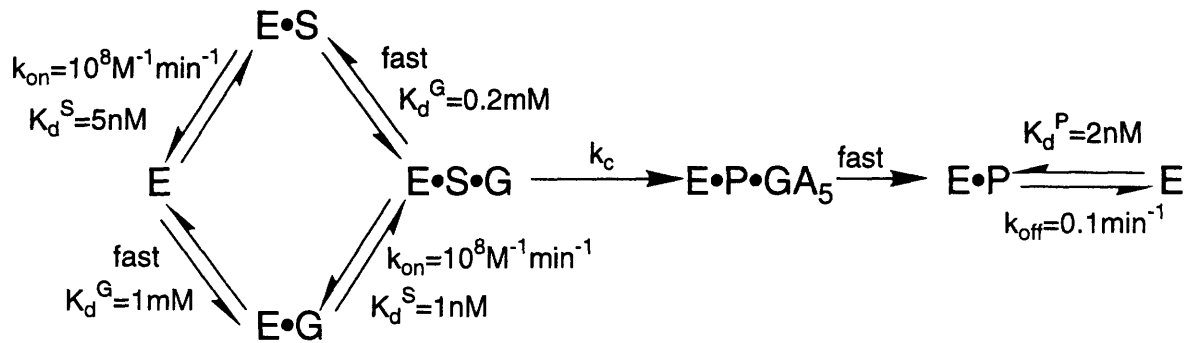


Figure 4. Kinetic mechanism for the *Tetrahymena* ribozyme endonuclease cleavage reaction. Adapted from Herschlag and Khosla (1994). E is the ribozyme, S the substrate (CCCUCUA₅), GA₅ and P (CCCUCU) the reaction products and G the guanosine cofactor.

The application of enzymological techniques has led to the elucidation of a detailed kinetic mechanism for the endonuclease reaction, which has provided a conceptual framework for the design and interpretation of additional studies (Fig. 4) (Herschlag and Cech, 1990a). Under multiple turnover conditions, the rate limiting step in the reaction cycle is release of the product bound to the guide sequence (see Fig. 3B), explaining why substrates that form mismatches with the guide sequence, and therefore bind less tightly, react faster under these conditions (Zaug *et al.*, 1988; Herschlag and Cech, 1990b). The slow rate of product release is consistent with the natural function of the intron where, between the two steps of self-splicing, the 5' exon, corresponding to the slowly released product, must be held in place until exon ligation can be completed (Waring *et al.*, 1986). Under single turnover conditions with subsaturating concentrations of ribozyme, cleavage is at least in part limited by substrate binding (Herschlag and Cech, 1990a). At saturating concentrations of ribozyme the chemical step or a conformational change following substrate binding, depending on pH, become rate limiting (McConnell *et al.*, 1993; Herschlag and Khosla, 1994). At low concentrations of guanosine cofactor, the chemical step also becomes rate limiting (Herschlag and Cech, 1990a). Different steps in the reaction cycle can therefore be assayed by varying the concentrations of ribozyme, substrate and cofactor, and the pH of the reaction buffer.

The ribozyme is most active in Mg^{2+} (Grosshans and Cech, 1989). In the absence of Mg^{2+} , activity is only observed with Mn^{2+} . The requirement for Mg^{2+} can in part be relieved by Ca^{2+} , Sr^{2+} and Ba^{2+} , but not by other divalent or any monovalent metal ions (Grosshans and Cech, 1989; Celander and Cech, 1991). Two types of specifically bound Mg^{2+} ions have been proposed (Grosshans and Cech, 1989). First, a small number (probably one or two) directly participating in catalysis, which can only be substituted by Mn^{2+} (Grosshans and Cech, 1989; Pyle, 1993; Steitz and Steitz, 1993). Second, a slightly

larger number (3-8) that stabilize the active structure of the RNA, but do not participate in catalysis (Grosshans and Cech, 1989; Celander and Cech, 1991; Wang and Cech, 1994). These can be substituted by Ca^{2+} , Sr^{2+} or Ba^{2+} as well as Mn^{2+} , but not by monovalent ions (Grosshans and Cech, 1989; Celander and Cech, 1991). Direct interactions of one oxygen near the scissile phosphodiester bond with Mg^{2+} have been identified (Piccirilli *et al.*, 1993), and a catalytic mechanism involving two Mg^{2+} ions has been proposed (Steitz and Steitz, 1993). While the exact catalytic mechanism is still unclear, substrate destabilization is one of the strategies used by the ribozyme to induce bond cleavage (Narlikar *et al.*, 1995).

By determining the steady state concentrations of spliced and unspliced precursor RNA, it was deduced that *in vivo* splicing proceeds significantly faster than *in vitro* (Brehm and Cech, 1983). This has prompted speculation that additional factors, most likely proteins, promote splicing *in vivo*. Any such proteins cannot be specific to *Tetrahymena*, since an equivalent increase in splicing rate is observed when the intron is introduced into *E. coli* (Zhang *et al.*, 1995b). It remains to be established what limits the rate of splicing *in vivo*: folding of the intron into its active global structure, productive binding of the substrate sequence in the active site, or catalysis.

Structure

A combination of phylogenetic comparisons, chemical and enzymatic mapping and site directed mutagenesis studies was used to define the secondary structure of the intron (Burke *et al.*, 1987). The representation shown in Fig. 5A shows the RNA as a continuous sequence, and highlights the existence of several long range base pairing interactions (such as P3 and P7). Base pairing interactions are numbered from the 5' end as P (for pairing) 1 through P9. A set of core helices is conserved in all group I introns, and these include P3, P4, P6 and P7. In addition to these highly conserved interactions,

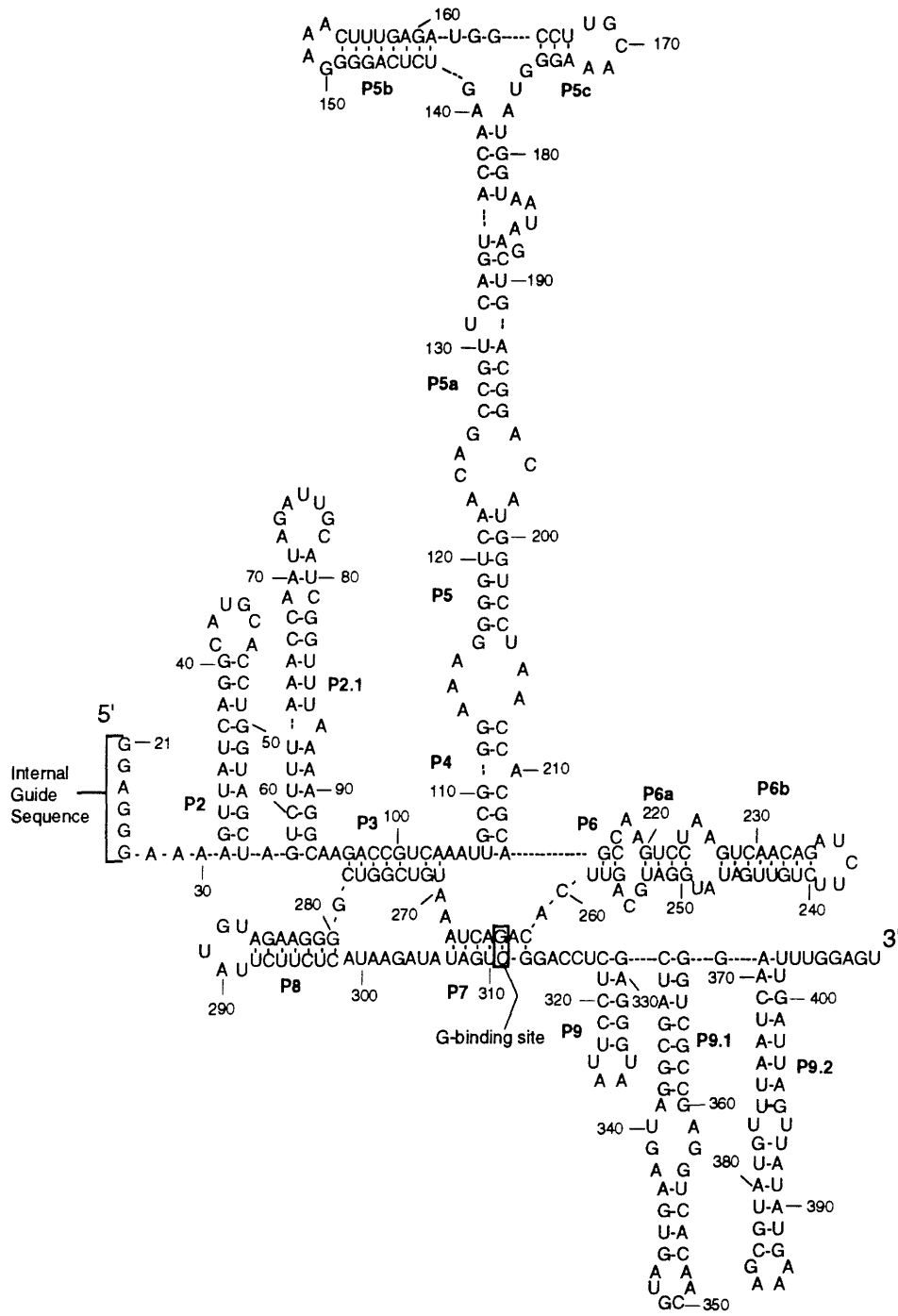


Figure 5. (A) *Tetrahymena* ribozyme secondary structure.
 Representation according to Burke *et al.* (1987).

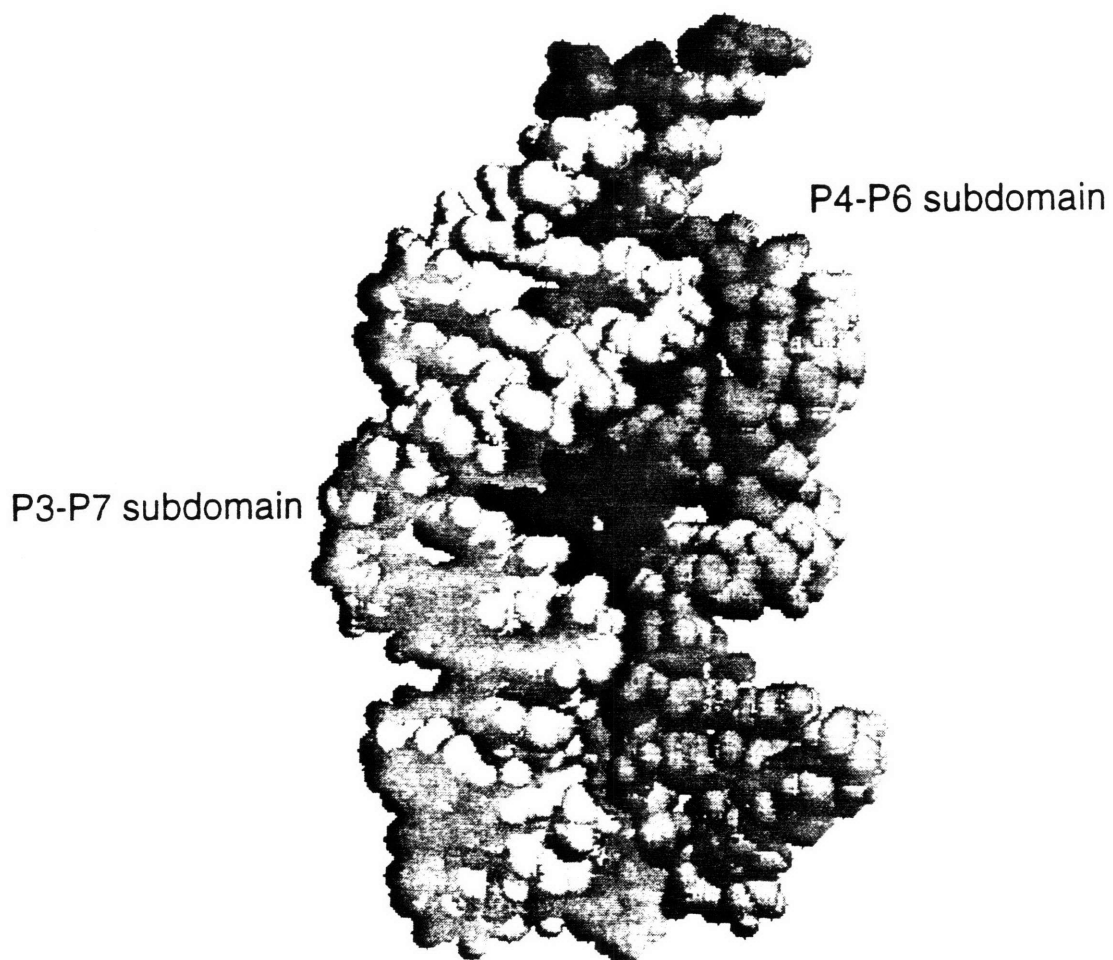


Figure 5. (B) Space-filling model of the three dimensional architecture of the *Tetrahymena* ribozyme core (including P3, P4, P5, P6/P6a, P7, P8 and P9). Adapted from Michel and Westhof (1990). The two subdomains are shown in different shades, and nucleotides proposed to form the triple helical scaffold by interacting with base-pairs in P4-P6 are highlighted in black.

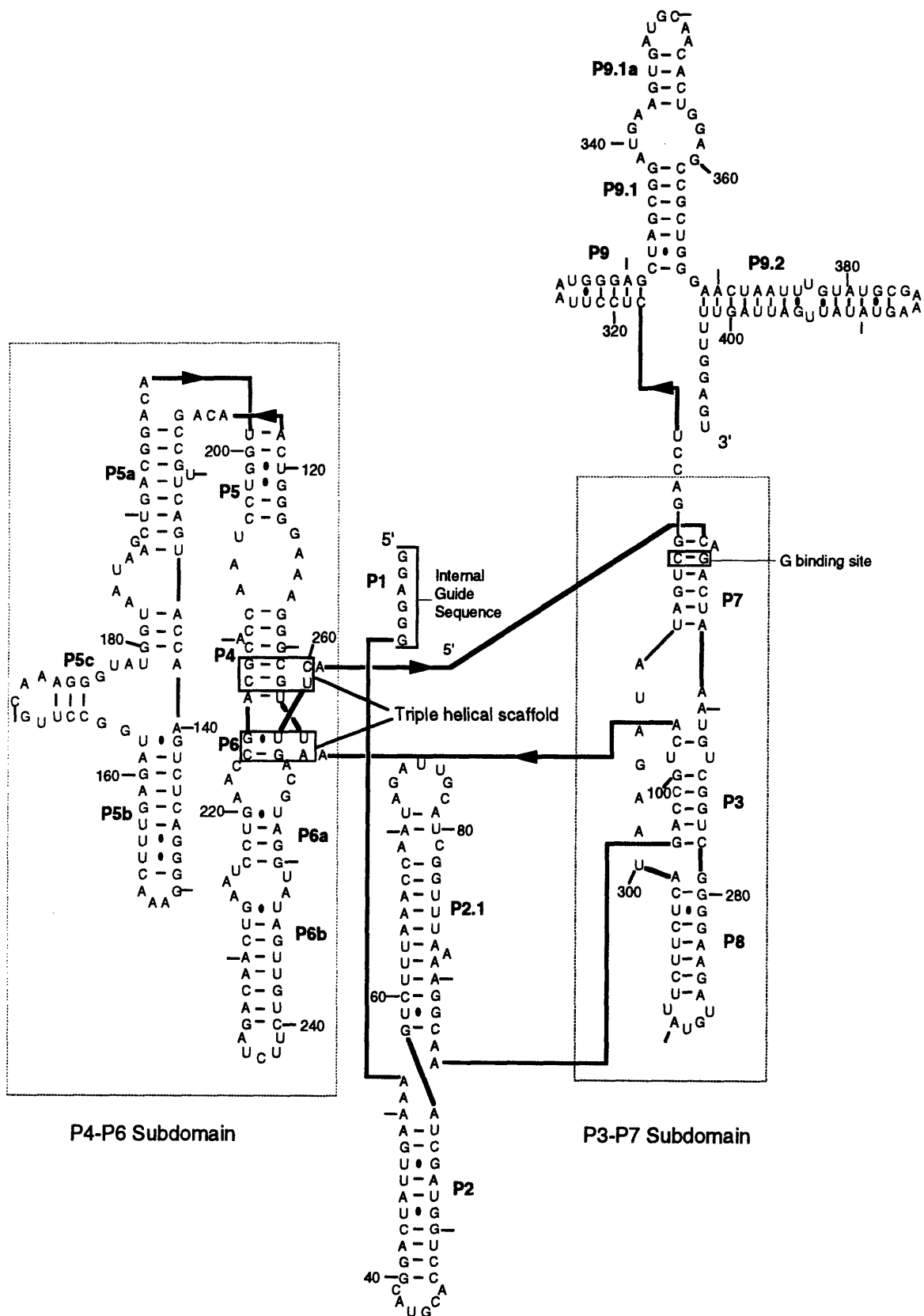


Figure 5. (C) *Tetrahymena* ribozyme secondary structure.
Representation according to Cech *et al.* (1994).

most group I introns contain additional helices peripheral to the core that are less conserved and define subclasses of related introns (such as the P9.1, P9.2 and P5abc interactions in the *Tetrahymena* intron) (Michel and Westhof, 1990). P1 is the interaction between the internal guide sequence and the substrate, and is not present in the L-21 ribozyme in the absence of substrate.

Similar techniques used to derive the secondary structure were employed to obtain a model for the three dimensional architecture of the core of the group I intron (Fig. 5B) (Michel and Westhof, 1990). This model is of low resolution, and does not include peripheral elements, but so far has been supported by experimental results (Downs and Cech, 1990; Michel *et al.*, 1990; Christian and Yarus, 1992; Wang and Cech, 1992; Chen *et al.*, 1993; Christian and Yarus, 1993; Murphy *et al.*, 1994). It is likely that the actual structure of the core region closely resembles this model. In the model it is apparent that there are two main structural subdomains which together form the core of the ribozyme (Michel and Westhof, 1990). One subdomain includes the interactions P4, P5 and P6 (the P4-P6 subdomain), and the second includes P3, P7 and P8 (the P3-P7 subdomain). Electron micrographs have shown that the entire RNA forms a globular structure in the presence of Mg^{2+} (Wang *et al.*, 1994; Nakamura *et al.*, 1995). While the two structural subunits have so far mostly been referred to as domains, the term subdomains is more analogous to the terminology used in the protein field, where a domain is a globular structure, and they will therefore be referred to as subdomains here. The formation of a globular structure is also indicated by the observation that in the presence of Mg^{2+} , regions in the core of the RNA are protected from cleavage by hydroxyl radicals produced by Fe(II)-EDTA (Latham and Cech, 1989). Similar protection is not observed even at high concentrations of Na^+ , demonstrating that, unlike in tRNA, Mg^{2+} is required for higher order folding (Celander and Cech, 1991). This role of Mg^{2+} is supported by the

localization of Mg^{2+} binding sites to regions where the backbones of the two subdomains approach each other in the three dimensional model (Christian and Yarus, 1993).

Consistent with the notion that secondary structure can form in the absence of Mg^{2+} , probing base pairing with small chemical probes showed that high concentrations of Na^+ (up to 1 M) do allow formation of most of the helical elements (Jaeger *et al.*, 1990).

Each of the subdomains is formed by stacking of a set of helical elements, and the active site is formed in a cleft between the two subdomains (Michel and Westhof, 1990; Murphy *et al.*, 1994). The guanosine cofactor binding site, which had previously been identified as being in the P7 helix, is positioned close to the active site and the scissile bond in the structural model (Michel *et al.*, 1989a; Yarus *et al.*, 1991). Its position with respect to other parts of the structure was confirmed by chemical cleavage studies using guanosine linked to Fe(II)-EDTA (Wang and Cech, 1992). A more recent, alternative representation of the secondary structure (Fig. 5C) more closely reflects the accumulated knowledge of the three dimensional structure, and more clearly outlines the modular organization of the intron (Cech *et al.*, 1994).

Phylogenetic and mutational studies have suggested the existence of a triple helical scaffold formed by interactions between base pairs in P4 and P6 and nucleotides in adjacent single stranded regions (Flor *et al.*, 1989; Michel *et al.*, 1990). P4 and P6 are believed to stack upon each other (Murphy *et al.*, 1994), and triple interactions are thought to involve two base pairs in P4 and two base pairs in P6, forming four contiguous levels of triples (Michel *et al.*, 1990; Green and Szostak, 1994). It is this triple helical scaffold which connects the two subdomains and most likely defines their orientation with respect to each other. Mutagenesis experiments suggested that the interactions consist of conventional base triples (Michel *et al.*, 1990). NMR studies of model compounds, however, revealed the alternative possibility that they may involve interactions between bases and the

backbone as well (Chastain and Tinoco, 1992; Chastain and Tinoco, 1993). These studies, along with *in vitro* selection experiments (Green and Szostak, 1994), also suggest that the triples may form cooperatively, so that disruption of one level will lead to destabilization of at least one additional level. The importance of the triple helical scaffold was underscored recently by the demonstration that it is required for the formation of an active complex if the two subdomains are synthesized as separate molecules and allowed to associate through tertiary interactions only (Doudna and Cech, 1995). If the subdomains were designed such that the scaffold was unable to form, they could still associate, but much more weakly than in the presence of the scaffold, and the resulting complex was inactive.

The three dimensional model does not include a number of peripheral helical elements that are phylogenetically less conserved, and little is known about their structure. Deletion experiments have established that none of the extensions are absolutely required for catalytic activity (Doudna and Szostak, 1989; Beaudry and Joyce, 1990; van der Horst *et al.*, 1991; Jaeger *et al.*, 1993; Laggerbauer *et al.*, 1994). In their absence, however, the structure of the intron core can be greatly weakened, and increased concentrations of Mg^{2+} are required to preserve some activity.

The most well defined extension is the P5abc extension that is part of the P4-P6 subdomain. A bend has been identified in the internal loop between P5 and P5a which allows an interaction between the loop capping P5b and a base pair in P6a, stabilizing the P4-P6 subdomain (Murphy and Cech, 1993; Murphy and Cech, 1994). It has also been demonstrated that the entire P5abc extension can be replaced by the CYT-18 protein (Mohr *et al.*, 1994). This protein is a tyrosyl-tRNA synthetase and is required for activity in a group I intron from *Neurospora*. In its cognate intron, which does not possess a P5abc extension, the protein binds to the RNA core, stabilizing it (Guo and Lambowitz, 1992; Mohr *et al.*, 1992). Completely unrelated RNA sequences have also been identified which

can substitute for the wild type sequence of P5abc, reinforcing the modular organization of the intron (Williams *et al.*, 1994). Together, these data indicate that one primary role of the P5abc extension is to stabilize the core of the ribozyme.

A second extension at the 3' end of the *Tetrahymena* intron which includes the P9.1 and P9.2 helices (the P9.1-P9.2 extension) has also been investigated more specifically (Banerjee *et al.*, 1993; Lagerbauer *et al.*, 1994). While deletion of this extension increases the Mg^{2+} requirement for stable formation of the core structure, at saturating Mg^{2+} concentrations the catalytic activity of the ribozyme appears undiminished compared to wild type (Lagerbauer *et al.*, 1994).

Binding of the oligoribonucleotide substrate to the ribozyme occurs in two steps. The substrate first base pairs to the internal guide sequence, resulting in an 'open' complex, before the resulting P1 helix then docks into the active site to form a 'closed' complex (Young *et al.*, 1991; Bevilacqua *et al.*, 1992; Herschlag, 1992; Wang *et al.*, 1993). A significant fraction of the binding energy in the closed complex is not derived from base pairing, but from tertiary interactions between the P1 helix and the core of the ribozyme (Pyle *et al.*, 1990). Interactions occur between phosphates on both strands of the helix and bases in the core (Bevilacqua and Turner, 1991; Pyle and Cech, 1991; Pyle *et al.*, 1992; Herschlag *et al.*, 1993; Strobel and Cech, 1993; Strobel and Cech, 1994), and also involve the exocyclic amine of the G in the G•U wobble pair preceding the cleavage site (Knitt *et al.*, 1994; Pyle *et al.*, 1994; Strobel and Cech, 1995; Strobel and Cech, 1996). It is this exocyclic amine that defines the cleavage site. A stopped flow fluorescence study of substrate binding showed that the two binding steps proceed rapidly (in milliseconds), and can be kinetically resolved (Bevilacqua *et al.*, 1992). In *Tetrahymena* pre-rRNA, base pairing of nucleotides preceding the 5' splice site to the IGS to form P1 competes with an

alternative interaction of these nucleotides with upstream sequences (Woodson and Cech, 1991). The need to disrupt this interaction can slow the rate of splicing.

The tertiary interactions between P1 and the core in part explain the slow rates of product release and substrate dissociation, which are much slower than expected for dissociation of a six base-pair interaction (Herschlag and Cech, 1990a). As a result, substrate cleavage is much faster than dissociation so that even substrates that form mismatches with the guide sequence, resulting in weaker binding and faster dissociation, are still cleaved very efficiently (Zaug *et al.*, 1988; Herschlag and Cech, 1990b). Sequence specificity of substrate cleavage can be increased by shifting the equilibrium between the open and closed complexes of substrate and ribozyme in favor of the open complex, allowing substrates to dissociate more rapidly (Herschlag, 1991). This can be accomplished by changing the length of the single stranded region between the IGS and P2 (Young *et al.*, 1991), which serves as a tether between the P1 helix and the core, or directly by eliminating tertiary interactions between P1 and the core (Pyle and Cech, 1991; Pyle *et al.*, 1992). Weakening the docking of the P1 helix into the active site cleft does, however, result in a decrease of the fidelity of substrate cleavage, reflecting the importance of interactions between P1 and the core not only in stabilizing the closed complex, but also in defining the register in which P1 docks (Young *et al.*, 1991; Herschlag, 1992; Strobel and Cech, 1994).

Folding

The Mg²⁺ requirement for higher order structure formation in the *Tetrahymena* ribozyme has allowed the examination of its equilibrium folding pathway by probing the structure of the RNA at different concentrations of Mg²⁺ (Celander and Cech, 1991). Protection from Fe(II)-EDTA cleavage was observed in the P5abc extension at slightly lower concentrations than in other regions of the molecule, leading to the proposal of an

equilibrium folding intermediate in which only the P5abc extension is properly folded (Celander and Cech, 1991). The existence of such an intermediate was further confirmed when Mg^{2+} -induced folding of the P4-P6 subdomain was examined in isolation (Murphy and Cech, 1993; Murphy and Cech, 1994). This study showed first, that this subdomain can fold into a higher order structure in the absence of the rest of the molecule, and that this structure appears similar to that formed in the context of the whole intron and second, that folding of the P5abc extension into a discrete higher order conformation appears to be required for structure formation in the rest of the P4-P6 subdomain. This hierarchical relationship was confirmed and extended by mutational studies in the context of the entire intron which revealed that here also, the P5abc extension must form before the P4-P6 subdomain can fold, and further that the P4-P6 subdomain must fold before the remainder of the core, including the P3-P7 subdomain (Laggerbauer *et al.*, 1994). These observations define an equilibrium folding pathway of the ribozyme at different Mg^{2+} concentrations that parallels very closely the kinetic pathway outlined in the following chapters.

When melting of the structure of the ribozyme is monitored by following changes in UV absorbance, a minor low temperature transition is followed by a major transition at higher temperature (Banerjee *et al.*, 1993). In combination with chemical modification experiments it was established that the low temperature transition represents loss of tertiary structure, while the high temperature transition is the concerted loss of secondary structure. Similar to what had been observed in tRNA, tertiary structure is therefore less stable than secondary structure. In particular, it was shown that an interaction between the 3' terminal P9.1-P9.2 extension and nucleotides in the loop of P2 or P2.1 is lost during the low temperature transition (Banerjee *et al.*, 1993). Equivalent studies on another group I intron gave similar results, with the low temperature transition again being due to interactions

between a 3' terminal element of the intron with nucleotides close to the core (Jaeger *et al.*, 1993). More recently, chemical modification agents were also used to follow kinetic folding of the RNA (Banerjee and Turner, 1995). Folding rates were deduced from changes in the rate of modification of the RNA upon folding. While these experiments were performed at low temperatures where the intron is not very active, and give only an indirect measure of folding itself, they did suggest that different parts of the RNA fold at different rates, and that folding of some regions can be very slow. Slow folding was found to be consistent with measures of the rate of gain of full catalytic activity of the ribozyme.

Together, these studies suggest that folding of the *Tetrahymena* ribozyme may be a complex process including one or more intermediates, but provide only very little information about kinetic folding. The early observation that, to obtain fully active ribozyme at 50°C, the RNA must be preincubated in Mg²⁺ for several minutes before the reaction is initiated by adding substrate suggested that even at functionally relevant temperatures formation of the active structure of the ribozyme might be rather slow (Herschlag and Cech, 1990a). More explicit kinetic investigations were required, however, to begin to define a kinetic folding pathway for this RNA, and gain an understanding of the rules governing the process of higher order RNA folding.

2. A KINETIC OLIGONUCLEOTIDE HYBRIDIZATION ASSAY TO STUDY RNA FOLDING

Concept

Measuring folding rates

The stability of the native conformation of large, highly structured RNA molecules generally depends on divalent metal ions, usually Mg^{2+} (Pan *et al.*, 1993; Pyle, 1993) (see chapter 1). In the absence of divalent metals, many RNAs exist in a partially denatured state where some local secondary structure elements, such as hairpins, are stable, but many tertiary interactions are not. The addition of Mg^{2+} can thus be used as a means to initiate higher order folding and formation of the native structure.

At equilibrium, Mg^{2+} -induced conformational changes in the structure of group I introns and other RNAs can readily be detected using conventional chemical probes such as dimethyl sulfate, diethyl pyrocarbonate (Jaeger *et al.*, 1990; Jaeger *et al.*, 1993), and Fe(II)-EDTA mediated hydroxyl radical cleavage (Latham and Cech, 1989; Celander and Cech, 1991; Heuer *et al.*, 1991), as well as by ultraviolet crosslinking (Wang and Cech, 1994). Following the timecourse at which these changes take place, however, is less easily accomplished. To measure folding rates directly, the time required to probe the RNA conformation must be short compared to the time it takes the RNA to fold. Current protocols for all the above techniques require that the RNA is exposed to the probe or to UV light for relatively long periods, making them ill-suited for kinetic experiments.

Spectroscopic techniques to follow higher order folding are also not readily available. The size of RNAs of interest and the absence of protons with the appropriate exchange properties preclude NMR analysis by hydrogen/deuterium exchange similar to that successfully used to study protein folding. Changes in UV absorbance report mainly the formation of secondary structure, and are not very sensitive to higher order folding. There are no intrinsic fluorescent moieties in RNA, and the application of fluorescence spectroscopy would require the introduction of chemical groups that can act as fluorescent probes. Therefore, none of the standard techniques used to study RNA structure and conformation are readily amenable to following the kinetics of higher order folding.

The base pairing potential of RNA suggested an alternative strategy to circumvent these difficulties. It was reasoned that if Mg^{2+} induced significant changes in the accessibility of the RNA to small chemical probes, hybridization of complementary oligodeoxynucleotides should also be affected. Upon Mg^{2+} -induced folding, parts of the RNA sequence previously accessible should become inaccessible to binding by short DNA oligonucleotides. In earlier studies, the hybridization of such probes to structured RNAs under various equilibrium conditions had been examined (Uhlenbeck *et al.*, 1970; Weller and Hill, 1992; Cload *et al.*, 1993; LeCuyer and Crothers, 1994), and significant changes in accessibility were detected. Hybridization of two complementary strands of nucleic acids is known to be very rapid, and potentially faster than tertiary structure formation, suggesting that this technique might allow a kinetic analysis of higher order folding. RNA bound by oligonucleotide probes can be detected by making use of the enzyme RNase H, which selectively cleaves the RNA strand of a RNA:DNA helix, and will therefore cleave the RNA at the site of probe hybridization.

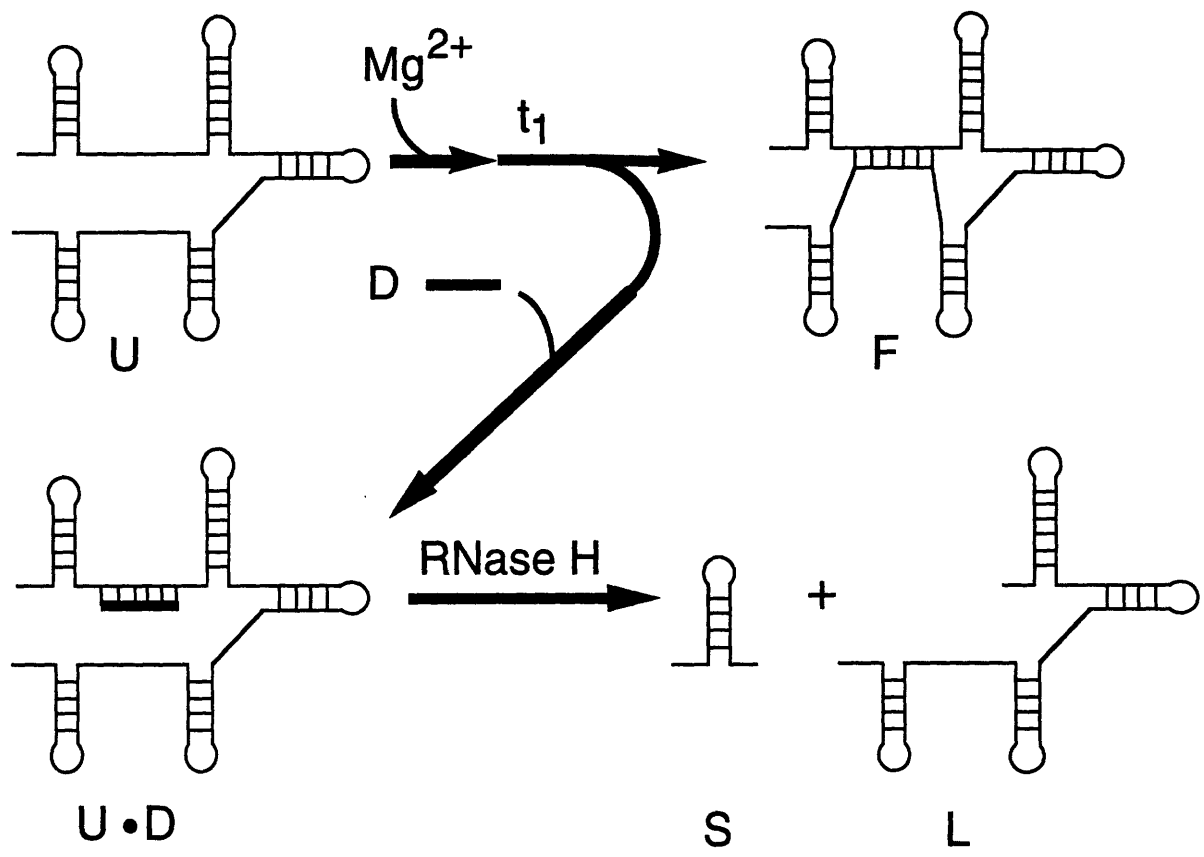


Figure 6. Schematic representation of the kinetic oligonucleotide hybridization assay.

Experimental approach

An experimental protocol was designed which allows measurement of the time dependence of changes in the accessibility of the RNA to specific oligonucleotide probes (Fig. 6). Folding is initiated by the addition of Mg^{2+} to ^{32}P -labeled RNA equilibrated in the absence of Mg^{2+} . Samples from the folding reaction are taken after increasing times t_1 , and added to an excess of oligonucleotide probe and RNase H, setting up a kinetic competition between continued folding and probe binding. Under the appropriate conditions the fraction of the RNA that remains unfolded (U) is rapidly bound by the probe (D) and cleaved by RNase H at the site of hybridization (resulting in fragments S and L). RNA that has already folded (F) is inaccessible, and remains uncleaved. Electrophoretic separation of cleaved and uncleaved RNA allows quantitation of the fraction cleaved, and thus the fraction unfolded, as a function of time.

The most revealing application of oligonucleotide hybridization and subsequent RNase H cleavage is a rapid quench experiment to measure the fraction of RNA that remains unfolded at any given time after addition of Mg^{2+} . The rate of decrease of this fraction, manifested by the amount of RNA cleaved by RNase H, directly corresponds to the folding rate of the RNA (Fig. 7). For the fraction of RNA cleaved in the assay to be equal to the fraction still unfolded at the time oligonucleotide and RNase H are added, and thus to obtain a snapshot of progress toward the folded state, four kinetic conditions must be met. First, the rates of probe hybridization and RNase H cleavage must be rapid compared to the folding rate. Second, sufficiently high probe concentrations must be used so that all accessible RNA is bound. Third, RNase H cleavage must be much faster than dissociation of the RNA•probe hybrid. Fourth, unfolding under native conditions must be slow compared to folding.

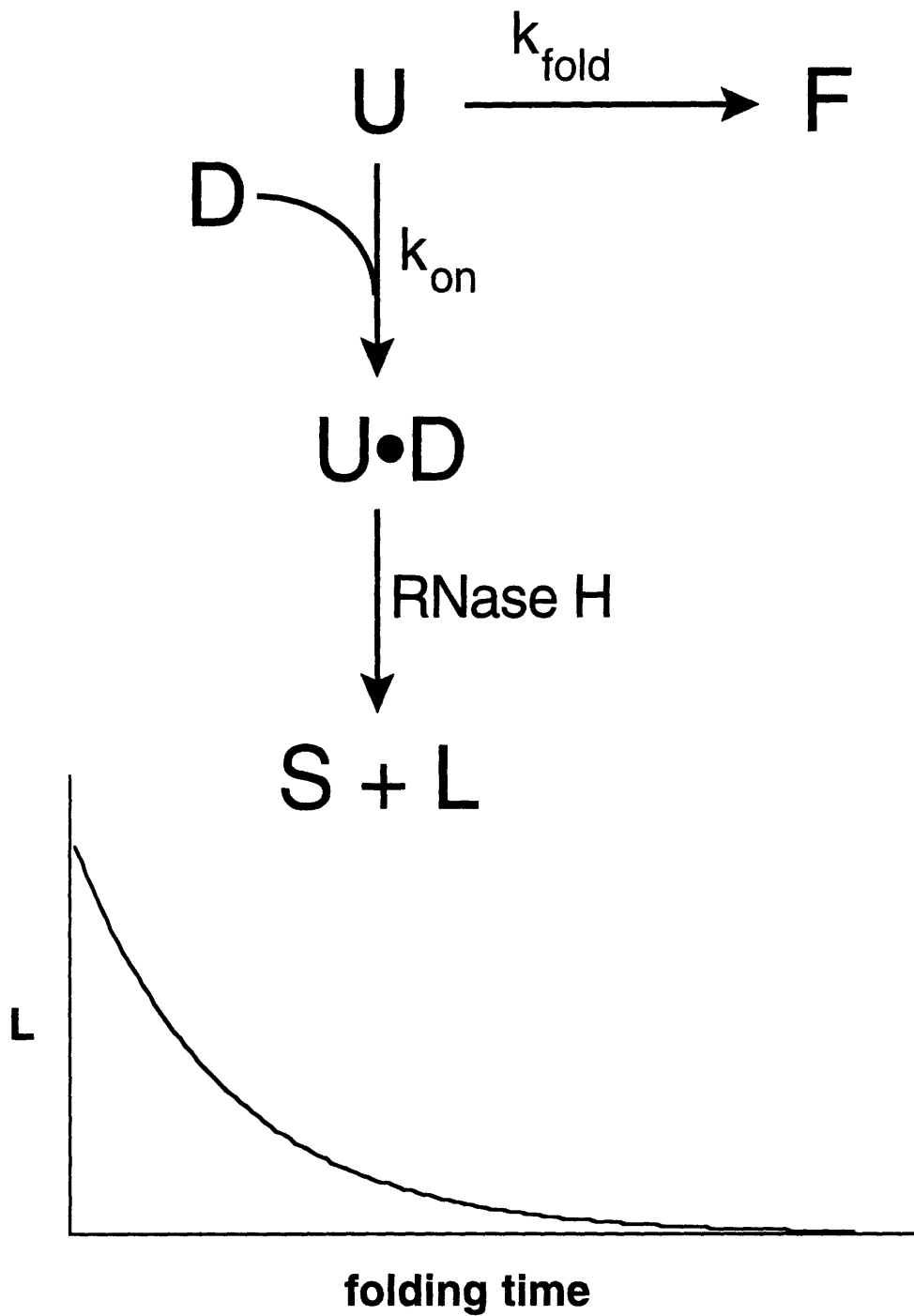


Figure 7. Kinetic scheme of the oligonucleotide hybridization assay as a rapid quench experiment.

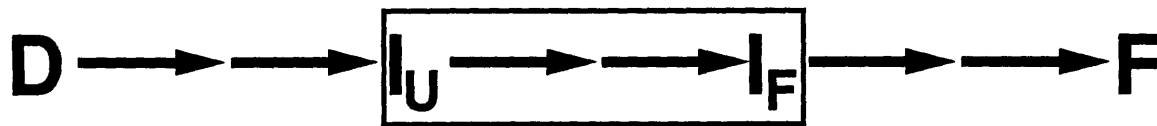


Figure 8. Only a part of the folding pathway of RNA can be probed by the kinetic oligonucleotide hybridization assay.

Limitations

In principle, if oligonucleotides complementary to different regions are used, it should be possible to probe folding of each part of a large RNA specifically and independently. In practice, only those sequences can be studied which are accessible to DNA probe hybridization and RNase H cleavage in the absence, but not in the presence, of Mg^{2+} , and where binding of the probe can compete successfully with structure formation when probe and Mg^{2+} are added at the same time. Therefore, the entire folding process cannot be followed from the denatured (D) to the native (F) state, but only the progression from the Mg^{2+} -free form of the RNA (I_U) to the first folding intermediate in which binding of the DNA probe or cleavage by RNase H, or both, are blocked (I_F) (Fig. 8). There certainly are steps preceding the addition/binding of Mg^{2+} , including the formation of most short range secondary structure, and there may be further rearrangements after the formation of the first inaccessible intermediate. Furthermore, only those folding events accompanied by a change in accessibility can be detected. Thus, only a part of the folding pathway can be explored by following Mg^{2+} -induced changes in the accessibility of RNA to external probes.

Feasibility

To determine whether there are regions of the *Tetrahymena* ribozyme suitable for kinetic analysis by the oligonucleotide hybridization assay, 21 oligonucleotides complementary to different sequences within the ribozyme were screened to survey most of the conserved core (Fig. 9). Each oligonucleotide was tested for its ability to hybridize, in competition with structure formation, when added at high concentration simultaneously with Mg^{2+} . A second experiment tested each probe for binding after the RNA had been

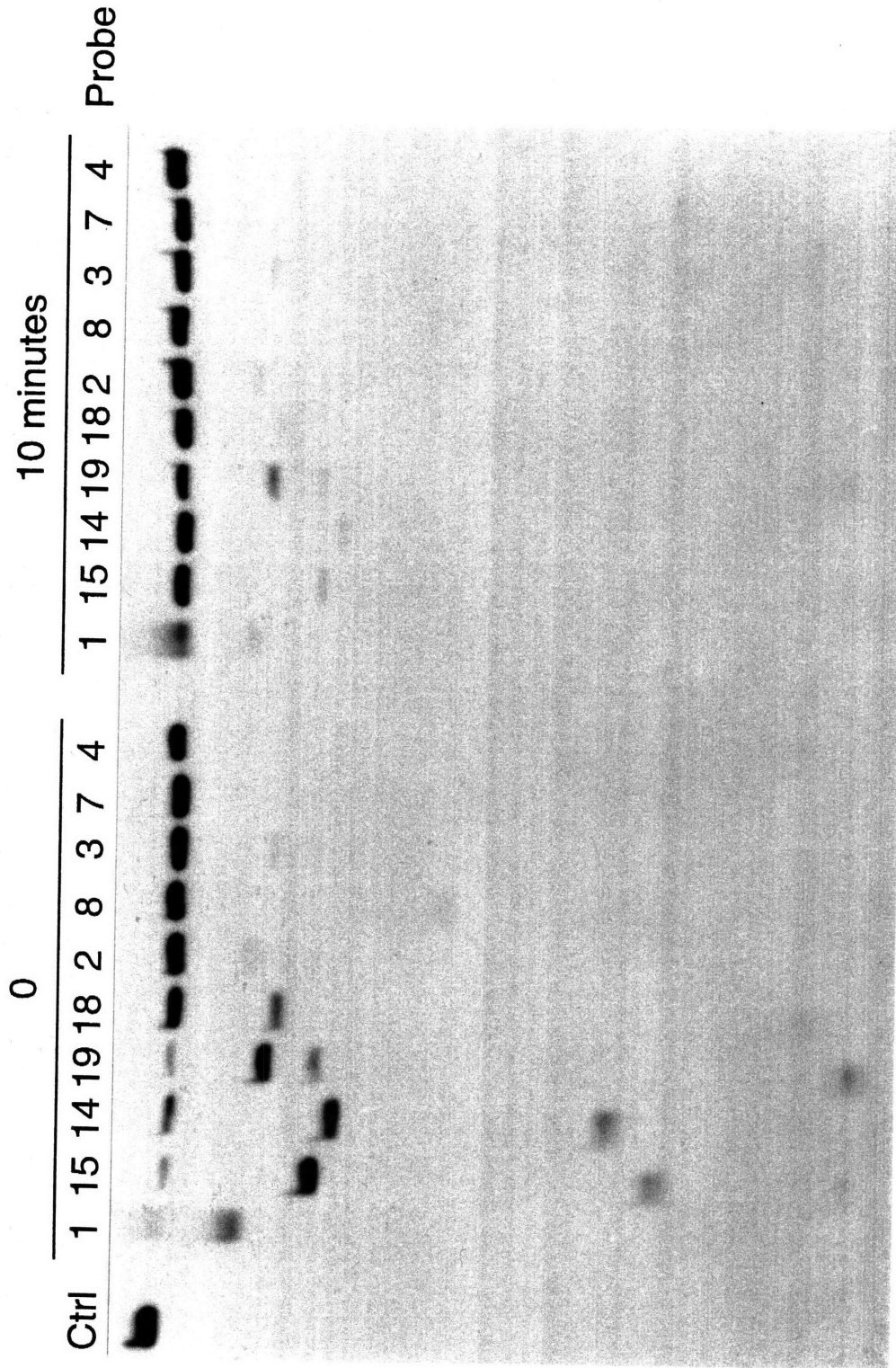


Figure 10. Initial screening of oligodeoxynucleotide probes complementary to the *Tetrahymena* ribozyme. Probes were either added to the RNA together with Mg^{2+} (0 minutes), or after the RNA had been equilibrated in Mg^{2+} for ten minutes. Final probe concentrations were $\sim 300 \mu M$.

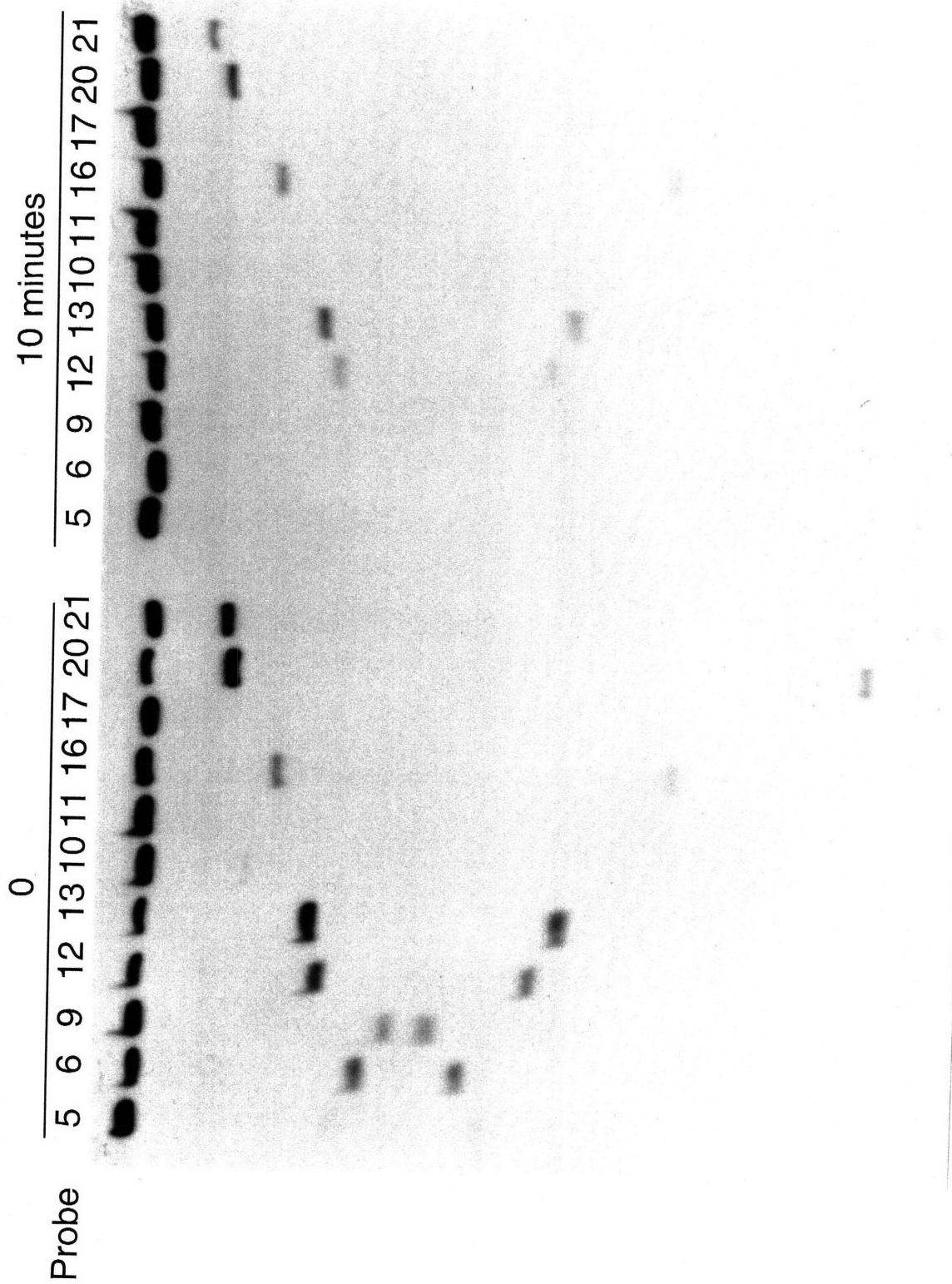


Figure 10. (contd.)

equilibrated in Mg^{2+} for ten minutes. These two extremes represent the initial and final points of a folding time course.

The probes can be grouped into three classes (Fig. 10). The first class gave no observable RNase H cleavage when added with, or after, Mg^{2+} . Members of this class generally target secondary structures that would be expected to form in the absence of Mg^{2+} (such as probes **10**, **11** and **17**), and the lack of cleavage was therefore not surprising. Although the failure to elicit cleavage likely is due to the target site being unavailable for probe binding in the absence of Mg^{2+} , it may also be due to Mg^{2+} -induced structure formation in these regions being much faster than oligonucleotide hybridization, or to the RNA•DNA complex being sterically inaccessible to RNase H. The second class competed successfully with structure formation when added simultaneously with Mg^{2+} , resulting in almost complete cleavage of the RNA, but gave almost no cleavage after equilibration in Mg^{2+} . This class includes probes targeting sequences involved in long-range interactions, such as the P3 and P7 helices (probes **1**, **14**, **15** and **19**; see Fig. 9), that might be expected to require Mg^{2+} for their stabilization, if not their formation. The third class of probes gave incomplete cleavage at the zero time point, and significant remaining cleavage at the ten minute time point (for example probes **6**, **9**, **12** and **18**). The observation of these three different classes of probe behavior suggested that there are distinguishable folding events that can be addressed by oligonucleotide hybridization.

Controls

A series of control experiments were performed to determine whether the kinetic conditions outlined above were met for the probes exhibiting the most significant changes in accessibility. To ensure that probe binding was fast and that all the accessible RNA was bound by probe, a series of concentrations of oligonucleotide was tested in experiments equivalent to the zero time point described above to determine the minimal concentration

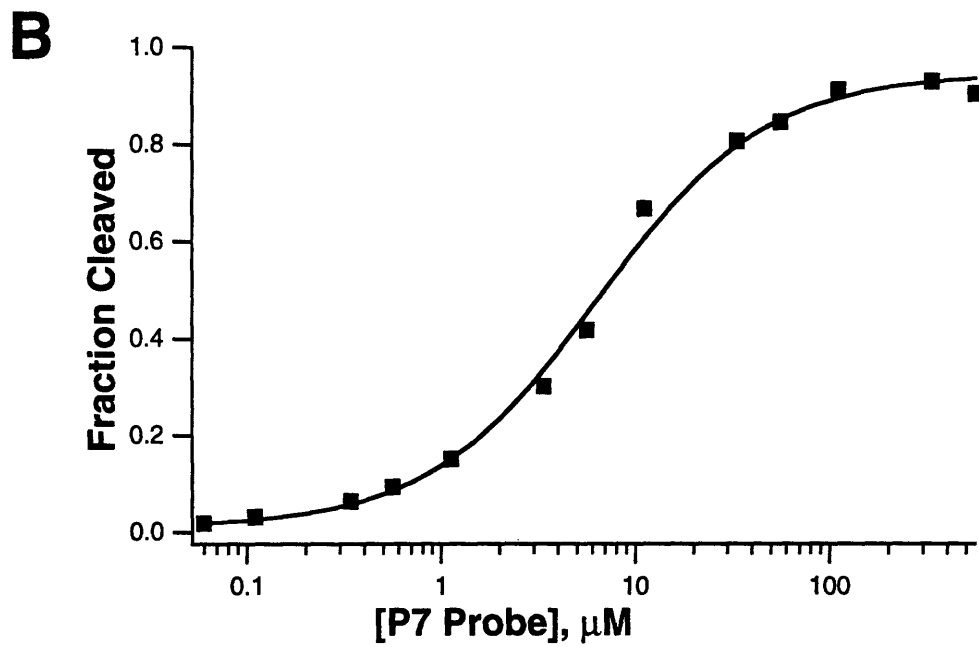
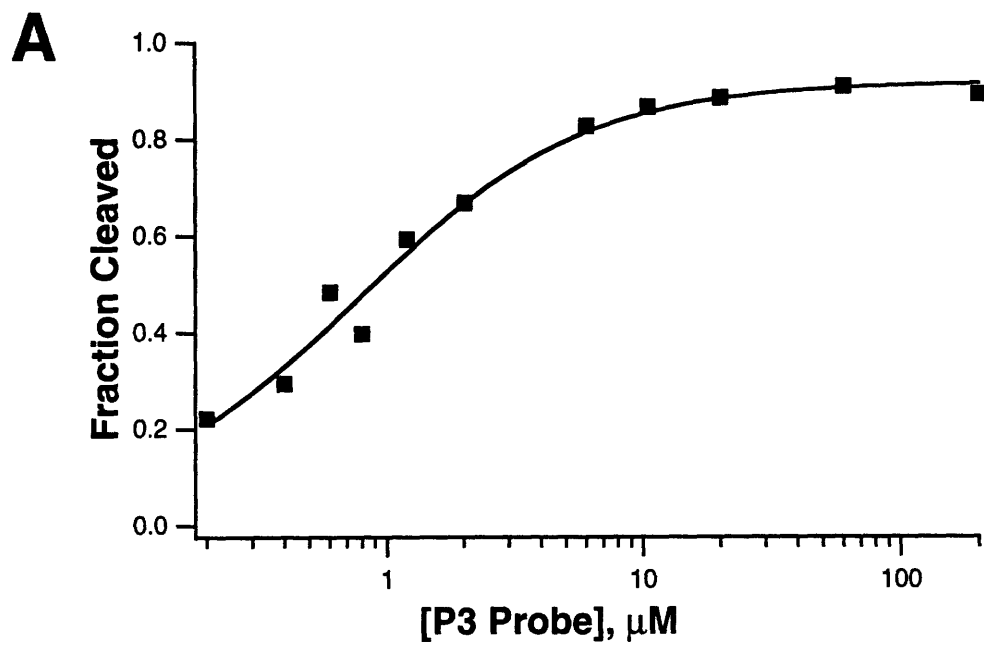


Figure 11. Oligonucleotide probe concentration dependence.
(A) Probe 15, targeting P3.
(B) Probe 19, targeting P7.

necessary for maximal cleavage. At 20 to 70 μM oligonucleotide more than 90 percent of the RNA was cleaved for probes targeting P3 and P7 (Fig. 11, probes **15** and **19**). Similar oligonucleotide concentrations for P4 and P6a probes (probes **8** and **12**) produced lower, but still saturated, cleavage. To ensure that RNase H cleavage was fast, reactions equivalent to the zero time point described above were performed in larger volumes using varying concentrations of RNase H, and aliquots removed and quenched after different times to determine the RNase H concentration necessary to produce maximal cleavage rapidly. At 0.1 units/ μl RNase H (corresponding to a large stoichiometric excess of enzyme over RNA) the reaction was always complete in less than 30 seconds. Under these conditions, therefore, hybridization and cleavage were rapid compared to the measured folding rates (see chapter 3). To confirm that RNase H cleavage was faster than probe dissociation, the rate of decrease in the extent of cleavage after incubation of pre-formed probe•RNA complex under folding conditions, corresponding to the rate of probe dissociation, was measured. For probes ten nucleotides long, cleavage decreased less than 10 percent after one minute when P3, P4, or P7 were targeted (Fig. 12), and less than 20 percent after 30 seconds when P6a was targeted. Dissociation rates varied from probe to probe, and were considerably faster than expected for ten base-pair helices, suggesting that the RNA can actively displace the oligonucleotides. Even the accelerated dissociation rates, however, are slower than RNase H cleavage (see above). A probe length of ten nucleotides was chosen because shorter probes did not form sufficiently stable complexes, while longer probes anneal to target sequences that contribute to more than one structural feature in the RNA. To determine whether significant unfolding of the RNA occurred under native conditions, fully folded ribozyme was incubated in the presence of excess probe and RNase H. Less than 20 percent cleavage was observed after two minutes (using probe **15**), indicating that the RNA remains in the folded, inaccessible conformation

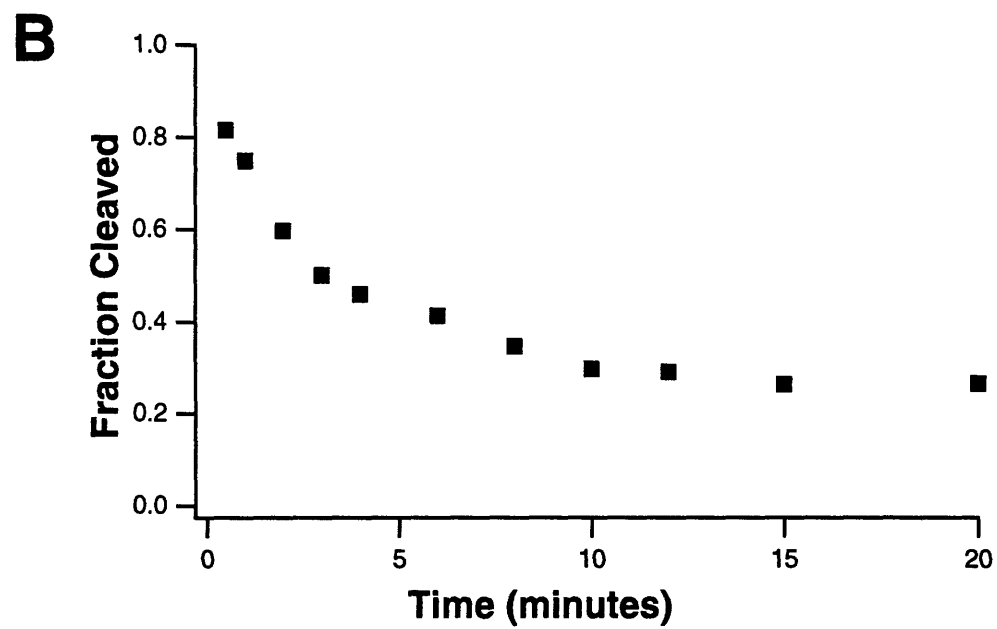
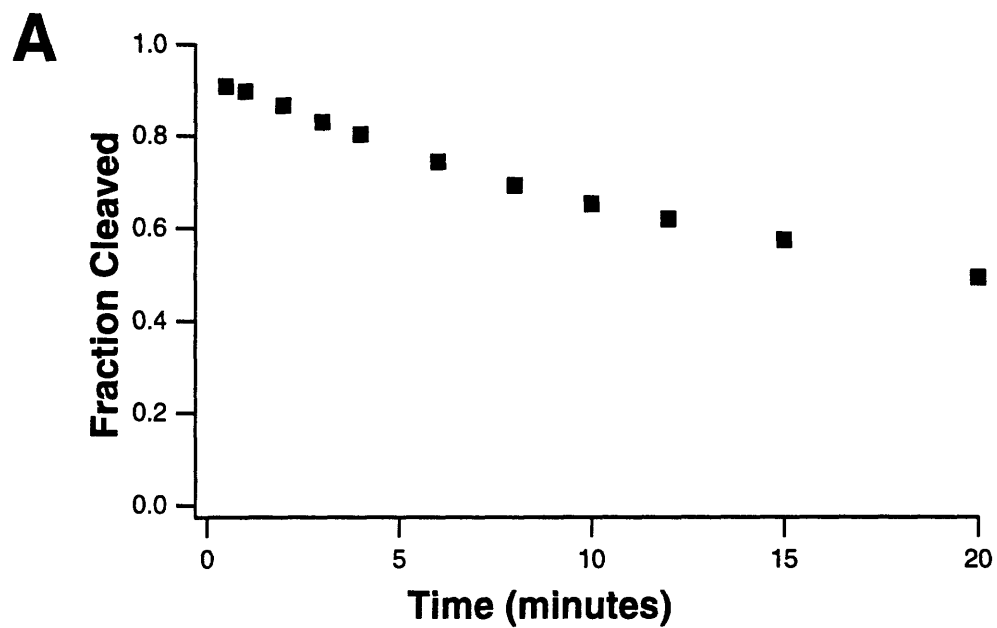


Figure 12. Oligonucleotide probe dissociation under native conditions.
(A) Probe 15, targeting P3.
(B) Probe 19, targeting P7.

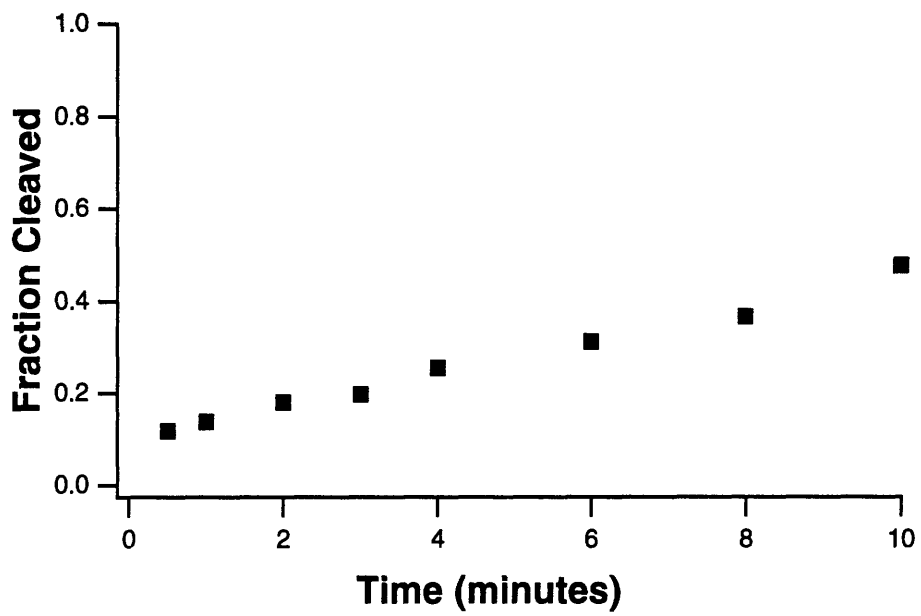


Figure 13. Unfolding of the *Tetrahymena* ribozyme under native conditions.

(Fig. 13). These results show that if probe binding and RNase H cleavage are performed for 30 seconds at high concentrations of oligonucleotide and RNase H, all the kinetic conditions outlined above are met, and the fraction of RNA cleaved corresponds directly to the fraction that remains unfolded at the time of probe addition. Rapid quench experiments performed under these conditions with probes targeting P3, P7, P4 or P6a therefore yield accurate information on ribozyme folding rates.

3. KINETIC INTERMEDIATES ON THE FOLDING PATHWAY OF THE *TETRAHYMENA* RIBOZYME

Identification of intermediates

In the initial screening (see chapter 2), oligonucleotides complementary to both strands of the P3 helix (probes **1**, **15**) showed the second class of behavior (Fig. 10), with accessibility to probe binding being Mg^{2+} -dependent. P3 is part of a pseudoknot constituting one subdomain of the core of the intron and is essential for ribozyme activity (Burke *et al.*, 1986; Williamson *et al.*, 1987; Couture *et al.*, 1990). It is a long-range base-pairing interaction, with its two strands far apart in the linear sequence of the RNA (see Fig. 5A), and it was therefore anticipated that it might exhibit interesting folding behavior. To begin to explore the kinetics of folding, a detailed analysis of formation of the P3 helix was undertaken.

Slow Folding of P3

After Mg^{2+} was added to RNA equilibrated in Tris•HCl buffer, the accessibility of P3 to probe **15** decreased with time (Fig. 14). The observed decay fits well to a single exponential (Fig. 15A) and is dependent on the presence of Mg^{2+} . When folding buffer containing Na^+ only was added, no decrease in cleavage was observed. Folding could also be initiated by adding Mg^{2+} to RNA that had already been equilibrated in Na^+ . The measured rate is independent of the concentrations of RNA, oligonucleotide probe and RNase H, and the time allowed for RNase H cleavage. Between 10 and ~400 mM Na^+ the extent and rate of decrease of cleavage were constant. At higher Na^+ concentrations, the

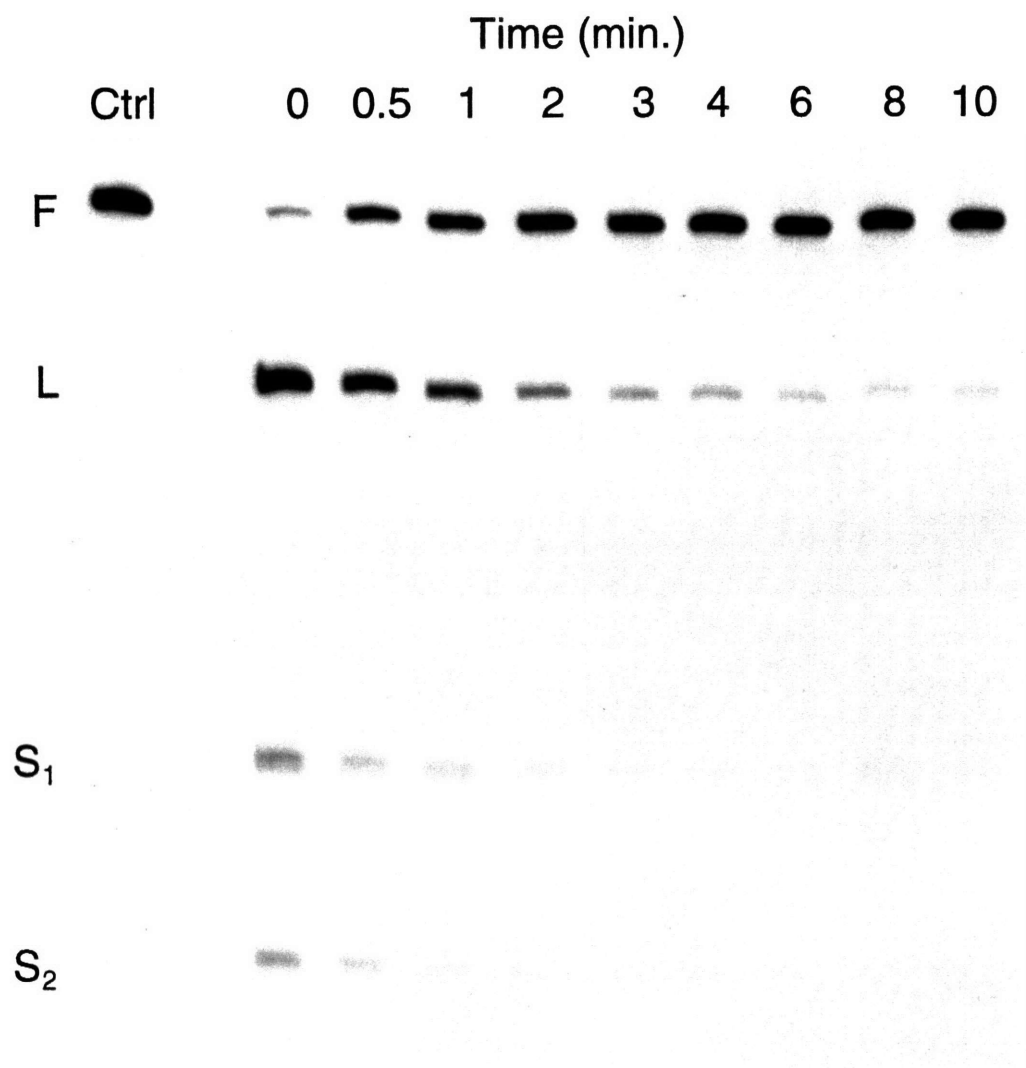


Figure 14. Folding kinetics of P3 measured by the oligonucleotide hybridization assay. Bands L, S₁ and S₂ are the products of RNase H cleavage at the site of hybridization of probe 15. Band F is the full length ribozyme.

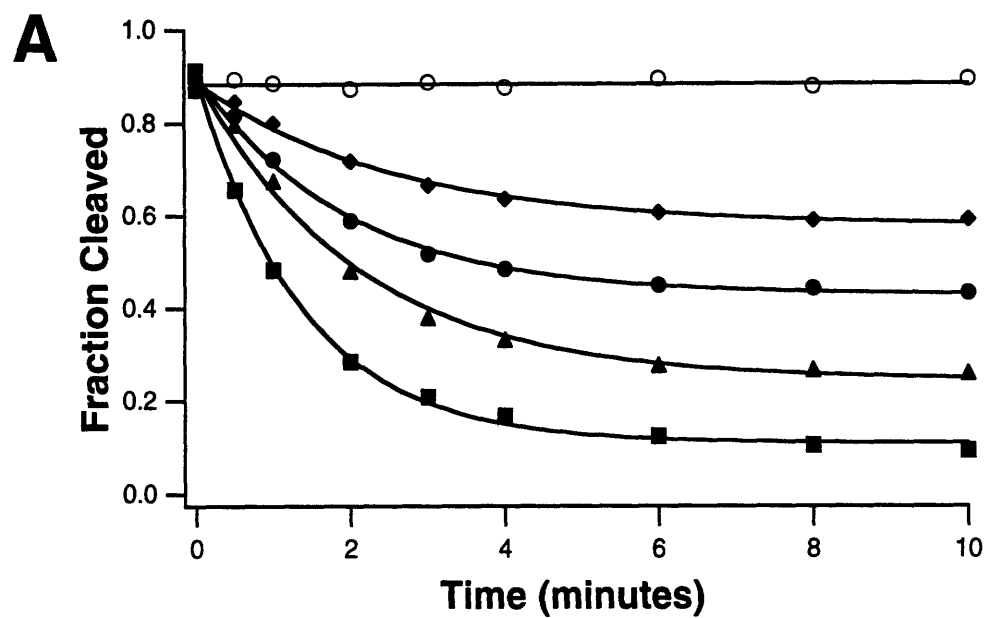


Figure 15. Formation of P3. **(A)** Kinetics of P3 formation at different Mg^{2+} concentrations. (■) 10 mM, (▲) 1.5 mM, (●) 1.0 mM, (◆) 0.8 mM, (○) 0.05 mM. Final probe concentration (probe 15) was 20 μM . The observed rate constant (k_{obs}) was obtained by a fit of the data to a single exponential.

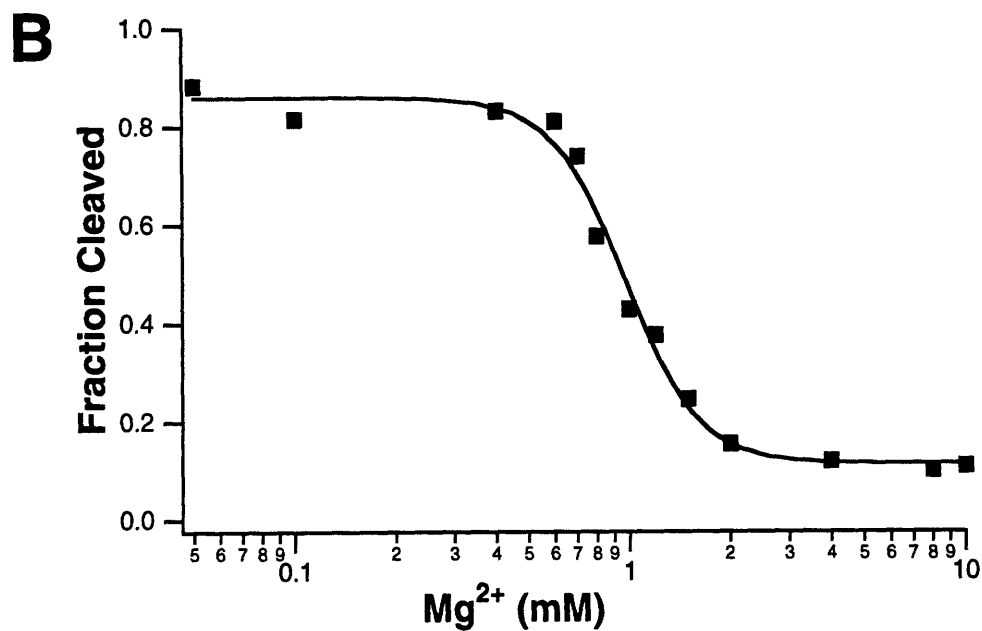
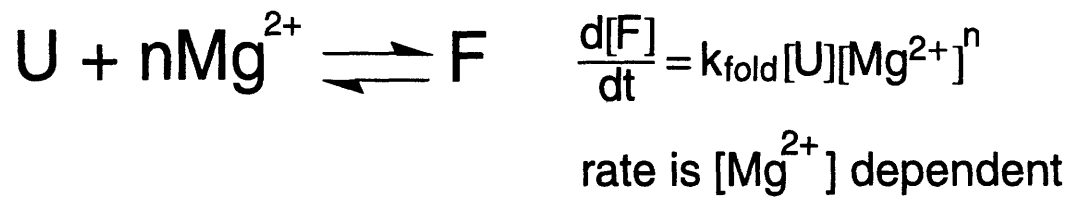


Figure 15. Formation of P3. **(B)** Equilibrium Mg^{2+} concentration dependence of P3 formation. Each point represents the endpoint of a kinetic experiment as in (A). The data were fit to an expression of simple two state binding of n Mg^{2+} ions. The fit gave values of 0.97 mM for the midpoint of the transition ($[\text{Mg}^{2+}]_{1/2}$) and 3.97 for n .

fraction of RNA competent for cleavage decreased in a gradual way, even in the absence of Mg^{2+} . This was not due to the inability of RNase H to function at such high salt concentrations (before cleavage, Na^+ was diluted to 100 mM, where RNase H still cleaves rapidly). The decreased accessibility under these high salt conditions is most likely due to Na^+ -induced stabilization of secondary structure, since Fe(II)-EDTA mapping has shown that the native tertiary structure does not form in the presence of monovalent ions only (Celander and Cech, 1991). To confirm the identity of the cleavage products, RNA was 5' or 3' end-labeled and subjected to probe binding and RNase H cleavage. The products were compared directly with an RNA sequencing ladder generated by I_2 -cleavage of phosphorothioate containing ribozyme (Rudinger *et al.*, 1992). S_2 (Fig. 14) is due to cleavage at sequences in P7 that have partial homology to the target sequence in P3. Since no intermediate resulting from cleavage at the P7 site only was ever detected, it is assumed that S_2 can only derive from S_1 , after the initial cleavage event in P3. Therefore, the presence of S_2 does not alter the quantitation of the fraction cleaved, which is obtained from the ratio of $L/(L + F)$ only (Fig. 14). Together, these results suggest that the observed rate reflects the rate limiting step for P3 folding and becoming inaccessible to probe binding and/or RNase H cleavage. Therefore, there is at least one slow step on the folding pathway of the *Tetrahymena* ribozyme which occurs on the minute time scale ($k_{obs} = 0.72 \pm 0.14 \text{ min}^{-1}$; the value shown is derived from 17 independent experiments performed over a period of 10 months with several preparations of each component). While the most straightforward interpretation is that the observed decrease in cleavage is due to actual helix formation or stabilization, other structural rearrangements may also indirectly result in the target sequence becoming inaccessible.



simple binding



binding to
intermediate

Figure 16. Minimal kinetic mechanism for P3-P7 formation. An intermediate is required to explain the $[Mg^{2+}]$ independence of the folding rate.

Requirement for a transient intermediate

The simplest mechanism that could account for the Mg^{2+} requirement is a two state transition involving the binding of Mg^{2+} upon folding. If such a one-step mechanism is correct, then the observed rate of folding will be dependent on the Mg^{2+} concentration. To test this possibility, the rate at which P3 becomes inaccessible was measured at different Mg^{2+} concentrations (Fig. 15A). At lower $[Mg^{2+}]$ the extent of cleavage once equilibrium had been reached increased, consistent with a shift in the equilibrium from the folded to the unfolded conformation. The midpoint of this transition ($[Mg^{2+}]_{1/2}$) occurred when the Mg^{2+} concentration was 0.97 mM (Fig. 15B). A similar transition is seen when the Mg^{2+} -dependence of ribozyme activity, global structure formation or active site formation is examined (Celander and Cech, 1991; Wang and Cech, 1994), indicating that these observations correlate closely with results obtained by independent methods. The rate of approach to equilibrium, however, was found to be independent of Mg^{2+} concentration (Fig. 15A). Probes to both strands of P3 (probes **1**, **15**) gave identical results, supporting the interpretation that the decrease in cleavage is due directly to formation or stabilization of the P3 helix. The only difference was that a threefold higher concentration of probe **1** was required compared to probe **15** to achieve the same extent of cleavage, and fulfill all the kinetic conditions outlined above. Apparently, the formation of P3 occurs via a more complex mechanism where rapid Mg^{2+} binding follows a slow, Mg^{2+} -independent step (Fig. 16). Thus, there is a transient intermediate on the folding pathway, and the rate limiting step is Mg^{2+} -independent.

Activity assay for folding

To compare the rate of formation of P3 with the rate at which the active structure is formed the well characterized endonuclease activity of this RNA was exploited (Herschlag and Cech, 1990a). The amount of active ribozyme can be determined from the magnitude

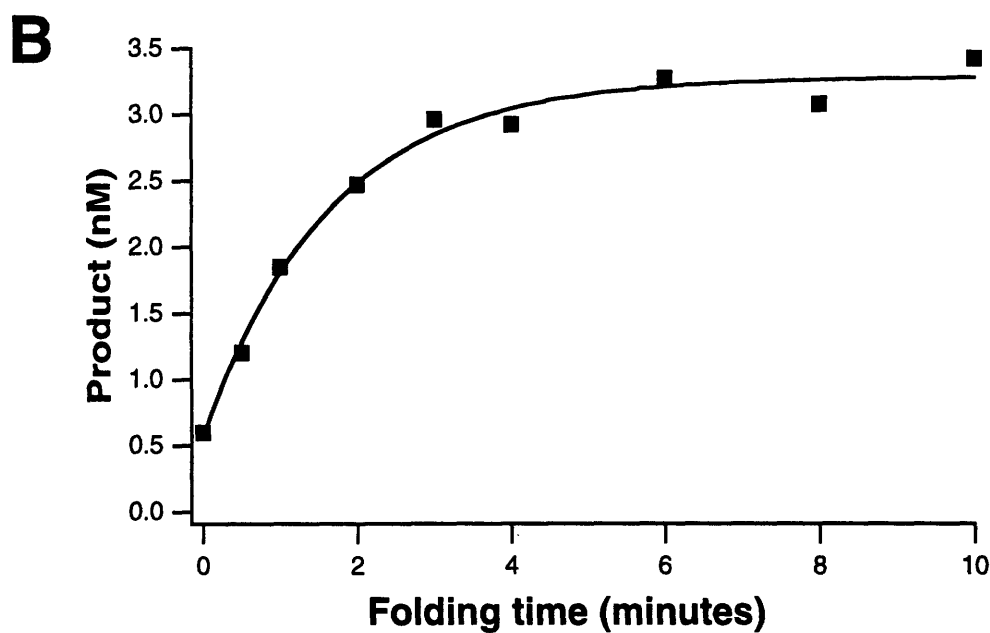
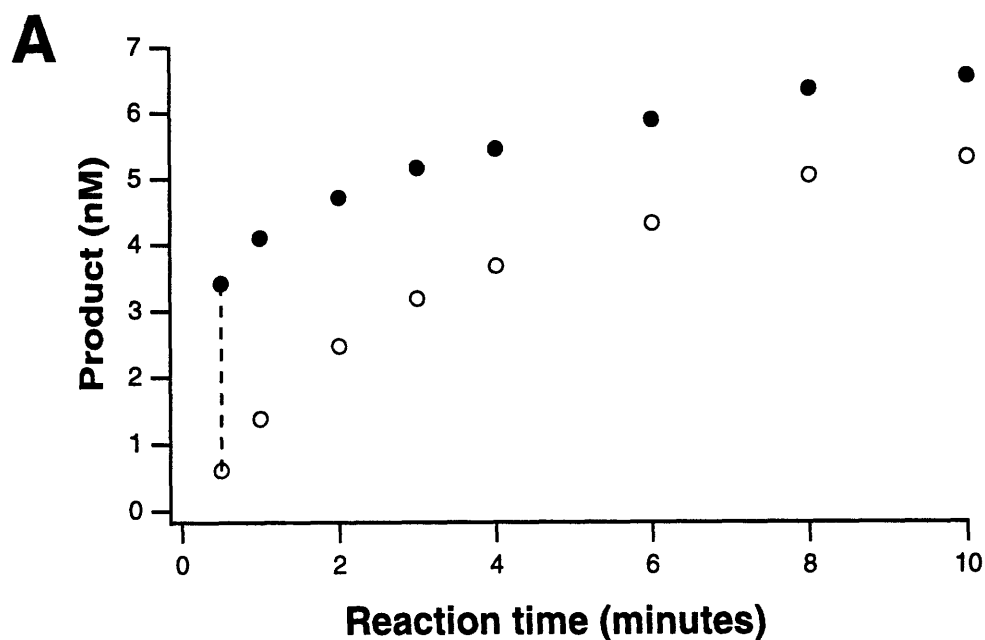


Figure 17. Ribozyme endonuclease assay for folding.
(A) Time course of substrate cleavage for folding times of 0 minutes (○) and 10 minutes (●) before the addition of substrate.
(B) Extent of substrate cleavage after 30 seconds as a function of the time allowed for the ribozyme to fold before initiation of the reaction.

of the kinetic burst corresponding to the first turnover in a multiple turnover reaction (Herschlag and Cech, 1990a). The relative amount of active ribozyme present was measured after increasing folding times following the addition of Mg^{2+} (Fig. 17). Full activity was obtained at a rate ($k_{obs.} = 0.61 \text{ min}^{-1}$) comparable to that of P3 becoming inaccessible to probe binding as determined by oligonucleotide hybridization. It was therefore concluded that P3 formation is the rate limiting step in formation of the active structure.

Folding of P7

To obtain information about the folding kinetics of different regions of the RNA, probes targeting other sequences were examined. P7 is the second stem of the central pseudoknot in the conserved core of group I introns. Together with P3, P7 forms part of one of two stacked helical subdomains in the model of the three-dimensional structure of the *Tetrahymena* ribozyme (Fig. 5C) (Michel and Westhof, 1990). In the initial screen, probes complementary to both strands of P7 (probes 14, 19) fell into the second class (Fig. 10), with accessibility to probe binding and/or RNase H cleavage being Mg^{2+} -dependent. In all experiments defining the kinetic mechanism of P3 formation, the probes for P7 gave results similar to those obtained with probes targeting P3. In 10 mM Mg^{2+} , the rate constant of P7 becoming inaccessible to oligonucleotide probe and/or RNase H cleavage was 0.74 min^{-1} , and this rate constant remained constant, within experimental error, at lower Mg^{2+} concentrations (Fig. 18A). The midpoint of the transition in the equilibrium Mg^{2+} concentration dependence was 1.4 mM, close to that found for P3 formation (0.97 mM, see above) (Fig. 18B). P3 and P7 are thus indistinguishable both kinetically and at equilibrium. Although these experiments so far could not yet determine whether they form cooperatively, sequentially, or independently, the P3 and P7 helices

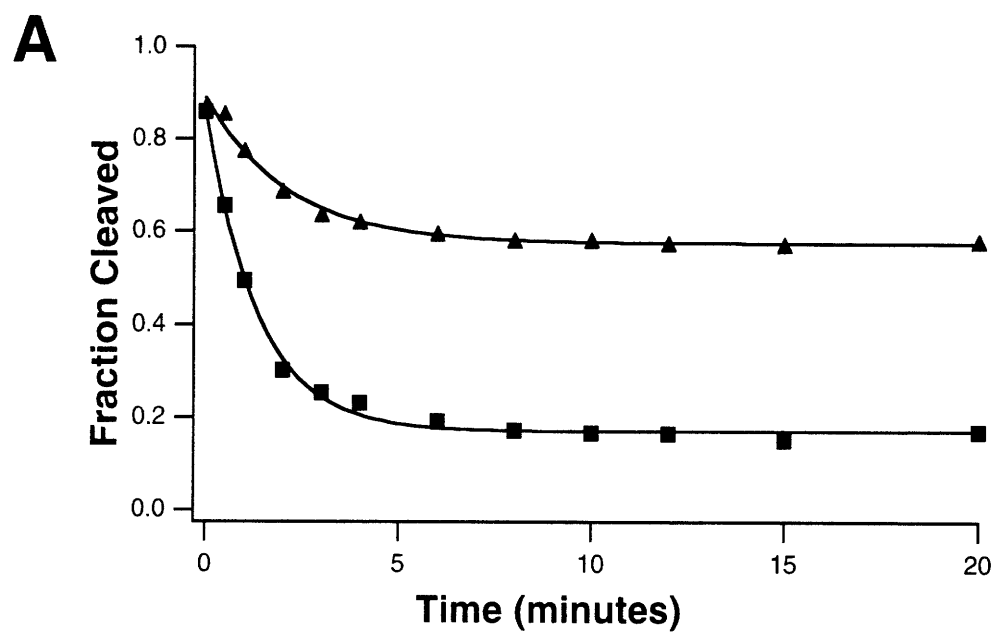


Figure 18. Formation of P7. **(A)** Kinetics of P7 formation at 10 mM (■) and 1 mM (▲) Mg²⁺. Final probe concentration (probe 19) was 60 μM.

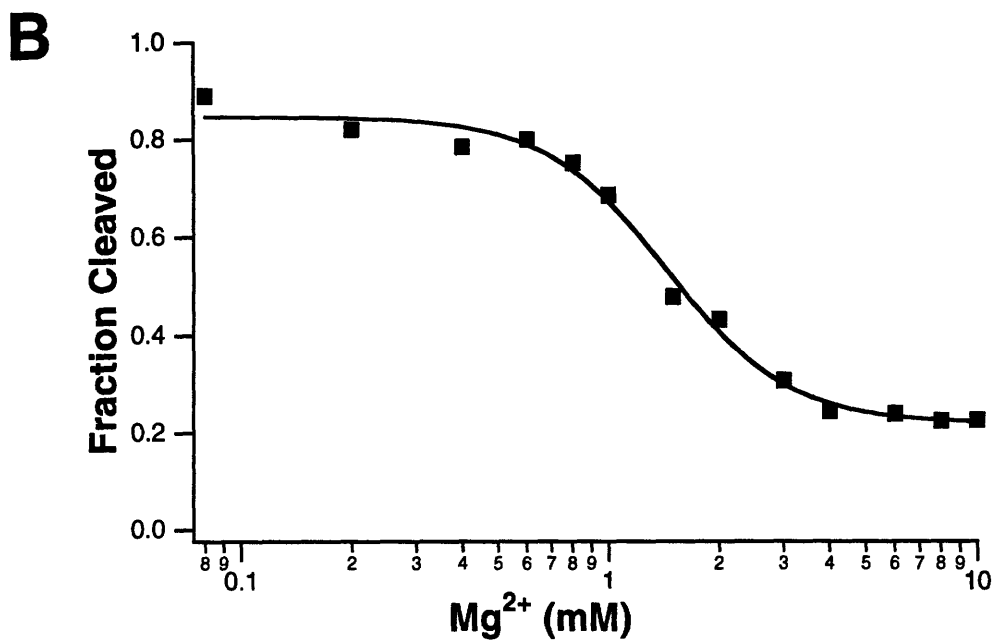


Figure 18. Formation of P7. **(B)** Equilibrium Mg²⁺ concentration dependence. The data were fit to an expression of simple two state binding of n Mg²⁺ ions. The fit gave values of 1.4 mM for the midpoint of the transition and 2.60 for n .

clearly become inaccessible to probe binding and RNase H cleavage on the same time scale and at similar Mg^{2+} concentrations.

Two further tests of the kinetic mechanism of rapid Mg^{2+} -binding following the rate limiting, Mg^{2+} -independent step were performed with probes targeting P3 and P7. First, unfolding was monitored after dilution of the Mg^{2+} -equilibrated sample, and both reverse steps were fast. Second, in a double-jump experiment, re-folding was initiated immediately after unfolding by adding back Mg^{2+} , and was found to occur at the same rate as Mg^{2+} -initiated folding. These two observations further support the proposed two-step mechanism of P3-P7 formation.

Folding of P4-P6

A second helical subdomain in the model of the three-dimensional structure of the intron is formed by the P4, P5, and P6 stems (see Fig. 5B, C) (Michel and Westhof, 1990). This subdomain has been shown to be independently stable (Murphy and Cech, 1993; Murphy and Cech, 1994). When oligonucleotides complementary to sequences in P4 (probe 8) and P6a (probe 12) were added to the RNA simultaneously with Mg^{2+} , no more than 40 to 50 percent cleavage was observed. These probes therefore are members of the third class (Fig. 10). If the limited cleavage were due to similar rates for folding and probe binding, it should have been possible to increase the extent of cleavage by increasing the probe concentration. At higher concentrations of either probe, however, the extent of cleavage remained unchanged. Alternatively, the population of unfolded molecules may be heterogeneous, with only a subpopulation able to bind oligonucleotides.

When either of the probes for P4 or P6a was added to the RNA before Mg^{2+} , cleavage increased to almost 80 percent. In a kinetic experiment a slow folding phase appeared to be preceded by a rapid burst (Fig. 19), providing evidence that an equilibrium between at least two conformations exists in the absence of Mg^{2+} . One of these is

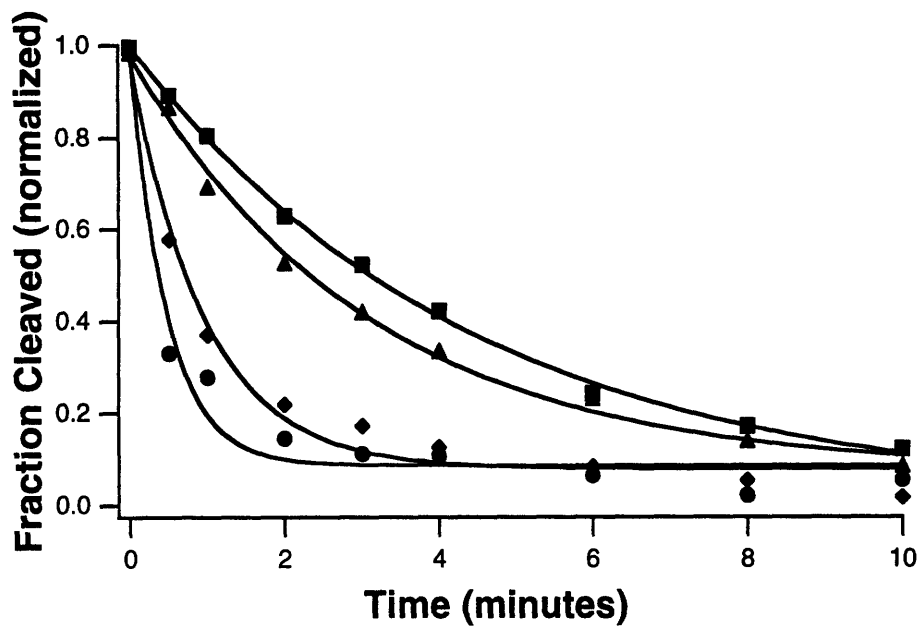


Figure 19. Folding kinetics of P3, P7, P4 and P6a at 27°C. The data for each probe was normalized to allow a direct comparison, and fit to a single exponential. Probes and concentrations used: (●) 12 (60 μM), (◆) 8 (60 μM), (▲) 19 (60 μM), (■) 15 (20 μM). For probes 8 and 12 the initial burst described in the text is reflected by the imperfect fit of the data to a single exponential.

inaccessible to binding by probes **8** and **12** or RNase H cleavage and can fold rapidly once Mg^{2+} is added, while the second conformation is accessible to the probes and must go through a slow step in order to fold. In the absence of bound Mg^{2+} , hybridization of oligonucleotide can drive the equilibrium toward the accessible conformation to give increased cleavage. An attempt to improve the resolution of the kinetic phases by folding the RNA at a lower temperature showed that the kinetic behavior of P4 and P6a at 37°C and 27°C was indistinguishable by kinetic oligonucleotide hybridization. The apparent temperature independence of P4 and P6a formation contrasts with P3 and P7, whose rates of formation were very sensitive to temperature (activation energy ~ 20 kcal/mol; this value was obtained from an arrhenius plot of the observed rates of formation of P3 and P7, measured using probes **15** and **19**, respectively, at five temperatures between 25°C and 37°C) (compare Figs. 15, 18 and Fig. 19). Thus, these two helical subdomains of the *Tetrahymena* intron clearly show different kinetic behavior during Mg^{2+} -induced folding, with the P4-P6 subdomain apparently forming faster than P3-P7.

A model for the kinetic folding mechanism

These results support a model for the folding pathway of the *Tetrahymena* ribozyme that includes several intermediates and involves the sequential formation and stabilization of helical subdomains (Fig. 20). The first intermediate, I_0 , consists of short-range secondary structures, such as hairpins, that form in the absence of Mg^{2+} . Experiments with probes targeting P4 and P6a show that I_0 is in a slow equilibrium with another conformation (I_1). At present, it is not possible to distinguish whether I_1 is a productive intermediate, or is on a nonproductive branch of the folding pathway. Rapid Mg^{2+} binding by either I_0 or I_1 leads to formation of I_2 and is reflected by the rapid burst observed in the kinetic experiments probing P4 and P6a. I_2 corresponds to an intermediate in which P4 and P6 are inaccessible to the probes, while P3 and P7 are still accessible. I_3 is the transient intermediate inferred

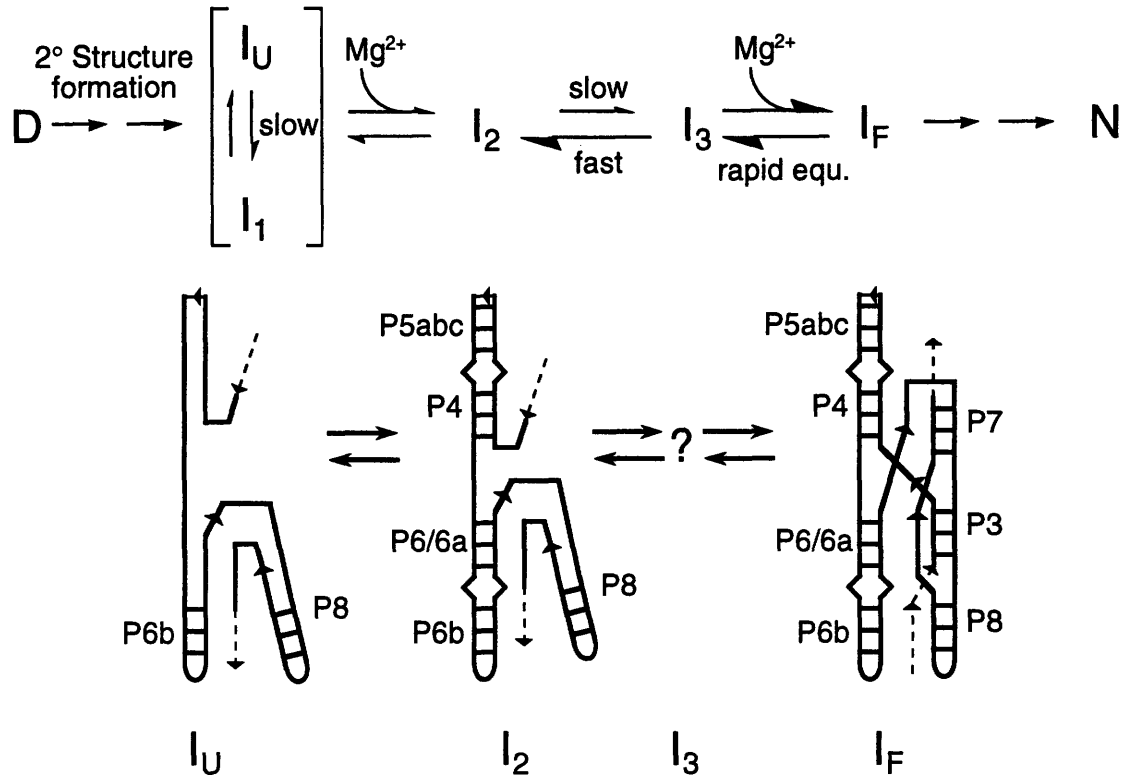


Figure 20. Model for the folding pathway of the *Tetrahymena* ribozyme.

from the Mg^{2+} concentration independence of the rate of P3 and P7 formation (see Figs. 15, 18). Only the presence of such an intermediate in the mechanism of P3-P7 formation can explain the requirement for Mg^{2+} , since the rate limiting step itself does not involve metal binding. The species I_3 may be one in which P3 and P7 are formed, but are not yet stable or in the correct orientation relative to P4-P6. Binding of Mg^{2+} to I_3 may allow the two helical stacks to come together in a conformation in which P3 and P7, as well as P4 and P6, are protected (I_f). Additional fast rearrangements may lead to the final structure (N), but since the rate at which full activity is gained is the same as the rate at which P3 and P7 become inaccessible, formation of the P3-P7 subdomain appears to be the overall rate limiting step on the folding pathway.

Implications

The described folding mechanism (Fig. 20) has three main implications. First, Mg^{2+} binding to I_3 must in fact be significantly stronger than suggested by the transition midpoint ($[Mg^{2+}]_{1/2}$) observed in the experiments with probes targeting P3 and P7, as well as in previous studies of the activity and structure of the *Tetrahymena* ribozyme (Celander and Cech, 1991; Wang and Cech, 1994). To overcome the unfavorable equilibrium between I_2 and I_3 , high concentrations of Mg^{2+} are required, even though I_3 binds Mg^{2+} tightly. With probes targeting P3 or P7 more than 90 percent cleavage is observed if they are added together with Mg^{2+} , suggesting that the equilibrium between I_2 and I_3 is at least tenfold in favor of I_2 , and thus that the true K_d of magnesium is at least ten times lower than the observed midpoint of the transition ($[Mg^{2+}]_{1/2}$) of about 1 mM. From the experimental data, an approximate upper limit of 100 μ M for the midpoint of the transition of Mg^{2+} binding to I_3 can therefore be estimated. This value is consistent with the magnitude of dissociation constants measured for high affinity Mg^{2+} binding sites in tRNA (Schimmel and Redfield, 1980; Pan *et al.*, 1993). Second, the folding of this RNA involves several

discrete steps, reflected in the inferred kinetic intermediates. Third, and most fundamental, the folding pathway includes a slow step taking place on the time scale of minutes, and this rate limiting step itself does not involve binding of Mg^{2+} .

Comparison to previous studies

These results support and expand upon previous investigations of RNA folding. Studies on tRNA by Crothers and others have demonstrated the existence of folding intermediates (Cole and Crothers, 1972; Cole *et al.*, 1972; Crothers *et al.*, 1974; Stein and Crothers, 1976a; Stein and Crothers, 1976b). Lynch and Schimmel, in a fluorescence study (Lynch and Schimmel, 1974), have shown that the rate limiting step for Mg^{2+} -induced folding of tRNA, as with the *Tetrahymena* intron, does not involve Mg^{2+} binding. In contrast to the findings for the ribozyme, however, in tRNA Mg^{2+} binding precedes the slow, Mg^{2+} independent, steps. For both group I introns and tRNA, secondary structure generally precedes formation of tertiary interactions, which can be quite slow (Lynch and Schimmel, 1974; Banerjee *et al.*, 1993; Jaeger *et al.*, 1993; Banerjee and Turner, 1995). The rate of helix formation in nucleic acids is known to be extremely rapid ($k = 10^7$ - 10^9 $M^{-1}min^{-1}$) (Scheffler *et al.*, 1968; Pörschke and Eigen, 1971). Short range helices, and thus the formation of I_{ij} from the denatured RNA, are expected to form at such fast rates.

The presence of a slow step on the folding pathway of the *Tetrahymena* ribozyme is consistent with the reported observation that prior incubation in Mg^{2+} is necessary to obtain fully active ribozyme (Herschlag and Cech, 1990a). Studies on the thermodynamics of Mg^{2+} -induced folding of this RNA (Celander and Cech, 1991; Lagerbauer *et al.*, 1994) have provided evidence for two equilibrium folding intermediates; one in which a substructure involving P5abc is formed, and one in which only the P4-P6 subdomain, but not P3-P7, is formed. These observations, together with the demonstration that the P4-P6 subdomain is stable independently, but only in the presence of Mg^{2+} (Murphy and Cech,

1993; Murphy and Cech, 1994), support the plausibility of the proposed kinetic intermediate I_2 . In these studies it was shown that the P5abc substructure facilitates the formation of the P4-P6 subdomain, and that the P4-P6 subdomain stabilizes the catalytic core, including P3-P7. It can therefore be inferred that the P5abc intermediate should precede I_2 in the kinetic mechanism. Since, however, only those folding events can be monitored by the kinetic hybridization assay that are accompanied by a change in accessibility to the oligonucleotide probes, and formation of the P5abc intermediate does not appear to cause such a change, it cannot be definitively placed on the kinetic folding pathway. The localization of Mg^{2+} interactions to sites between the two helical stacks formed by P4-P6 and P3-P7 (Christian and Yarus, 1993; Murphy and Cech, 1993; Murphy and Cech, 1994) may explain the Mg^{2+} requirement for association of the subdomains. The general mechanism of helix formation followed by docking into the correct orientation follows a precedent set by kinetic and thermodynamic studies of an oligonucleotide substrate binding to the intron, where hybridization to the internal guide sequence precedes movement of the resulting P1 helix into the active site (Bevilacqua *et al.*, 1992; Herschlag, 1992; Wang *et al.*, 1993).

A framework for RNA folding

This model provides a conceptual framework for studying kinetic RNA folding intermediates. The active structures of large RNA molecules appear to consist of stacks of helices forming subdomains and to be stabilized by intersubdomain tertiary contacts. It has been suggested that these structures are formed in a hierarchical process, with individual helices forming first and providing nucleation sites for formation of the stacked subdomains, which then come together in the final structure (Westhof and Michel, 1992; Cech, 1993; Murphy and Cech, 1993). The experiments described above provide direct kinetic evidence for such a hierarchical folding pathway in the *Tetrahymena* ribozyme. The

observed intermediates suggest that subdomain folding generally occurs in two steps. Base-pairing interactions form first, followed by formation of tertiary contacts and stabilization through specific binding of Mg^{2+} .

In summary, the studies described so far allowed the presentation of a model for the folding pathway of the *Tetrahymena* ribozyme in which folding proceeds via several intermediates, with the P4-P6 subdomain being formed before, and potentially required for, formation of the P3-P7 subdomain. The rate limiting step is formation of P3-P7, and in itself does not involve binding of Mg^{2+} . Although only a part of the overall folding process could be examined, it is clear that folding is not a simple two-state transition, but takes place via a complicated mechanism involving both Mg^{2+} -dependent and Mg^{2+} -independent steps. The rapid transcription rates of RNA polymerases (von Hippel *et al.*, 1984) would result in the synthesis of an RNA the size of a group I intron in a matter of seconds. This may affect the formation of higher order structure (Emerick and Woodson, 1993; Lewin *et al.*, 1995), which can be much slower, and probably takes place post-transcriptionally rather than during transcription.

4. PROBING THE DETAILS OF THE FOLDING MECHANISM BY SITE-DIRECTED MUTAGENESIS

The model for the folding mechanism of the *Tetrahymena* ribozyme presented in chapter 3 documents the basic order of subdomain assembly, but does not reveal the details of the molecular rearrangements represented by the observed kinetic folding steps. To further refine the model, site directed mutagenesis was used to address three important issues in the folding mechanism. First, does the P3-P7 subdomain in fact constitute a cooperative kinetic folding unit? Second, what is the nature of the rate limiting step? Third, is formation of the two subdomains in fact hierarchical? The results presented in this chapter show that the P3 and P7 helices form in an interdependent manner, that the slow step in the minimal kinetic folding pathway involves both the formation of a triple helical scaffold that orients the two subdomains with respect to each other, and the stabilization of the P3-P7 subdomain by a nonconserved 3' terminal extension of the intron. The integrity of the P4-P6 subdomain, furthermore, is required for proper formation of P3-P7, while disruptions in P3-P7 have little effect on P4-P6 formation.

Interdependence of P3 and P7

In the studies described above, the P3 and P7 helices became inaccessible to oligonucleotide probes at the same rate and at similar Mg^{2+} concentrations, suggesting that the P3-P7 subdomain folds as a unit. It was not possible, however, to determine whether this subdomain forms cooperatively, or whether individual elements can fold independently. To test for possible interdependence of formation of the P3 and P7 helices, equilibrium and kinetic folding was examined for ribozymes containing mutations

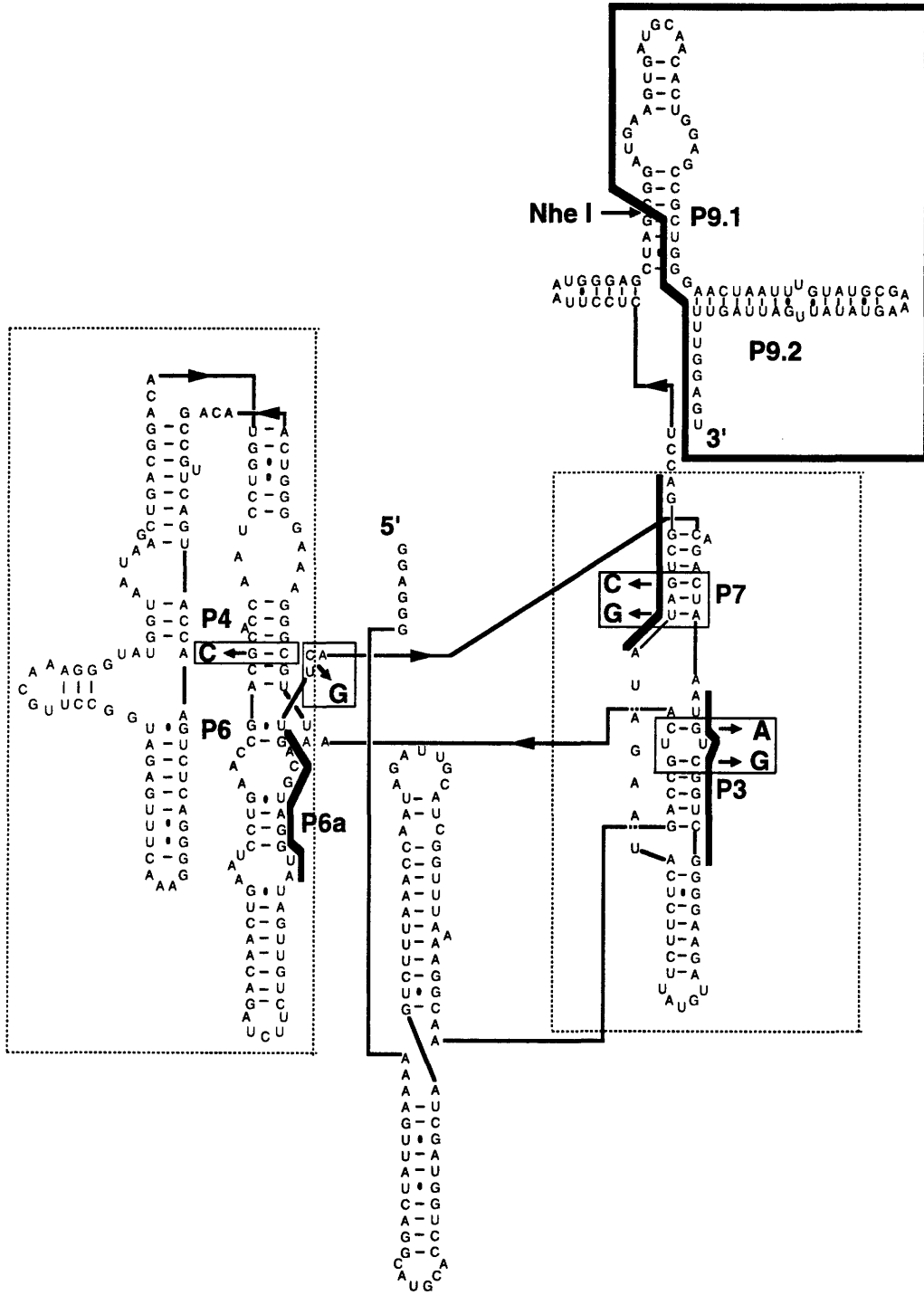


Figure 21. Mutations in the *Tetrahymena* ribozyme designed to reveal details of the folding pathway. The probes were adjusted to maintain complementarity to the targeted sequences. Substitutions present in any one mutant RNA are boxed in. P3 mutant: G272A/C274G; P7 mutant: U307G/G309C; triple scaffold mutant: C260G; P4 mutant: G212C. The site of truncation in the Nhe I deletion mutant is indicated, along with the P9.1-P9.2 extension. The two structural subdomains are outlined.

disrupting two base pairs in either the P3 or the P7 helix (Fig. 21). The particular mutations were chosen because their effect has previously been studied in the context of a self-splicing version of the intron to show the importance of the P3 and P7 stems for catalytic activity (Burke *et al.*, 1986; Williamson *et al.*, 1987). These studies showed that the mutant introns, while inactive at 2 mM Mg²⁺ (where the wild type is active), still retained catalytic activity at 10 mM Mg²⁺, suggesting that the mutations do not result in a global disruption of the structure of the RNA at the higher Mg²⁺ concentration.

P3 mutant

In the P3 mutant (G272A/C274G), both the extent and the rate of formation of the P3 helix were found to be indistinguishable from wild type at saturating (10 mM) Mg²⁺ (Fig. 22A). Examination of the Mg²⁺ concentration dependence of folding at equilibrium, however, revealed an increased requirement for Mg²⁺ to form the P3 helix in the P3 mutant compared to wild type (Fig. 22B). This observation shows that the P3 helix is destabilized by these mutations, and is consistent with the reported increased Mg²⁺ requirement for catalytic activity in this mutant (Williamson *et al.*, 1987). When formation of the P7 helix was monitored in the P3 mutant at saturating (10 mM) Mg²⁺, both the rate and extent of folding were again similar to wild type (Fig. 22C). Following formation of the P7 helix in the P3 mutant at low Mg²⁺ concentration showed an increased requirement for Mg²⁺. At equilibrium at 3.5 mM Mg²⁺, P7 remained accessible in about half the RNA molecules (Fig. 22C), suggesting that the apparent midpoint of the equilibrium transition is very close to that for the formation of the P3 helix in this mutant (compare the plateau levels at 10, 3.5 and 2 mM Mg²⁺ in Fig. 22C with the equilibrium transition for the P3 mutant in Fig. 22B). Therefore, disruption of base-pairs in the P3 helix affects formation of the P3 and P7 helices equally, and P7 appears unable to form in the absence of P3.

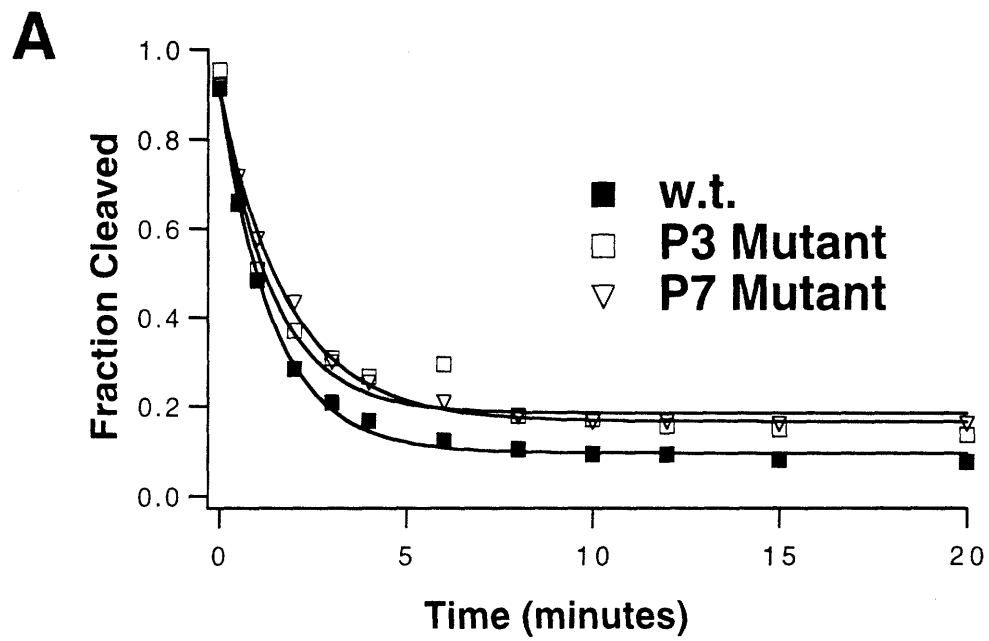


Figure 22. Interdependence of the P3 and P7 helices. **(A)** Kinetics of P3 helix formation at 10 mM Mg^{2+} .

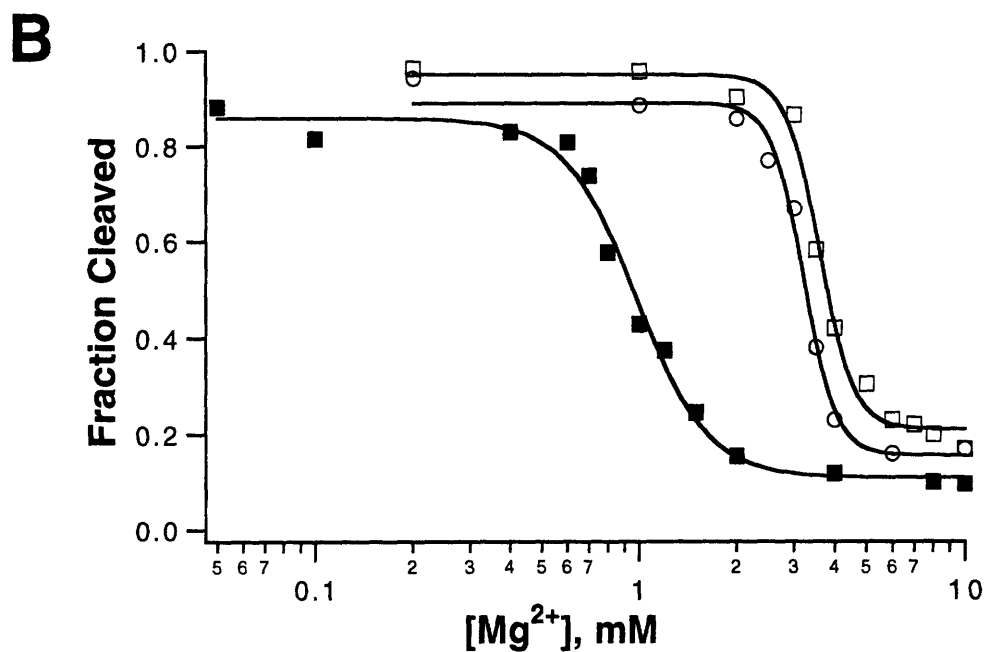


Figure 22. Interdependence of the P3 and P7 helices. **(B)** P3 helix formation at equilibrium in the wild type (■), P3 mutant (□) and P7 mutant (○). Each point represents the endpoint of a kinetic experiment performed at a different Mg^{2+} concentration. Values for $[Mg^{2+}]_{1/2}$ and n were 0.97 mM and 4.0 for wild type, 3.6 mM and 8.1 for the P3 mutant and 3.2 mM and 8.8 for the P7 mutant, respectively.

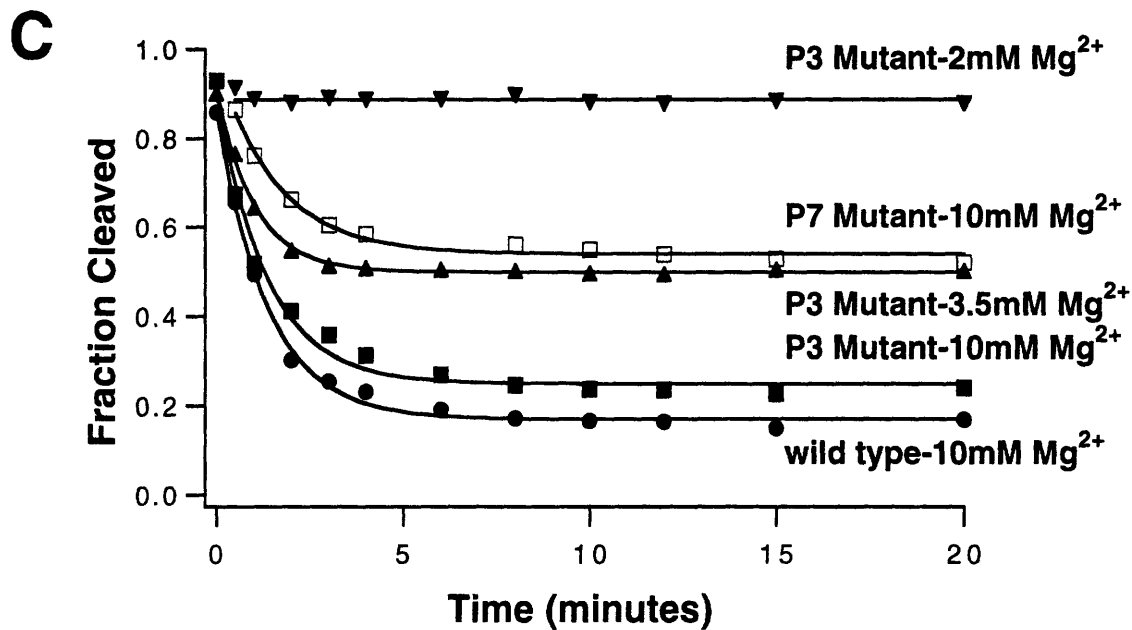


Figure 22. Interdependence of the P3 and P7 helices. **(C)** Kinetics of P7 helix formation in the wild type, P3 mutant and P7 mutant at different concentrations of Mg²⁺ at 37°C. In the P3 mutant at 3.5 mM Mg²⁺, approximately half of the RNA is inaccessible to the P7 probe at equilibrium. At 2 mM Mg²⁺, where in the wild type the P7 helix is largely inaccessible at equilibrium (see Figure 18B), it is completely accessible in the P3 mutant. Concentrations used for the P3 probe were: 20 μM (wild type, P3 mutant), 30 μM (P7 mutant); for the P7 probe: 60 μM (P3 mutant), 10 μM (P7 mutant).

P7 mutant

A more severe effect was observed in the context of the P7 mutant (U307G/G309C). Even at 10 mM Mg^{2+} , the P7 helix was partially accessible to the oligonucleotide probe, suggesting it was unable to form completely or stably (Fig. 22C). Higher Mg^{2+} concentrations did not increase the extent of protection, indicating that the partial accessibility was not due to a simple shift in the transition midpoint. Formation of the P7 helix is therefore impaired in the P7 mutant, consistent with the decreased catalytic activity under saturating conditions (see below). Since the mutant is catalytically active, it was concluded that the inaccessible fraction represents properly folded molecules. The results do not, however, address the nature of the accessible, presumably misfolded, fraction. There may be an equilibrium between the two conformations, but the rate of interconversion must be slow compared to the timescale of RNase H cleavage, or cleavage would be driven to completion. When P3 helix formation was probed in the P7 mutant, the equilibrium midpoint of the Mg^{2+} -dependent transition was shifted to a higher concentration compared to wild type (Fig. 22B), while at saturating Mg^{2+} concentrations both the rate and extent of folding were unchanged from wild type (Fig. 22A). Therefore, the mutations in P7 also destabilize the P3 helix, although not to the same extent as the P7 helix itself.

Together, these results show that disruption of either the P3 or P7 helix has an effect on both, and that P3 and P7 therefore are interdependent. Their formation is not truly cooperative, however, since, in the P7 mutant, the P3 helix appears to be stable under conditions where the P7 helix is partially accessible to oligonucleotide probes.

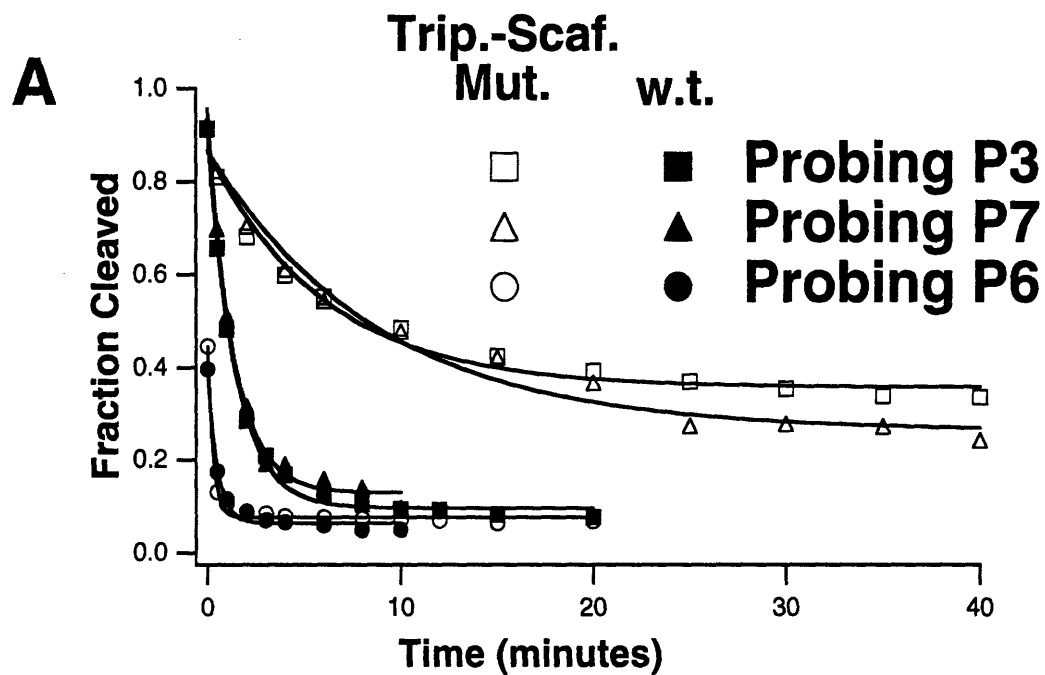
Nevertheless, it is apparent that the P3 and P7 helices do indeed constitute a kinetic folding unit of interdependent structural elements, and it is proposed that this folding unit includes the entire P3-P7 subdomain (see Fig. 21).

Importance of the triple helical scaffold

From earlier results it was known that the rate limiting folding step precedes stable formation of the P3-P7 subdomain (see chapter 3). The formation of tertiary interactions was a likely possibility for the molecular event corresponding to this slow step. A triple helical scaffold has been shown to form between base pairs in the P4 and P6 helices and adjacent single stranded regions (see Fig. 5C) (Michel *et al.*, 1990). This scaffold directly connects the two subdomains in the primary sequence of the RNA and is located at the subdomain interface in the model of the three dimensional structure (see Fig. 5B) (Michel and Westhof, 1990). To test whether formation of the triple-scaffold was involved in the slow folding step, one of the bases proposed to interact with a base pair in P4 was mutated from a pyrimidine to a purine (triple-scaffold mutant, C260G) (Fig. 21). Examination of group I intron phylogeny (Michel and Westhof, 1990) and the results of *in vitro* selection experiments (Green and Szostak, 1994) showed that if base pair three in P4 is C•G (C109•G212 in the *Tetrahymena* ribozyme), the nucleotide which is proposed to form a triple interaction with this base pair is never G (C260 in the *Tetrahymena* ribozyme). The C260G mutation is therefore expected to disrupt at least one triple interaction in the scaffold, and may destabilize adjacent triples as well (Michel *et al.*, 1990; Michel and Westhof, 1990; Chastain and Tinoco, 1993; Green and Szostak, 1994).

Decreased folding rate

When the kinetics of formation of the P3 helix were examined in the triple-scaffold mutant, the observed folding rate constant ($k_{\text{obs.}} = 0.15 \pm 0.02 \text{ min}^{-1}$, where the error encompasses the range of rates measured in three independent experiments) was decreased almost fivefold from the wild type ($k_{\text{obs.}} = 0.72 \pm 0.14 \text{ min}^{-1}$, where the error is the standard deviation for 17 independent experiments, as shown in chapter 3) (Fig. 23A). The same decrease in the rate constant was observed when formation of the P7 helix was



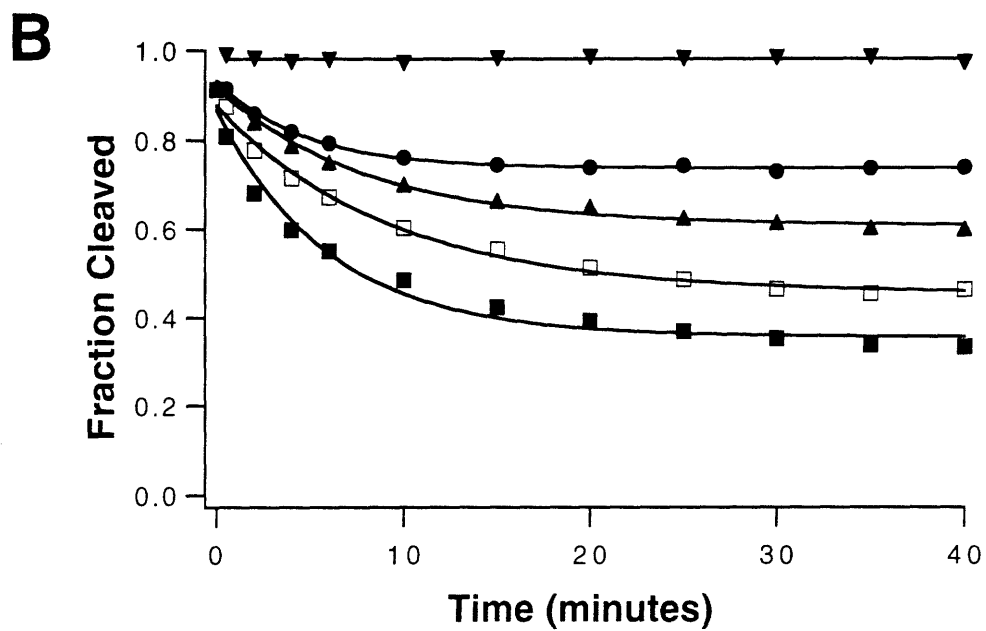


Figure 23. Slow folding of the triple-scaffold mutant compared to wild type.
(B) Kinetics of P3 formation at 10 mM (■), 4 mM (□), 3 mM (▲), 2.5 mM (●) and 1 mM (▼) Mg²⁺.

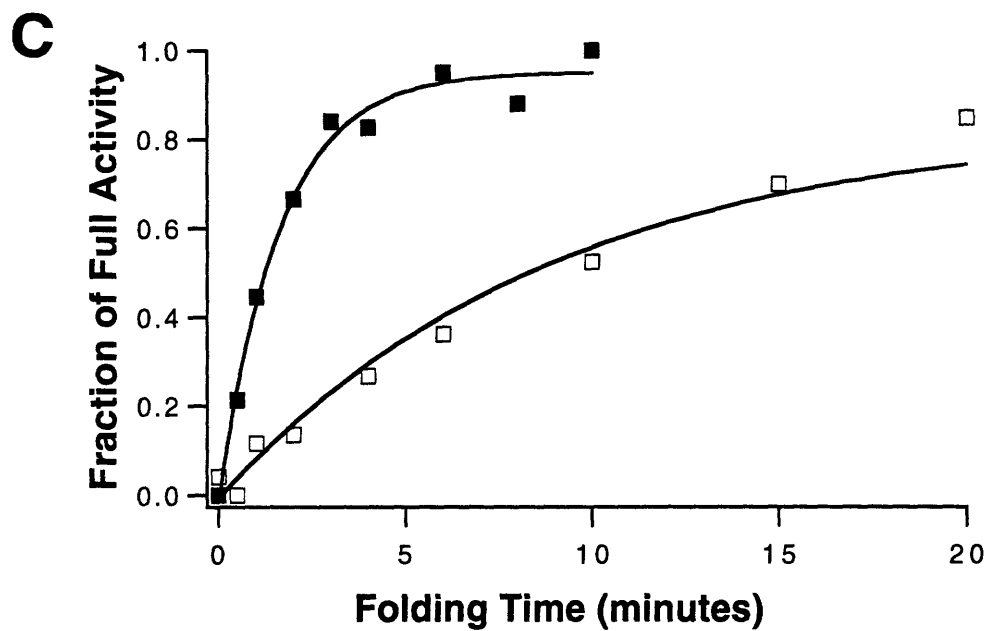


Figure 23. Slow folding of the triple-scaffold mutant compared to wild type. **(C)** Kinetics of active structure formation in the wild type (■) and the triple-scaffold mutant (□) measured by the ribozyme activity assay at 37°C. The data were fit to a single exponential (wild type: $k_{\text{obs.}} = 0.61 \text{ min}^{-1}$; triple-scaffold mutant: $k_{\text{obs.}} = 0.11 \text{ min}^{-1}$) and normalized to allow a direct comparison.

probed (Fig. 23A), further confirming the interdependence of the P3 and P7 helices. The rate decrease may result either from an effect on the slow step itself, or from an effect on a different step that leads to a change in the rate limiting step. To test whether a rearrangement preceding the formation of the P3-P7 subdomain had become rate limiting, the kinetics of folding of the P4-P6 subdomain were examined. No change was observed compared to the wild type (Fig. 23A), suggesting that P3-P7 formation remained rate limiting. To test whether the Mg^{2+} binding step following the slow step had become rate limiting the rate of P3-P7 formation was measured at different Mg^{2+} concentrations. The apparent rate constant was found to be independent of the Mg^{2+} concentration, within the error of the experiments, showing that Mg^{2+} binding had also not become rate limiting (Fig. 23B). Therefore, within the previously proposed minimal kinetic scheme, the rate limiting step has not changed in this mutant. The decrease in the rate of P3-P7 subdomain formation in the triple-scaffold mutant is thus proposed to be due to a direct effect on the slow step, which apparently includes formation of the triple helical scaffold.

Slow gain of catalytic activity

The rate at which full catalytic activity is gained was measured to verify that the slower folding rate of the triple-scaffold mutant is functionally significant. The fraction of activity present after various folding times was compared to the activity of fully folded ribozyme. Full activity was obtained with a rate constant ($k_{obs.} = 0.11 \text{ min}^{-1}$) (Fig. 23C) which is nearly the same as the folding rate constant measured by the oligonucleotide hybridization assay using probes targeting the P3 or P7 helices (Fig. 23A), and more than fivefold slower than that measured for the wild type ribozyme ($k_{obs.} = 0.61 \text{ min}^{-1}$). This provides an independent confirmation of the decreased folding rate of the triple-scaffold mutant, and shows that there are no subsequent steps that are significantly slower than formation of the P3-P7 subdomain.

Catalytic activity of mutant ribozymes

The catalytic efficiency of all three mutant ribozymes was determined quantitatively at 10 mM Mg²⁺ to check for any extreme decrease in activity compared to the wild type RNA that might suggest severe structural perturbations. While the catalytic competency of the P3 and P7 mutants has been established previously (Burke *et al.*, 1986; Williamson *et al.*, 1987), these earlier studies employed a self-splicing version of the intron and were performed before a detailed kinetic framework for the endonuclease reaction was available (Herschlag and Cech, 1990a). To determine the relative catalytic activity of the wild type and mutant ribozymes the kinetic parameter k_c was measured, which is the rate of substrate cleavage under single turnover conditions with saturating ribozyme and GTP, corresponding to the rate of reaction of the ribozyme•substrate•GTP complex (McConnell *et al.*, 1993). This rate reflects steps after substrate binding and before product release, and as such should be sensitive to changes in the core structure of the RNA. To slow the rate of the chemical cleavage step, and facilitate measurement of k_c , assays were performed at decreased pH (pH 5.5). Values for k_c were very close for the wild type ($k_c = 0.20 \pm 0.02 \text{ min}^{-1}$) and the P3 mutant ($k_c = 0.17 \pm 0.01 \text{ min}^{-1}$). The P7 mutant had somewhat decreased activity ($k_c = 0.034 \pm 0.004 \text{ min}^{-1}$). No heterogeneity due to the possible presence of two conformations suggested by the oligonucleotide accessibility data (see above) was detected in this mutant. Since activity was determined in ribozyme excess, however, the measured rate could be due solely to the inaccessible conformation. While the lower activity does suggest some structural defects, a much greater effect would be expected if the core structure was substantially altered. The activity of the triple scaffold mutant was essentially identical to the wild type ($k_c = 0.23 \pm 0.02 \text{ min}^{-1}$), indicating that this mutation results in a true kinetic folding mutant in which the final, active structure of the RNA is not significantly perturbed, although the rate of formation of active enzyme is slower than for

the wild type (Fig. 23C). None of the three mutants thus showed a severe defect in the ability to catalyze substrate cleavage at 10 mM Mg^{2+} , suggesting that the overall structure of the RNA is likely largely preserved in these molecules.

Role of the P9.1-P9.2 peripheral extension

In the *Tetrahymena* intron, the 3' end is formed by several helices that constitute the P9.1-P9.2 peripheral extension (Fig. 21). A ribozyme truncated at a point corresponding to a Nhe I site in the plasmid template (L-21 Nhe I) lacks the P9.1-P9.2 extension but, once folded, has catalytic activity equal to that of the full length L-21 Sca I RNA (which is produced from a plasmid template linearized with Sca I) (Laggerbauer *et al.*, 1994). At high Mg^{2+} concentrations at equilibrium, deletion of the extension introduces no apparent structural perturbation in the remainder of the molecule (Laggerbauer *et al.*, 1994). However, an increased concentration of Mg^{2+} is required to form the P3-P7 subdomain compared to full length RNA. The P9.1-P9.2 extension therefore appears to specifically stabilize the P3-P7 subdomain (Laggerbauer *et al.*, 1994).

Slow folding of L-21 Nhe I RNA

To explore the role played by a structural element outside of the catalytic core during folding, and to determine whether the stabilization of P3-P7 by the extension is important for the folding process as well as for stability of the final structure, the folding kinetics of the L-21 Nhe I ribozyme, in which the extension is deleted, were examined using an oligonucleotide probe targeting the P3 helix (Fig. 21). Since the P3-P7 subdomain forms in an interdependent manner, the probe targeting P3 reports on formation of the entire subdomain (see above). The folding rate constant in the L-21 Nhe I RNA was decreased compared to full length L-21 Sca I ribozyme (Fig. 24A) ($k_{obs} = 0.32 \pm 0.04 \text{ min}^{-1}$ for L-21 Nhe I vs. $0.72 \pm 0.14 \text{ min}^{-1}$ for L-21 Sca I). Because the rearrangement

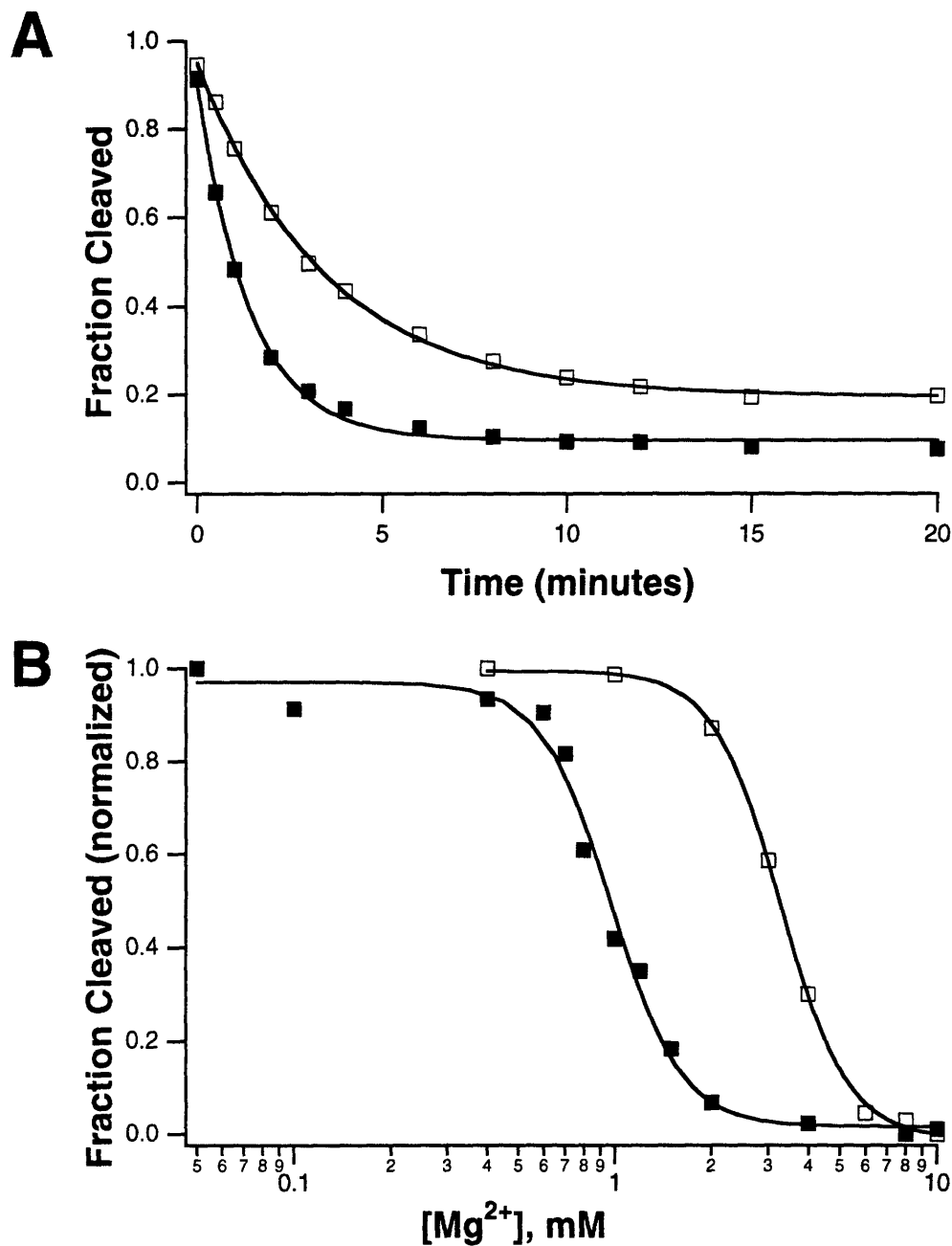


Figure 24. Formation of P3 in L-21 Sca I (■) and L-21 Nhe I (□) RNA.

(A) Kinetics of P3 formation at 10 mM Mg²⁺ and 37°C. Observed rate constants were $0.72 \pm 0.14 \text{ min}^{-1}$ for L-21 Sca I and $0.32 \pm 0.04 \text{ min}^{-1}$ for L-21 Nhe I. The error represents the standard deviation from 17 independent experiments for L-21 Sca I, as described in chapter 3, and the range of values obtained from three independent experiments for L-21 Nhe I. The probe concentration (probe 15) was 20 μM .

(B) Mg²⁺ concentration dependence of P3 formation at equilibrium. Transition midpoints are 0.97 mM for L-21 Sca I, as described (chapter 3), and 3.3 mM for L-21 Nhe I. The data were normalized to allow a direct comparison.

monitored with this probe is a slow step required for the stable formation of the P3-P7 subdomain (the formation of intermediate I_3 from I_2 , see Fig. 20), the decreased rate constant suggests that the stabilization of the P3-P7 subdomain by the P9.1-P9.2 extension may be involved in this slow step. Alternatively, deleting the 3' terminal extension may have resulted in a change of the rate limiting step.

To determine whether the Mg^{2+} -dependent formation of I_F from I_3 had become rate limiting (Fig. 20), the folding rate was measured with the P3 probe at different concentrations of Mg^{2+} . No change in the rate constant was observed, within experimental error, suggesting that Mg^{2+} binding had not become rate limiting. To determine whether an earlier step on the folding pathway had become rate limiting, the kinetics of formation of the P4-P6 subdomain and intermediate I_2 were monitored using a probe targeting P6/P6a (Fig. 21). The rate was found to be fast and unchanged from the L-21 Sca I RNA (Fig. 25), showing that formation of I_2 had also not become rate limiting. Together, these results indicate that within the minimal kinetic folding pathway the rate limiting step remained unchanged upon deletion of the P9.1-P9.2 extension.

Significance

The observed decrease in the folding rate constant upon deletion of the P9.1-P9.2 extension is only twofold, and therefore not very dramatic. The folding rate of the wild type RNA was measured numerous times with separate preparations of each reagent over a period of almost one year, and in none of these experiments was a rate constant as low as that observed for the L-21 Nhe I RNA ever measured. Conversely, the folding rate constant of the L-21 Nhe I ribozyme was independently determined several times using separate preparations of reagents, and the rate constant was never in the range observed for the wild type RNA. The effect is therefore highly reproducible, which strongly suggests that the twofold difference, although small, is real.

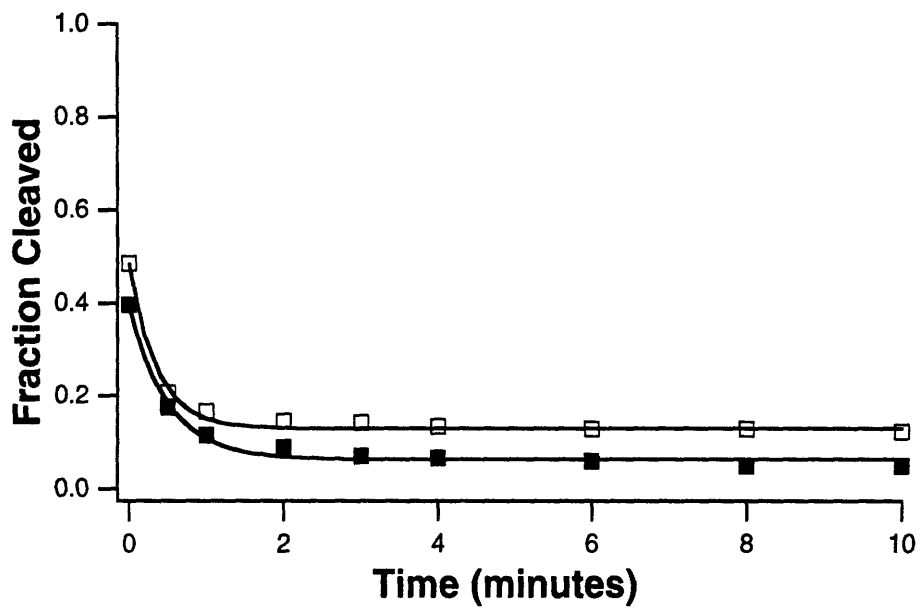


Figure 25. Kinetics of P4-P6 subdomain formation in L-21 Sca I (■) and L-21 Nhe I (□) RNA monitored by a probe targeting P6a (probe 12). The probe concentration was 60 μ M.

Equilibrium Mg²⁺-dependence

To compare the oligonucleotide hybridization data with the published equilibrium Mg²⁺-dependence of P3-P7 formation in the context of the L-21 Nhe I ribozyme (Laggerbauer *et al.*, 1994), the fraction of RNA in which P3 is accessible to the oligonucleotide probe was determined at equilibrium at a series of Mg²⁺ concentrations (Fig. 24B). The fraction of ribozyme folded at each Mg²⁺ concentration was obtained from the equilibrium endpoint of a kinetic experiment as described above. Each point therefore represents an independent experiment. The midpoint of the transition ($[Mg^{2+}]_{1/2} = 3.3$ mM) was shifted to a higher Mg²⁺ concentration compared to L-21 Sca I ($[Mg^{2+}]_{1/2} = 0.97$ mM, see chapter 3).

Hierarchical relationship between the two subdomains

None of the mutations described so far affected the kinetics of P4-P6 formation. When the equilibrium Mg²⁺ concentration dependence of P4-P6 formation was examined in the P3 and P7 mutants, only a moderate increase in the midpoint of the transition was observed (Fig. 26), consistent with published observations (Laggerbauer *et al.*, 1994). Mutations outside of P4-P6 that affect P3-P7 formation therefore have only a small effect on P4-P6 formation.

To determine the effect of changes in P4-P6, a mutation disrupting one base pair in P4 (G212C) was introduced into the ribozyme (Fig. 21). At 10 mM Mg²⁺ the P3-P7 subdomain was unable to form completely or stably in this mutant (identical results were obtained using probes targeting either P3 or P7) (Fig. 27). Higher Mg²⁺ concentrations again did not result in an increase in protection, suggesting the disruption of P3-P7 formation was not merely due to an increased requirement for Mg²⁺. When P4-P6 formation was examined in this mutant using a probe targeting P6/P6a, a change in the

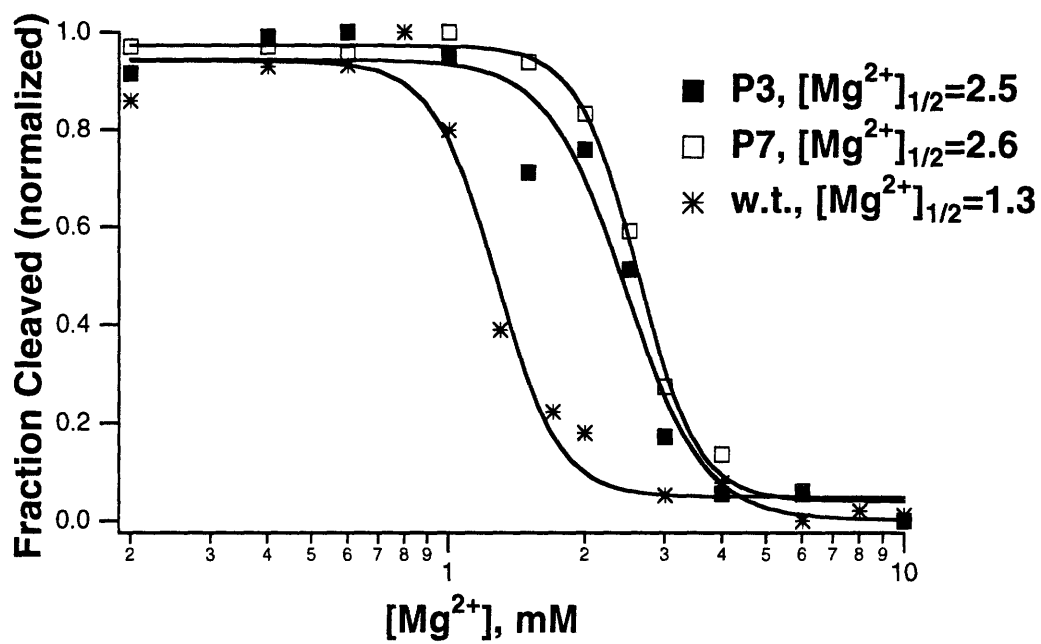


Figure 26. Effect of mutations in P3 and P7 on the equilibrium Mg^{2+} concentration dependence of P4-P6 formation monitored with a probe targeting P6a (probe 12). The probe concentration was 60 μM .

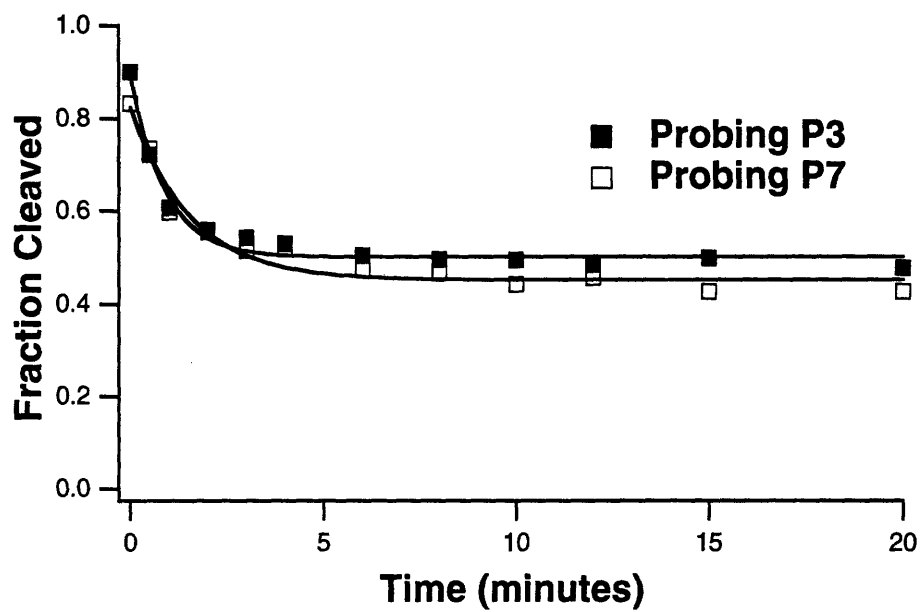


Figure 27. Formation of P3 and P7 in the P4 mutant (G212C) at 10 mM Mg^{2+} . Probe concentrations were 20 μM (probe 15) and 60 μM (probe 19).

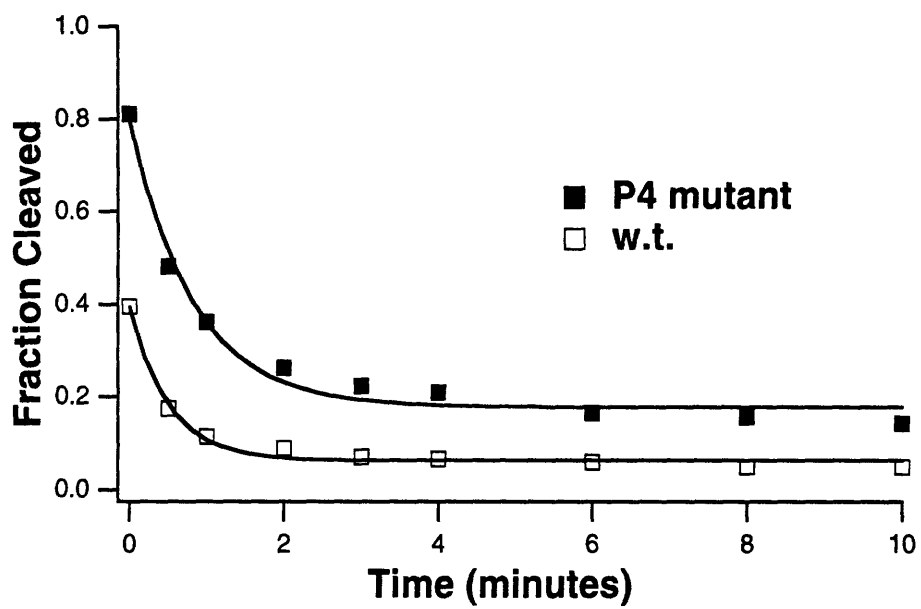


Figure 28. Formation of P4-P6 in the P4 mutant monitored with a probe targeting P6a (probe 12). The probe concentration was 60 μ M.

relative amplitudes of the fast and slow kinetic phases observed previously (see Fig. 19) was detected (Fig. 28). In the P4 mutant the fast phase had almost disappeared, suggesting a shift in the equilibrium between intermediates I_U and I_1 in favor of I_U . These results confirm the involvement of P4 during early folding events.

A refined model for the folding mechanism

Hierarchy and interdependence

The findings described in this chapter allow a refinement of the model for the kinetic folding mechanism presented in chapter 3. Together with other observations, the results confirm and emphasize the hierarchical nature of the folding pathway of the *Tetrahymena* ribozyme. There is no effect on the kinetic or equilibrium folding of the P4-P6 subdomain in the triple-scaffold mutant (Fig. 23A). In the P3 and P7 mutants, while the Mg^{2+} requirement for P4-P6 subdomain formation is slightly increased (Fig. 26), the kinetics of P4-P6 subdomain folding are unaltered compared to wild type. Mutations that affect the P3-P7 subdomain therefore have little or no effect on folding of the P4-P6 subdomain, while disruption of the P4-P6 subdomain significantly destabilizes the P3-P7 subdomain (Laggebauer *et al.*, 1994)(Fig. 27). The hierarchical formation of the two subdomains is thus a central feature of the folding pathway. The interdependence of the P3 and P7 helices also shows that the P3-P7 subdomain can form in a concerted fashion. This structural subdomain therefore corresponds to a kinetic folding unit, although the exact boundaries of this folding unit have not yet been determined. Equilibrium folding experiments have shown that the P4-P6 subdomain consists of two modules that form in a hierarchical manner (Celander and Cech, 1991; Laggebauer *et al.*, 1994; Murphy and Cech, 1994), and it is proposed that each of these modules may also represent a cooperative kinetic folding unit.

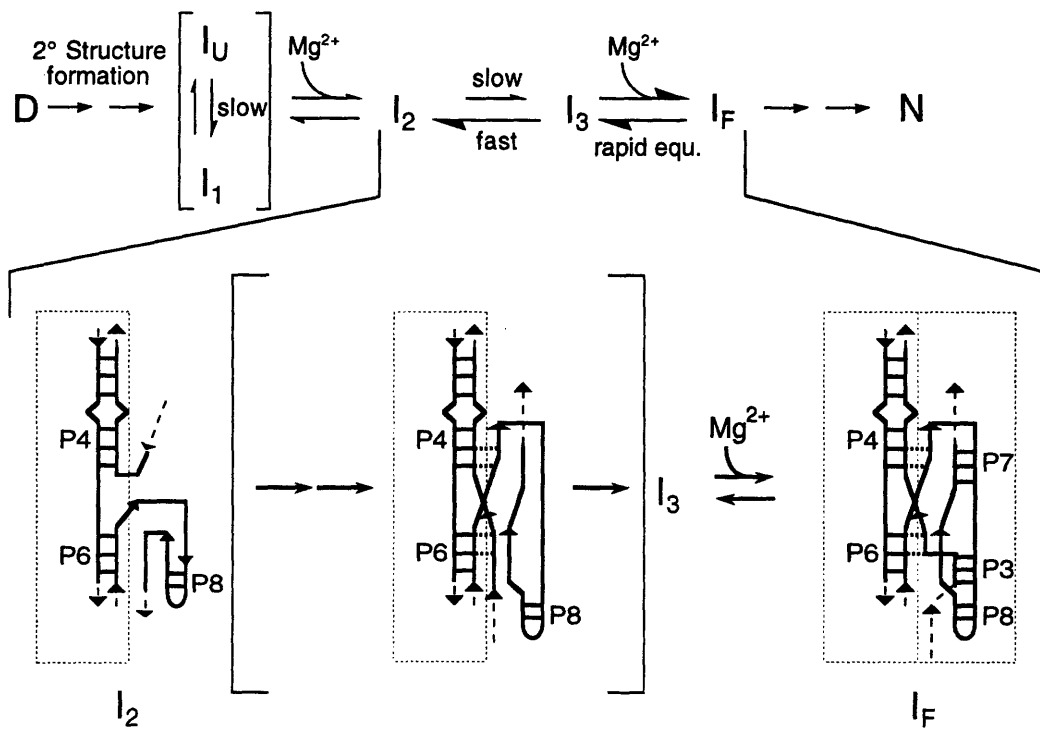


Figure 29. Refined model for the kinetic folding pathway of the *Tetrahymena* ribozyme. The two main structural subdomains are outlined.

Identity of I₃

The initial observations identified a transient intermediate during the formation of the P3-P7 subdomain, termed I₃, but did not provide any indication of its structural features (Fig. 20). In intermediate I₂, the P4-P6 subdomain but not the P3-P7 subdomain is folded, and formation of I₃ from I₂ was identified as the slow step in the minimal kinetic scheme. The effect of the triple-scaffold mutation on this kinetic step implies that while the scaffold is not yet present in I₂, it is formed in I₃ (Fig. 29). Formation of the triple helical scaffold promotes, and either precedes or is coincident with, the interaction of the two subdomains. Similarly, the decrease in the rate of the slow step upon deletion of the P9.1-P9.2 extension at the 3' end of the ribozyme (Fig. 24A) suggests that the interaction between P3-P7 and the extension takes place during the conversion of I₂ to I₃. In I₃, P3 and P7 are properly positioned to allow the rapid binding of Mg²⁺, resulting in the stabilization of the P3-P7 subdomain and formation of the active structure.

The slow folding step

The slow step in the proposed minimal mechanism (Fig. 29) may include several microscopic steps which cannot be resolved using the kinetic oligonucleotide hybridization assay. In addition to formation of the triple helical scaffold and the interaction between P9.1-P9.2 and L2/L2.1 (see below), base pairing in P3 and P7 and formation of other tertiary interactions may also occur during this step. At present it cannot be distinguished whether formation of the triple-scaffold and the stabilization of the P3-P7 subdomain by the P9.1-P9.2 extension are discrete steps, or whether they occur in a concerted rearrangement. A double mutant ribozyme, in which both the 3' terminal extension was deleted and the triple-scaffold was disrupted, was too unstable to allow careful analysis of folding kinetics (data not shown). It is also impossible to determine at this time which microscopic step occurring during the slow step in the kinetic mechanism is rate limiting.

Furthermore, it also cannot be ruled out that the mutations resulted in the stabilization of a non-native interaction, and that the rate limiting step is the slow disruption of such an interaction.

Role of the triple helical scaffold

The importance of the triple helical scaffold during folding provides further evidence that it orients the two subdomains with respect to each other to allow their proper association. A similar conclusion was drawn in a study of the requirements for formation of a functional complex where the two subdomains were combined as separate molecules (Doudna and Cech, 1995). In this system, association of the subdomains was weak in the absence of the triple helical scaffold, and the resulting complex was catalytically inactive. Inclusion of the scaffold in the complex strengthened the association and restored activity. The results presented here suggest that in the context of the whole intron disruption of the triple helical scaffold, in the case of the one mutant studied here, does not preclude the proper association of the subdomains, but merely slows their docking. The role of the triple helical scaffold in orienting the two subdomains was also suggested by an NMR study of model oligonucleotides (Chastain and Tinoco, 1993), *in vitro* selection experiments (Green and Szostak, 1994), and by the ability of the CYT-18 protein to suppress mutations in the scaffold (Mohr *et al.*, 1992).

The P9.1-P9.2 extension guides folding

Chemical modification experiments have provided evidence for a tertiary interaction between the 3' terminal extension and bases in the loops L2 or L2.1, which may help lock the P3-P7 subdomain in place (Banerjee *et al.*, 1993) and be responsible for the stabilization of P3-P7 (Laggerbauer *et al.*, 1994). In the *sunY* group I intron a similar stabilization of the core by elements of a 3' terminal domain was observed (Jaeger *et al.*, 1993). The results described above suggest that the proposed tertiary interaction between

the P9.1-P9.2 extension and L2 or L2.1 in the *Tetrahymena* intron is formed during the slow step of the minimal folding mechanism, and is therefore present in the transient intermediate I_3 . The identity of the nucleotides in P9.1-P9.2 and L2/L2.1 participating in the tertiary interaction has not yet been established, and it was therefore not possible to test the involvement of this interaction more specifically. It is proposed that the P9.1-P9.2 extension is important not only for the stabilization of the P3-P7 subdomain, but also helps guide folding of the RNA by limiting the mobility of P3 and P7 and, along with the triple helical scaffold, promoting the proper association of the P4-P6 and P3-P7 subdomains.

Determination of the Mg^{2+} dependence of formation of the P3 helix in the L-21 Nhe I RNA showed a shift of the transition midpoint to a higher Mg^{2+} concentration, and confirmed published observations (Banerjee *et al.*, 1993; Lagerbauer *et al.*, 1994) that the P3-P7 subdomain is destabilized in this truncated ribozyme. While the exact value of the transition midpoint ($[Mg^{2+}]_{1/2} = 3.3$ mM) differs somewhat from that observed using Fe(II)-EDTA as a footprinting probe (Lagerbauer *et al.*, 1994) ($[Mg^{2+}]_{1/2} = 1.83$ mM), the difference in the apparent midpoints is not surprising given the large difference in size between the two probes used (small hydroxyl radicals generated in the presence of Fe(II)-EDTA versus large oligonucleotides). The different temperatures at which the experiments were performed (42°C for the hydroxyl radical footprinting, 37°C for the kinetic oligonucleotide hybridization) may also contribute to this discrepancy. Although the formation of I_3 from I_2 itself does not involve binding of Mg^{2+} , the observed increase in the equilibrium Mg^{2+} requirement for P3-P7 subdomain formation can at least in part be accounted for by the decreased rate of formation of I_3 . Since rapid Mg^{2+} binding to I_3 drives the otherwise unfavorable conversion of I_2 to I_3 (Figs. 16, 29), a decrease in the rate of this conversion will result in an increase of the Mg^{2+} concentration required for formation of I_F , even if the dissociation constant of Mg^{2+} binding to I_3 is unaltered. The

kinetic results described here thus complement previous observations made at equilibrium (Banerjee *et al.*, 1993; Lagerbauer *et al.*, 1994).

In the cellular environment it may be that production of mature RNA, for transcripts containing group I introns, is limited by proper folding of the intron, rather than by catalysis. Partially or completely unfolded RNA may be bound rapidly and nonspecifically by proteins, which can inhibit formation of the active structure. Even a relatively minor increase in folding rate afforded by a peripheral extension may therefore confer a selective advantage by preventing the core of the RNA from becoming kinetically trapped in unproductive conformations or complexes.

The role of the P9.1-P9.2 extension during folding of the *Tetrahymena* ribozyme may have implications for the folding mechanisms of other large RNAs, many of which also consist of phylogenetically highly conserved core regions surrounded by less conserved peripheral extensions. Such an organization can be found in group I (Michel and Westhof, 1990; Cech, 1993) and group II introns (Michel *et al.*, 1989b), in the RNA component of RNase P (Darr *et al.*, 1992), and in ribosomal RNAs (Neefs *et al.*, 1993; De Rijk *et al.*, 1994). In several cases it has been demonstrated that peripheral extensions can stabilize the core region of a large RNA (Beaudry and Joyce, 1990; Jaeger *et al.*, 1991; Banerjee *et al.*, 1993; Jaeger *et al.*, 1993; Lagerbauer *et al.*, 1994), suggesting that some general functions of the extensions may be conserved between different RNAs. The importance of the P9.1-P9.2 extension during folding of the ribozyme suggests that, in addition to stabilizing the final conformation, nonconserved peripheral extensions may also guide the folding process in large, highly structured RNAs.

5. FOLDING OF RNASE P RNA

The studies described in this chapter were carried out in collaboration with Jing Wang, an undergraduate student in the Department of Molecular and Cellular Biology at Harvard University who performed his undergraduate thesis research in Prof. Jamie Williamson's laboratory.

RNase P as a second model system

The work described in the previous chapters provides considerable insight into the kinetic folding pathway of the *Tetrahymena* ribozyme. An unanswered question, however, was whether the folding properties of this group I intron are exceptional, or whether they illustrate features of the folding mechanism of large RNAs and even RNPs in general. To answer this question, folding of other RNAs and RNPs must be investigated, and compared to that of the *Tetrahymena* ribozyme. The RNA component of ribonuclease P (RNase P) is another well defined, relatively large, highly structured RNA and as such represents a second convenient model system in which to study kinetic RNA folding. A study of the folding mechanism of this RNA was therefore initiated.

RNase P is a ribonucleoprotein enzyme found in all organisms and plays a crucial role in the production of mature tRNA by cleaving nucleotides from the 5' end of pre-tRNA (Altman *et al.*, 1995). The RNA component of RNase P, like the *Tetrahymena* ribozyme, is a relatively large, highly structured molecule. While the RNA from many prokaryotes can catalyze pre-tRNA cleavage alone, in the absence of protein, the protein component of the enzyme enhances the reaction, and is required for activity in higher organisms (Darr *et*

al., 1992; Altman *et al.*, 1995). Like group I introns, RNase P RNAs from different species contain a core region of highly conserved interactions that is surrounded by additional structural elements that can vary substantially (Darr *et al.*, 1992; Haas *et al.*, 1994). The enzymes from *E. coli* and *B. subtilis* are among the best studied RNase P molecules, and for both the RNA alone is an efficient catalyst. They are of similar size as the *Tetrahymena* ribozyme (close to 400 nucleotides) and their secondary structure has been determined from phylogenetic comparison and biochemical studies (Darr *et al.*, 1992; Haas *et al.*, 1994)(Fig. 30). Models for the three dimensional architecture of the *E. coli* RNA have been proposed (Harris *et al.*, 1994; Westhof and Altman, 1994), and a detailed kinetic framework for the reaction of *B. subtilis* RNase P RNA with a pre-tRNA substrate is also available (Beebe and Fierke, 1994), providing a very powerful functional assay. Furthermore, an equilibrium study of the Mg^{2+} induced folding of the *B. subtilis* RNA has provided evidence for the existence of several subdomains of tertiary structure (Pan, 1995). This study also revealed that although optimal catalytic activity of RNase P RNA requires the presence of high concentrations of monovalent metal ions or Mg^{2+} (Reich *et al.*, 1988; Darr *et al.*, 1992; Brown *et al.*, 1993), stable higher order structure can be formed at relatively low concentrations of Mg^{2+} , even in the absence of additional monovalent metals.

As in the *Tetrahymena* ribozyme and most other large RNAs, at least two different classes of base pairing interactions are apparent in the secondary structure of RNase P RNA (Fig. 30). First, there are short range hairpin stem-loop interactions, which are expected to form very rapidly and at low salt concentrations, with no requirement for divalent metals. Second, there are long range interactions, such as the P2, P4 and P7 helices, where regions of the RNA far apart in the linear sequence fold back on each other. In analogy to the findings in the group I ribozyme, the long range interactions may require Mg^{2+} for their stable formation, and may form more slowly than short range interactions.

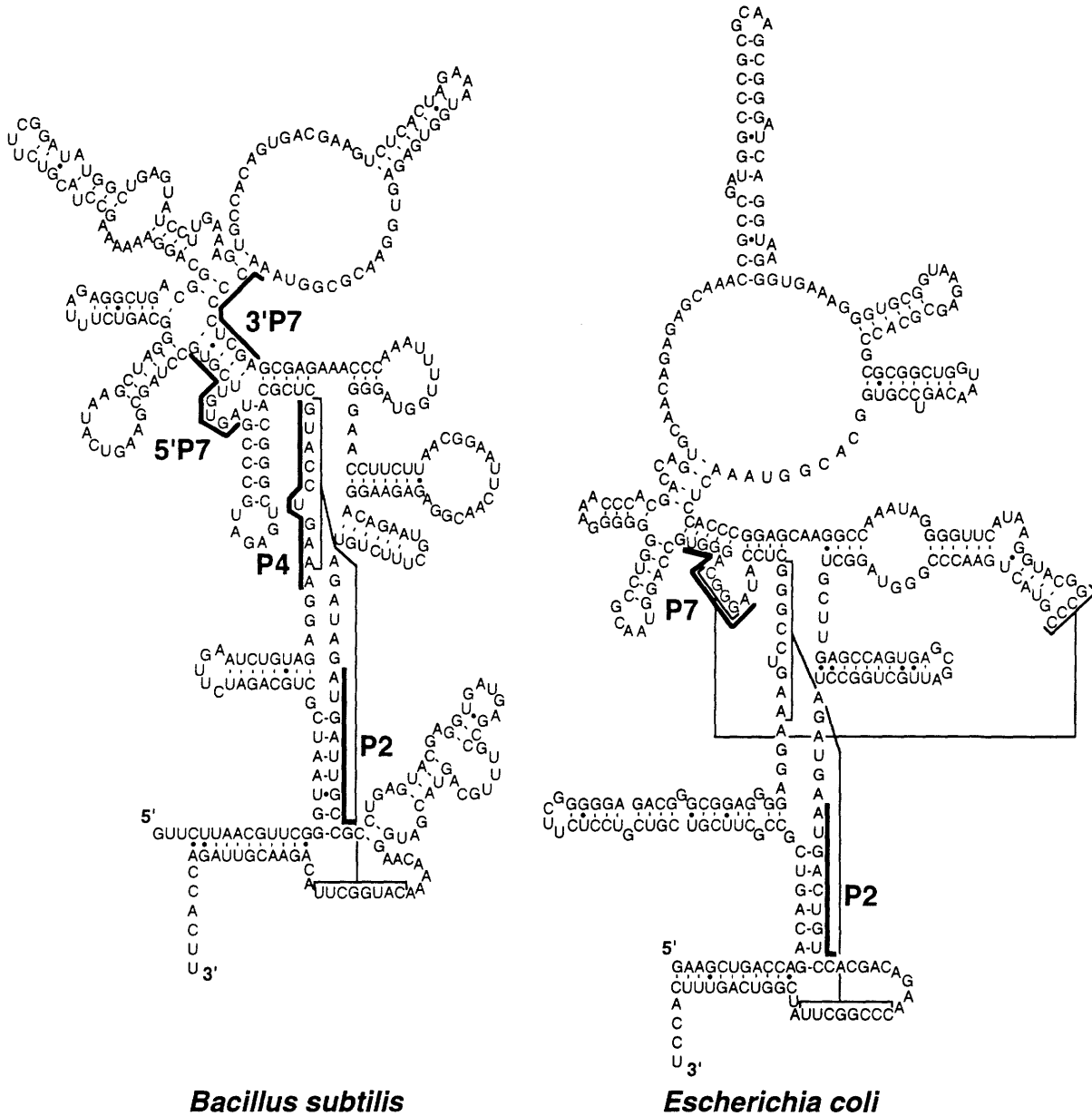


Figure 30. Secondary structures of RNase P RNA. Sequences complementary to oligonucleotide probes are indicated by bold lines. Probes are named according to the secondary structure element they target.

Adapting the kinetic oligonucleotide hybridization assay

Oligonucleotide screen

As a first step, the suitability of the kinetic oligonucleotide hybridization assay for studying folding of RNase P RNA was determined. As in the group I intron, formation of higher order structure in RNase P requires Mg^{2+} , while much of the secondary structure can form at very low concentrations of monovalent metal ions. The addition of Mg^{2+} together with a relatively high concentration of Na^+ or NH_4^+ , is therefore expected to allow formation of the native conformation from a partially folded state. To identify regions of the RNA which undergo changes in accessibility to oligonucleotide probes and/or RNase H upon addition of metal ions, a series of four probes complementary to sequences in *B. subtilis* RNase P RNA was tested (Fig. 30). The targeted nucleotides are involved in long range base pairing interactions, similar to those successfully studied with this assay in the *Tetrahymena* ribozyme (see chapter 3), which may be expected to require Mg^{2+} for their formation. As in the *Tetrahymena* intron, pairing regions in RNase P are numbered starting from the 5' end of the molecule. The probes were complementary to both strands of the P7 helix, one strand of the P4 helix or one strand of the P2 helix. P7 participates in a four helix junction and is part of what appears to be an independently folding subdomain (Pan, 1995). P4 is a long range interaction proposed to be a component of the active site (Harris and Pace, 1995), and P2 links regions near the 5' and 3' ends of the molecule. For each of these sequences it was asked whether the target is accessible to oligonucleotide binding and/or RNase H cleavage if the probe is added simultaneously with Mg^{2+} and a moderately high concentration of NH_4^+ , corresponding to the zero time point of a kinetic experiment, and whether it is accessible after the RNA has been equilibrated under folding conditions for 10 minutes (Fig. 31). Although maximum catalytic activity of RNase P

0 **10 minutes**
5'P7 P4 P2 3'P7 **5'P7 P4 P2 3'P7** **Probe**



RNA requires the presence of high concentrations of monovalent ions, RNase H activity is inhibited at high salt concentrations. The RNA was therefore folded in 100 mM NH_4Cl and 20 mM MgCl_2 , where RNase H retains full activity in the hybridization assay. While not optimal, these conditions still support RNase P catalytic activity (Tallsjö and Kirsebom, 1993; Kufel and Kirsebom, 1994), and should allow stable formation of the active structure (Pan, 1995). Both strands of P7 showed metal ion dependent differences in accessibility, with the effect being much more pronounced for the strand closer to the 5' end of the RNA (using probe 5'P7), indicating that P7 is not stable in the absence of metal ions. P2 also displayed differential accessibility with and without preincubation in folding conditions. P4, however, was completely inaccessible, even when the probe was added together with Mg^{2+} and NH_4^+ , suggesting that although it is a long range interaction, it is formed under less stringent conditions than other elements of the secondary structure.

Controls

The observed changes in the accessibility of the RNA upon folding suggested that RNase P is amenable to analysis by the oligonucleotide hybridization assay. The changes were substantial for the P2 and 5'P7 probes, and it was therefore decided to use these probes to investigate the kinetics of folding. As outlined in chapter 2 above, however, several conditions have to be met for the fraction of RNA cleaved by RNase H after various folding times to accurately reflect the fraction still unfolded at that time, and thus for the assay to report the actual folding rate in a kinetic experiment. First, binding of the probe and RNase H cleavage must be fast compared to folding, so that no additional folding takes place once probe and RNase H are added. Second, RNase H cleavage must be fast compared to dissociation of the RNA:probe complex to ensure that all of the RNA bound by the probe will be cleaved. Third, unfolding of the RNA under native conditions must be slow compared to folding, to prevent cleavage of RNA that had already folded at the time

probe and RNase H were added. Finally, high enough probe concentrations must be used to ensure rapid binding of all accessible RNA. Control experiments to ensure that each of these conditions was met in the standard assay procedure were performed with each probe used in kinetic experiments. The controls showed that probe binding and RNase H cleavage were complete in less than thirty seconds, that less than 10-20 percent of RNA:probe complex dissociated within 30 seconds after addition of the probe, and that no significant unfolding took place in the thirty seconds the RNA was exposed to probe and RNase H. The concentration of each probe necessary to produce maximal cleavage of the RNA was determined by measuring RNA cleavage at the zero time point, as described above, at a series of probe concentrations. The kinetic oligonucleotide hybridization assay therefore does accurately measure the rate of folding under the conditions used in the experiments reported here.

A complex folding pathway

Slow folding of P7

To measure the rate at which the changes in accessibility of P7 took place, the fraction of the RNA accessible to the P7 probe after different folding times was determined (Fig. 32A). Accessibility was found to decrease exponentially with a rate constant of $0.50 \pm 0.13 \text{min}^{-1}$ (the error represents the standard deviation from 5 independent experiments), suggesting that there is a slow step on the folding pathway of *B. subtilis* RNase P RNA and, furthermore, that this slow step is required for the stable formation of the P7 helix.

Next, it was asked whether the folding event monitored by this probe was dependent on Mg^{2+} by repeating the experiment at different Mg^{2+} concentrations, while keeping the NH_4^+ concentration constant (Fig. 32A). As the Mg^{2+} concentration was lowered, the fraction of RNA accessible to probe binding and/or RNase H cleavage at

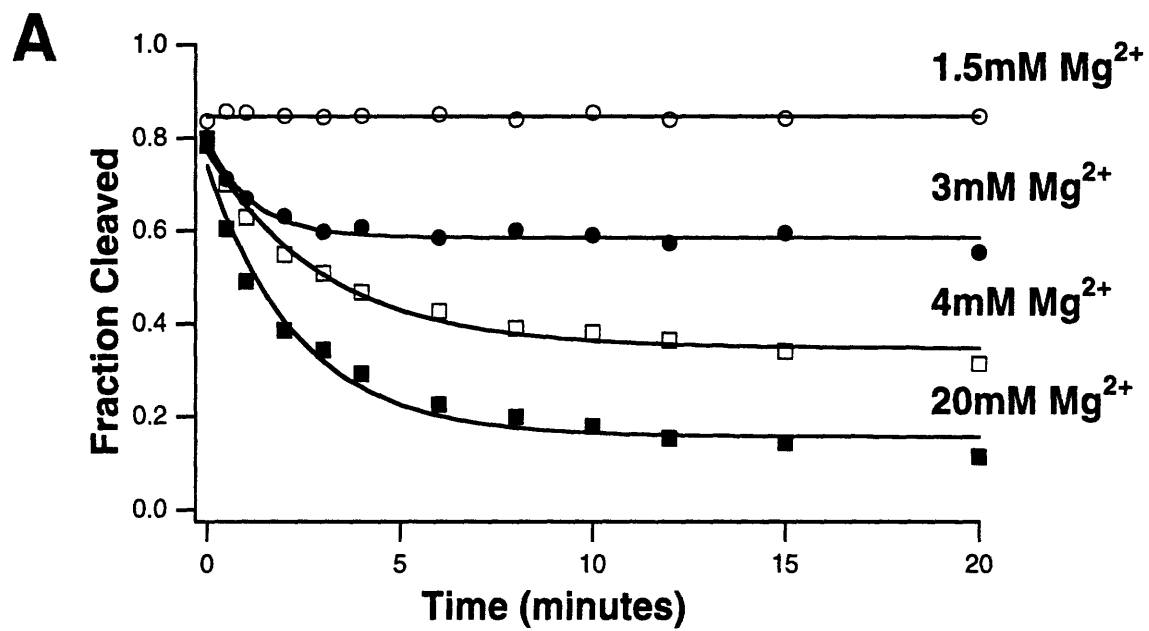


Figure 32. Folding of the P7 region in *B. subtilis* RNase P.
(A) Folding kinetics at different Mg²⁺ concentrations. The final probe concentration (P7 probe) was 70 μM.

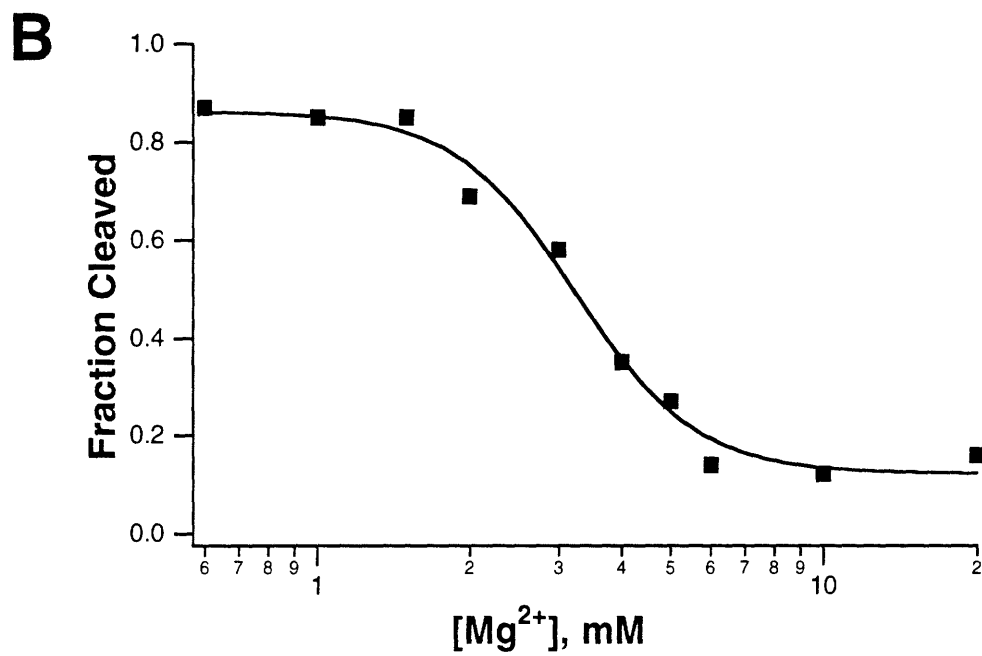


Figure 32. Folding of the P7 region in *B. subtilis* RNase P.
(B) Equilibrium Mg²⁺ concentration dependence. Each point represents the endpoint of a kinetic experiment as in (A). Values for [Mg²⁺]_{1/2} and n were 3.2 mM and 3.6, respectively.

equilibrium increased, indicating that the folding transition only took place in the presence of Mg^{2+} (Fig. 32A). However, the observed rate constant at which folding proceeded was independent of the Mg^{2+} concentration. Therefore, there must be at least two individual steps involved in this folding event (see Fig. 16): a slow, Mg^{2+} -independent step, the rate of which is measured in the hybridization assay, followed by a rapid Mg^{2+} binding step which drives the Mg^{2+} -independent step. These observations therefore reveal the presence of at least one kinetic intermediate on the folding pathway. The fraction of RNA cleaved at equilibrium at different Mg^{2+} concentrations was used to construct an equilibrium folding curve (Fig. 32B). The midpoint of the transition ($[Mg^{2+}]_{1/2} = 3.2$ mM) was very close to that reported by monitoring the protection of the RNA from Fe(II)-EDTA induced hydroxyl radical cleavage ($[Mg^{2+}]_{1/2} = 2-3$ mM)(Pan, 1995), suggesting that the two techniques may report the same folding transition.

Folding of P2

When the folding kinetics of the RNA were measured using a probe targeting the P2 helix, behavior very different from that of the P7 region was found. At 20 mM Mg^{2+} , a substantial fraction of the RNA remained accessible to the probe at equilibrium (Fig. 33). Increasing the Mg^{2+} concentration to 40 mM did not result in any additional protection. Conversely, omitting Mg^{2+} altogether, and folding the RNA in the presence of NH_4^+ only, still resulted in significant protection. Only when the NH_4^+ concentration was lowered to 10 mM, in the absence of Mg^{2+} , did the RNA become almost completely accessible to probe binding and/or RNase H cleavage (Fig. 33). Sodium was found to be as effective as NH_4^+ in inducing protection of the RNA, and the effect is therefore not specific for NH_4^+ . These observations indicate that different subregions of RNase P RNA exhibit distinctly different folding behavior upon the addition of metal ions.

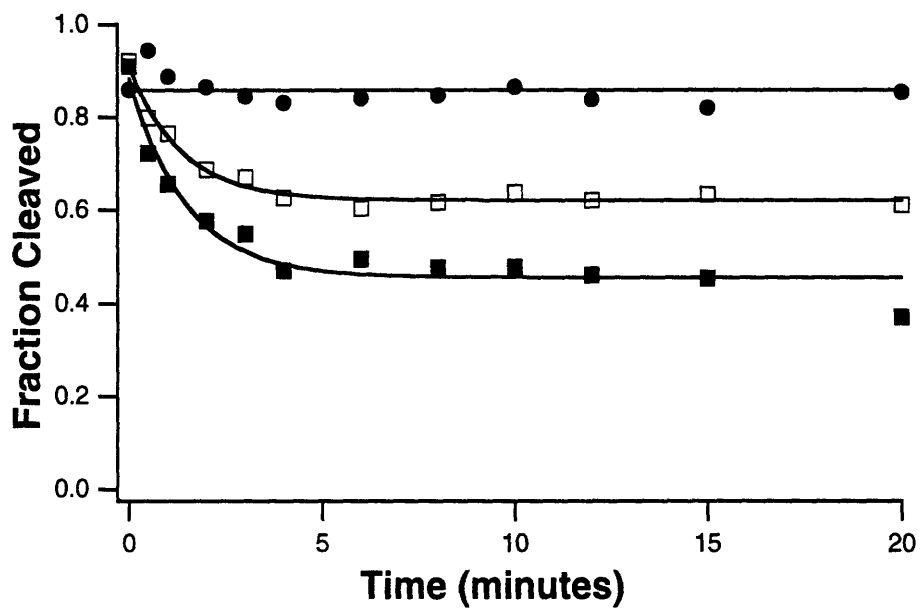


Figure 33. Formation of P2 in *B. subtilis* RNase P. Folding conditions were 20 mM MgCl₂/100 mM NH₄Cl (■), 0 mM MgCl₂/100 mM NH₄Cl (□), 0 mM MgCl₂/10 mM NH₄Cl (●). The final probe concentration (P2 probe) was 40 μM.

Folding of E. coli RNase P

To determine whether the folding behavior of RNase P is conserved between different species the experiments described above were repeated with the enzyme from *E. coli*, again using probes targeting P7 and P2. In *E. coli*, the secondary structure in the P7 region differs from that in *B. subtilis* RNA and nucleotides next to P7, targeted by the P7 probe, are directly involved in another long range interaction (Fig. 30). The folding rate constant measured with the P7 probe in *E. coli* (Fig. 34) ($k_{\text{obs.}} = 0.68 \pm 0.07 \text{ min}^{-1}$, where the error represents the standard deviation from 4 independent experiments) was nevertheless very similar to that found for *B. subtilis* RNA. As in *B. subtilis*, the apparent rate constant was independent of the concentration of Mg^{2+} , while the extent of folding at equilibrium declined at lower Mg^{2+} concentrations (Fig. 34A). The midpoint of the equilibrium transition ($[\text{Mg}^{2+}]_{1/2} = 2.4 \text{ mM}$ in *E. coli*, 3.2 mM in *B. subtilis*) was also similar in the two species (Fig. 34B).

The P2 region was again protected by NH_4^+ , even in the absence of Mg^{2+} (Fig. 35). In *E. coli*, however, the protection was almost complete, whereas in *B. subtilis* only partial protection of this region was observed. Furthermore, there was only a minimal loss of protection upon lowering the NH_4^+ concentration from 100 mM to 10 mM (Fig. 35). Even in the absence of any added metal ions, the P2 region was partially protected by the addition of buffer alone. The rate of the folding event monitored by the P2 probe appeared to change at different metal ion concentrations, but the changes were not systematic and not consistent enough to allow accurate quantitation. The data do suggest, however, that this region behaves qualitatively similar in both species of RNase P RNA examined, and that this behavior is distinctly different from that of the P7 region.

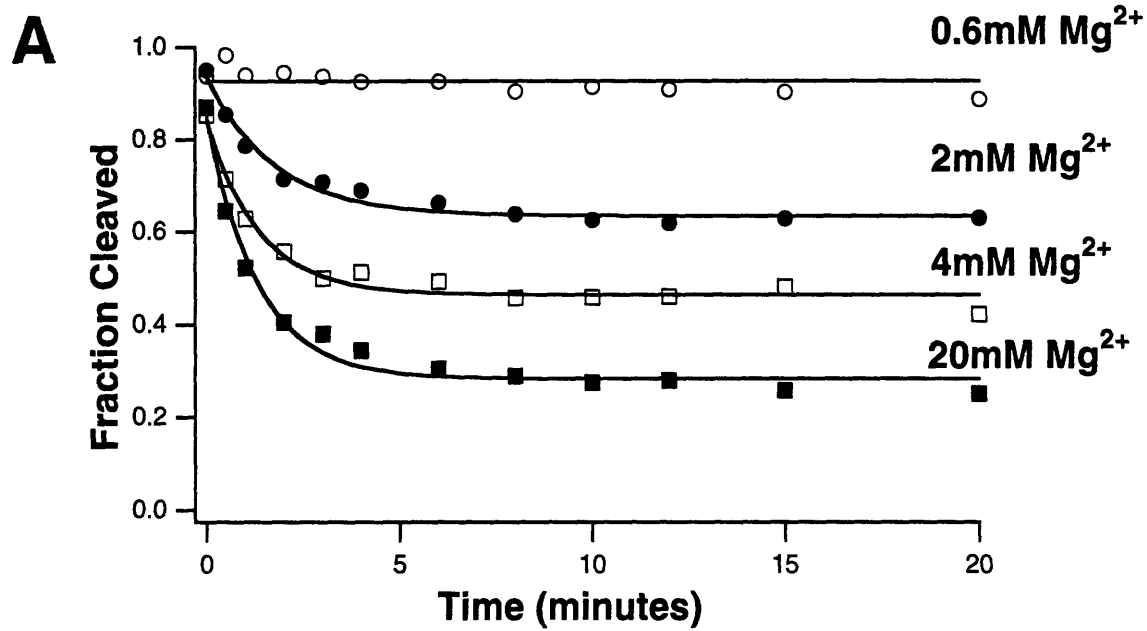


Figure 34. Folding of the P7 region in *E. coli* RNase P.
(A) Folding kinetics at different Mg²⁺ concentrations. The final probe concentration (P7 probe) was 70 μ M.

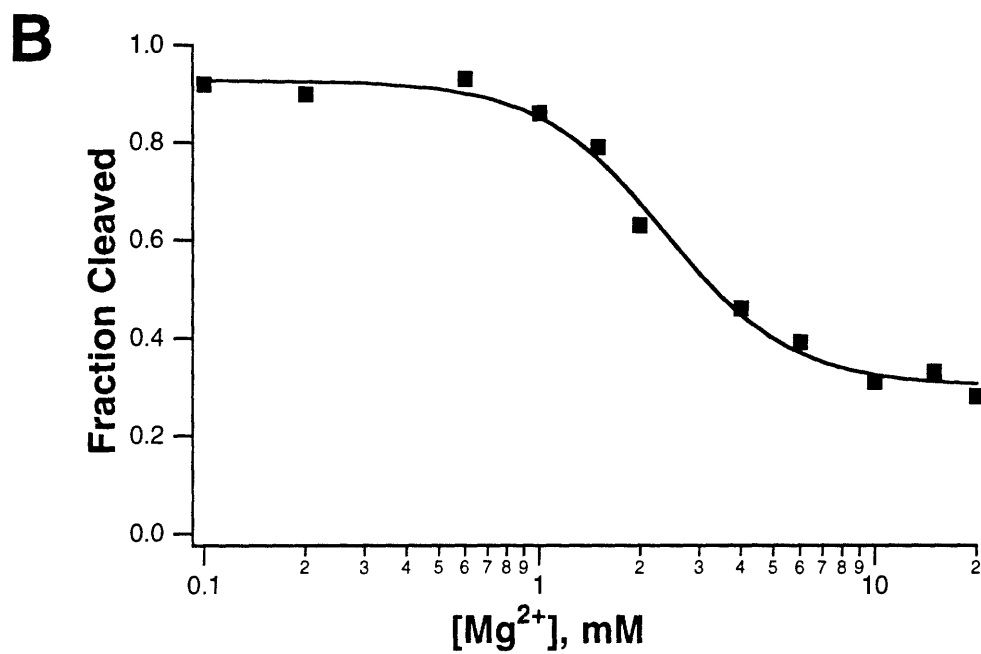


Figure 34. Folding of the P7 region in *E. coli* RNase P.
(B) Equilibrium Mg^{2+} concentration dependence. Each point represents the endpoint of a kinetic experiment as in (A). Values for $[Mg^{2+}]_{1/2}$ and n were 2.4 mM and 2.3, respectively.

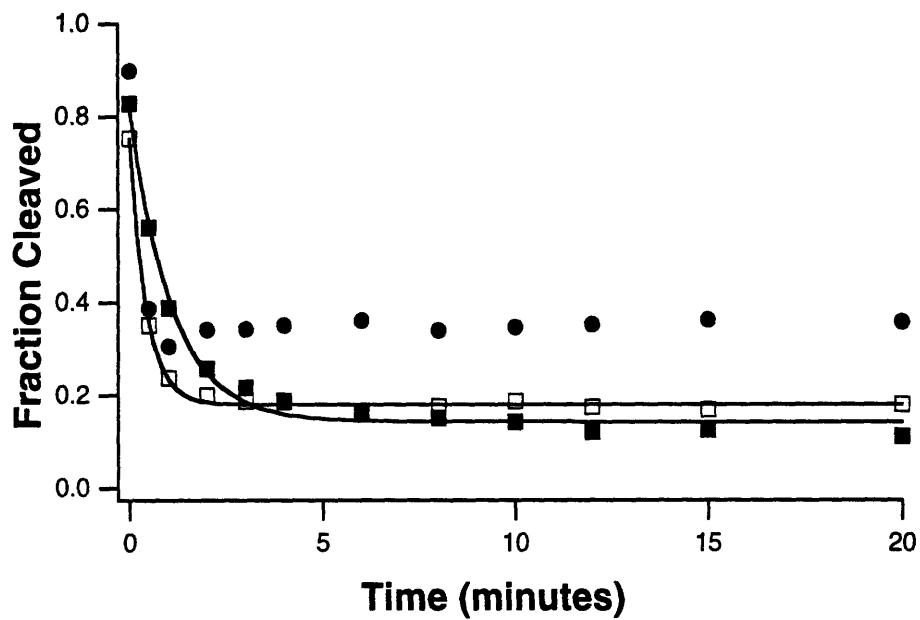


Figure 35. Formation of P2 in *E. coli* RNase P. Folding conditions were 20 mM MgCl₂/100 mM NH₄Cl (■), 0 mM MgCl₂/100 mM NH₄Cl (□), 0 mM MgCl₂/10 mM NH₄Cl (●). The final probe concentration (P2 probe) was 70 μM.

RNase P folding pathway

Together, these results show that the kinetic oligonucleotide hybridization assay can be applied to study RNase P folding. An initial screen of several regions of *B. subtilis* RNase P RNA revealed changes in the accessibility of the RNA to binding by short, complementary oligodeoxynucleotides and/or cleavage of the resulting complexes by RNase H upon the addition of metal ions. A series of control experiments show that information about folding kinetics can be gained from measuring the rates of these changes.

The most significant changes were observed with a probe targeting the P7 region of the RNA. When the rate at which these changes occurred was measured it was found to be slow, taking place on the timescale of minutes ($k_{\text{obs.}} = 0.50 \pm 0.13 \text{ min}^{-1}$) (Fig. 32A). The observation of a slow step with a probe targeting P7 and adjacent sequences suggests that structure formation in this region either is directly involved in the slow step, or that the rearrangement occurring during the slow step is required for structure to form in this section of the RNA. The Mg^{2+} concentration independence of the rate of this slow step shows that the structural rearrangement does not involve binding of Mg^{2+} . The existence of at least two steps must be invoked to explain the apparent contradiction of the requirement for Mg^{2+} , but Mg^{2+} independence of the slow step itself, providing evidence for at least one kinetic intermediate on the folding pathway.

The sequences targeted by the four probes used in the initial screen for changes in accessibility each exhibited a different pattern of protection (Fig. 31). The 5'P7 target sequence was almost completely accessible when the probe was added together with metal ions, but became almost completely inaccessible after incubation of the RNA under folding conditions. The 3' strand of P7, however, showed less dramatic changes in accessibility. This asymmetry observed for P7 is probably explained by the two probes targeting adjacent sequences which are part of different structural elements, in addition to nucleotides forming

P7 itself. P7 is one of the helices forming a four-way junction, and the 3'P7 probe is centered on this junction, possibly reporting mainly its formation or stabilization, rather than that of P7 itself. The results suggest that the four-way junction may be marginally stable in the absence of added metal ions, and become stabilized upon their addition. The 5'P7 probe, however, does not extend across the junction, but rather also targets nucleotides in a bulge adjacent to P7, and may therefore report predominantly the formation of structure involving this bulged region.

The equilibrium Mg^{2+} concentration dependence of folding detected with the 5'P7 probe is consistent with published results using Fe(II)-EDTA generated hydroxyl radicals as a probe for higher order structure formation (Pan, 1995). The midpoint of the transition ($[Mg^{2+}]_{1/2} = 3.2$ mM for oligonucleotide hybridization, 2-3 mM for hydroxyl radical footprinting), as well as the number of magnesium ions involved (approximately 3-4) was similar in both studies, suggesting that the two techniques report on the same folding event.

The P4 helix, which is a long range interaction and is proposed to be part of the active site (Harris and Pace, 1995), is inaccessible even if the probes are added together with metal ions, suggesting either that it is already stable in the absence of added metal, or that it forms extremely rapidly, much faster than binding of the DNA probe to its target sequence. In the Fe(II)-EDTA footprinting study (Pan, 1995), P4 was accessible in the absence of Mg^{2+} . This apparent discrepancy with the results shown here may be explained if base pairing in P4, which would afford protection from binding by the oligonucleotide probe, is stable early during folding and in the absence of Mg^{2+} , but if the higher order interactions, which afford protection from hydroxyl radical cleavage, do not occur unless Mg^{2+} is present. Protection of the P2 helix, which links sequences near the two ends of the RNA, is not absolutely dependent on Mg^{2+} , and occurs in the presence of moderate concentrations of monovalent ions. This suggests that base pairing in this helix may be

stabilized sufficiently to exclude the oligonucleotide probes even in the absence of higher order structure. The current representation of the RNase P secondary structure (Fig. 30) includes P2, while P4 is only indirectly indicated. The apparent greater stability and/or more rapid formation of P4 revealed by the results described here suggests that an alternative representation, which includes P4 and where P2 is indicated indirectly, may be more appropriate. The variable behavior exhibited by different regions of the RNA in response to metal ions suggests that folding of RNase P does not occur in a single, cooperative transition, but that structure is formed under different conditions in different parts of the RNA, and that the folding mechanism is likely to be complex.

Central features of RNA folding

While the secondary structure of RNase P RNA from *E. coli* differs in several places from that of *B. subtilis*, including the region surrounding the P7 helix (Fig. 30), the folding behavior of the two regions tested here in both molecules is very similar. There is a slow step involving structure formation in the P7 region in both RNAs, and this slow step, while only taking place in the presence of Mg^{2+} , does not involve binding of Mg^{2+} . The P2 region of both RNAs becomes at least partially inaccessible in the presence of only moderate concentrations of monovalent ions. In spite of the differences in degree, these findings imply at least some conservation of the folding mechanism between these two species of RNase P. While the extent of conservation remains to be defined, it is consistent with the hypothesis that the final structure and the process by which the structure is formed are linked, and do not evolve independently.

A comparison of RNase P folding with the folding mechanism of the *Tetrahymena* ribozyme, outlined in the preceding chapters, also reveals that several aspects of group I intron folding are conserved in RNase P. In both ribozymes, different regions exhibit distinctly variable patterns of accessibility to complementary oligonucleotide probe binding

and/or RNase H cleavage. In the *Tetrahymena* ribozyme, these differences reflect a complex kinetic folding mechanism with multiple steps and intermediates. It seems likely that a similarly complex folding pathway will be revealed for RNase P. The second clear similarity is the existence of a slow step not involving binding of Mg^{2+} , and taking place on the timescale of minutes, on the folding pathway of both RNAs. Both the rate of the slow step, and the concentration of Mg^{2+} required to allow it to proceed, are very similar in RNase P and the *Tetrahymena* group I intron. Furthermore, there are reports that folding of a group II intron (Griffin *et al.*, 1995) and another group I intron (Lewin *et al.*, 1995; Weeks and Cech, 1996) also is slow, taking place at timescales similar to those found here and for the *Tetrahymena* ribozyme. Results from the *Tetrahymena* ribozyme and from RNase P, together with these reports, suggest that slow kinetics and a complex folding pathway including multiple intermediates may be common features of higher order folding of large RNAs. The Mg^{2+} independence of the rate of the slow step in both RNase P and the group I ribozyme may also indicate that the nature of the rate limiting step itself may be similar in different RNAs. Furthermore, the recent demonstration that the RNA component of a simple RNP, composed of a group I intron and a specifically bound protein, must form its core structure before protein binding can occur shows that the kinetics of RNA folding can govern the kinetics of RNP assembly (Weeks and Cech, 1996). The results presented in this thesis may therefore be relevant not only to folding of RNAs, but also to assembly of multicomponent RNPs.

6. PARALLELS TO PROTEIN FOLDING AND FUTURE DIRECTIONS

Parallels to protein folding

Deciphering the processes by which proteins fold into their active conformations continues to be one of the central problems in biochemistry. Although this protein folding problem has been intensively investigated for more than twenty years, it is still a topic of vigorous debate and is far from being solved. New insights continue to emerge and reshape how protein folding is viewed. While there are obvious chemical differences between proteins and RNA, the basic problem of how to fold a linear polymer chain into a specific, globular conformation is shared between these two classes of macromolecules. It is therefore instructive to compare the insights on RNA folding afforded by the studies described in the previous chapters with current knowledge of protein folding mechanisms.

The induction of RNA structure formation by the addition of Mg^{2+} can be viewed as the 'inverse' of the denaturation of proteins by chemical agents. In the presence of denaturants such as urea, proteins are in a largely unfolded state. Formation of the folded structure is induced by reducing the denaturant concentration and can be followed either at equilibrium at different levels of denaturation, or kinetically by a concentration-jump experiment. Analogously, many RNAs are in a partially denatured state in the absence of Mg^{2+} , and the addition of a divalent metal to RNA is therefore equivalent to dilution of a denaturant for proteins, allowing application of the same experimental approaches used to study protein folding. The correlation of kinetic information with sequence by targeted hybridization obtained by the kinetic oligonucleotide hybridization assay resembles amide 1H - 2H exchange experiments (Woodward, 1994). For both, a snapshot of the folding

reaction is obtained by a rapid quench experiment. The state of accessibility to external probes is determined in a sequence specific way, yielding information not only on the rates of folding events, but also on which regions of the molecule are involved in a particular step. While large oligonucleotide probes are necessarily far less specific and sensitive than single protons, conceptually, the two techniques yield similar types of information.

More importantly than these experimental similarities, the folding pathway of the *Tetrahymena* ribozyme reveals significant parallels to protein folding mechanisms. For both RNA and proteins, short range secondary structure appears to form rapidly to yield a state in which much of the secondary structure is present, but which is still very flexible and lacks stable tertiary contacts (Lynch and Schimmel, 1974; Stein and Crothers, 1976a; Westhof and Michel, 1992; Christensen and Pain, 1994; Weeks and Cech, 1995). The native structure is then formed from this fluid state by the successive formation and stabilization of larger folding units which generally correspond to identifiable structural subunits. In proteins these folding units correspond to small sets of α -helices, turns, and β -sheets (Serrano *et al.*, 1992; Jennings and Wright, 1993; Dobson *et al.*, 1994; Bai *et al.*, 1995; Dobson, 1995; Wu *et al.*, 1995), while in RNA they consist of sets of stacked helices (Celander and Cech, 1991; Westhof and Michel, 1992; Lagerbauer *et al.*, 1994; Murphy and Cech, 1994; Pan, 1995; Weeks and Cech, 1995; Weeks and Cech, 1996). In both RNA and proteins, these subdomains seem to form in a hierarchical manner, where the presence of the fast forming elements may be required for formation of the slower folding subdomains (Lynch and Schimmel, 1974; Stein and Crothers, 1976a; Serrano *et al.*, 1992; Jennings and Wright, 1993; Lagerbauer *et al.*, 1994; Bai *et al.*, 1995; Loh *et al.*, 1995). The formation of specific long range contacts that allow the folding units to interact then occurs late on the folding pathway (Stein and Crothers, 1976a; Serrano *et al.*, 1992; Dobson, 1995).

Although these parallels are striking, the details of how folding is achieved diverge due to the different nature of the forces responsible for the stability of proteins and RNA. For proteins the early formation of secondary structure is accompanied or preceded by a hydrophobic collapse and the resulting state has been termed a molten globule (Christensen and Pain, 1994). Extensive secondary structure can be detected in this collapsed form in the absence of rigid tertiary packing, but individual secondary structure elements are rarely stable in isolation. In RNA the negatively charged backbone prevents such an early collapse and the compact state is achieved mainly through binding of Mg^{2+} after most of the secondary structure has already formed (Stein and Crothers, 1976a; Wang *et al.*, 1994; Nakamura *et al.*, 1995). Short range secondary structure elements, such as simple hairpins, can generally form in isolation. In spite of these differences, the results outlined here suggest that in both polymers, short range secondary structure forms fast and early, structural subdomains fold as kinetic units in a hierarchical fashion, and specific tertiary interactions between subunits form late during the folding process. These fundamental parallels in the mechanisms used by proteins and RNA to achieve their native conformation indicate that at least some of the strategies of macromolecular folding transcend chemical boundaries.

The intrinsic importance of observed intermediates to productive protein folding is still unresolved after many years of intensive investigations, and the protein folding field continues to evolve rapidly (Baldwin, 1996). By comparison, studies of RNA folding, particularly kinetic studies, are still relatively sparse. It is likely that views of RNA folding will similarly evolve as more data becomes available. The results presented here give only a first indication of the parallels between RNA and protein folding, and additional studies may well extend these similarities. The common folding strategies also suggest that the

accumulated insights into protein folding mechanisms may be expected to productively guide future studies of RNA folding.

Future directions

Several issues in the folding mechanism of the *Tetrahymena* ribozyme remain unresolved, and are likely to be fruitful areas for future investigations. First, what is the exact nature of the overall rate limiting step? While several interactions have been identified here as being formed during the slow step on the folding pathway, there is no direct evidence to show whether the formation of any of these is rate limiting, or whether there are additional rearrangements which occur during this step. Second, do any of the observed intermediates contain non-native interactions, and, if they exist, what, if any, is the role of such non-native intermediates during folding? Evidence is mounting that, at least in some cases, the folding rate of proteins is limited by the disruption of non-native interactions, and that in their absence the intrinsic rate of protein folding is extremely rapid (Sosnick *et al.*, 1994; Kiefhaber, 1995; Kotik *et al.*, 1995). Results from early studies of tRNA hinted at the possibility of a similar role for non-native interactions during tRNA folding under some conditions (Cole *et al.*, 1972; Yang and Crothers, 1972). Third, do the two structural modules that compose the P4-P6 subdomain (Murphy and Cech, 1993; Murphy and Cech, 1994) also correspond to kinetic folding units, and what is the temporal order of their formation? Studies at equilibrium show that the P5abc module is required for formation of the P4 and P6 helices (Murphy and Cech, 1993; Murphy and Cech, 1994), but this hierarchy remains to be confirmed kinetically. Fourth, is the intermediate I_1 a productive intermediate (see Fig. 29), or does it represent a kinetic trap, and thus a nonproductive branch of the folding pathway? Results with the P4 mutant (see chapter 4) suggest that the P4 helix contributes to the equilibrium between I_U and I_1 , but do not show whether I_1 can be bypassed entirely. Fifth, how exactly does the P9.1-P9.2 extension

accelerate folding of the intron, and do other peripheral extensions play a similar role?

Sixth, what are the sites in which Mg^{2+} binds during the two Mg^{2+} -dependent folding steps in the model? In addition to these specific issues, many of the molecular details of the rearrangements occurring during each kinetic step remain to be defined.

To address these questions, additional mutations will have to be introduced into the ribozyme, and their folding properties investigated. In some cases, specific mutations likely to be informative can be identified. In other cases, *in vitro* selection strategies will have to be applied to search for mutations with the desired properties. For example, the mutations in P5abc and P6 which allowed the establishment of the hierarchical relationship between the two structural modules of the P4-P6 subdomain at equilibrium (Murphy and Cech, 1993; Murphy and Cech, 1994) are also likely to provide insight into the kinetics of P4-P6 subdomain formation. Similarly, the significance of the proposed interactions between the P9.1-P9.2 extension and the loops L2.1 and/or L2.2 (Banerjee *et al.*, 1993) can be tested by mutating the specific nucleotides thought to be involved. Additional mutations in P3, P7, and in the triple helical scaffold should also be constructed and tested to further confirm the conclusions reached here. To determine the nature of the rate limiting step, on the other hand, the interactions being formed in the transition state must be identified. One way to accomplish this is to find mutations which accelerate folding. Such mutations are expected either to stabilize interactions formed during the rate limiting step, or, if the breaking of interactions (native or non-native) is rate limiting, destabilize these interactions. An *in vitro* selection strategy based on the oligonucleotide hybridization assay is most likely to identify such mutations.

To thoroughly investigate the folding properties of interesting mutant ribozymes and to provide additional details about folding of the wild type intron, new techniques, in addition to the kinetic oligonucleotide hybridization assay, should be developed. One

drawback to using oligonucleotides as probes is their large size. Small chemical probes are likely to provide higher resolution information of events occurring during each folding step. It is likely that standard chemical modification protocols, used to study RNA structure at equilibrium (Jaeger *et al.*, 1990; Celander and Cech, 1991; Christian and Yarus, 1992; Rudinger *et al.*, 1992; Banerjee *et al.*, 1993; Christian and Yarus, 1993), can be adapted to provide kinetic information. In addition to the higher resolution afforded by such small probes, such techniques may also make it possible to access faster timescales through use of a rapid quench apparatus. Fluorescence spectroscopy provides a second means for accessing faster timescales, although significant effort is required to incorporate fluorescent probes at specific sites and correlation of spectroscopic changes with structural rearrangements is difficult. The ability to monitor fast folding events will be important to define the events leading to P4-P6 subdomain formation more clearly.

The identification and characterization of a series of additional mutant ribozymes using a variety of techniques will reveal more detailed features of *Tetrahymena* ribozyme folding, and will significantly refine the current model for the folding mechanism of this RNA. Additional effort is also required, however, to determine how far this folding mechanism is conserved in other RNAs and RNPs. The studies described in chapter 5 suggest that RNase P represents a second convenient model system in which to study RNA folding, and that at least some features of the folding mechanism are conserved between group I introns and RNase P. To determine the extent of this conservation, the folding pathway of RNase P should be defined using techniques and strategies similar to those applied to the *Tetrahymena* intron. In addition to studying folding of the RNA itself, it also will be crucial to determine the effect of the RNase P protein on folding. Does the protein actively participate, or does it merely bind to and stabilize the final structure?

For any future *in vitro* studies of RNA folding, it will continue to be useful to examine the results in the context of protein folding mechanisms. Such comparisons should help reveal how far the parallels apparent now extend, and may guide the interpretation and design of additional experiments.

Finally, the *in vitro* results described in this thesis may make it possible to begin to examine RNA folding *in vivo*. There has been much speculation on the role of proteins in folding and assembly of RNAs and RNPs *in vivo*. Experiments suggest that nonspecific RNA binding proteins can accelerate folding *in vitro*, possibly by preventing misfolding (Herschlag, 1995). These results, however, are based on functional assays and do not address folding directly. They also do not show a clear relation between *in vitro* observations and the *in vivo* situation. The availability of kinetic folding mutants, combined with a strategy that allows measuring the rate of interconversion of different length RNA species directly *in vivo* (as, for example, described in (Decker and Parker, 1993)) should make it possible to determine whether folding follows the same basic mechanism *in vitro* and *in vivo*, or whether additional factors are involved.

APPENDIX: MATERIALS AND METHODS

Construction of mutant ribozymes

Plasmid templates for the transcription of mutant ribozymes were constructed using standard recombinant PCR techniques (Higuchi, 1990). A fragment of the wild type pT7L-21 plasmid (Zaug *et al.*, 1988) containing most of the coding sequence for the intron and flanked by an Sph I site in the region coding for P2 and a Hind III site 29 base pairs downstream from the Sca I site used to linearize the template for transcription was replaced with a an equivalent fragment containing the desired mutation(s). The mutant fragment was obtained through two successive PCR steps (Fig. 36). The first step (PCR#1) consisted of two separate PCR reactions performed with different primers. One reaction included a 5' primer complementary to the sequence upstream from the Sph I site, terminating within that site (*5' Primer*), and a mutagenic 3' primer 20 nucleotides long and complementary to the sequence surrounding the site of the desired mutation, except for the mutation itself (*Mut 3'*). The second reaction (PCR#2) included a mutagenic 5' primer (*Mut 5'*) complementary to the mutagenic 5' primer used in the first reaction, and a 3' primer complementary to the sequence downstream from the Hind III site, terminating within that site (*3' Primer*). For the first step, 50 μ l reactions were carried out using 200 ng pT7L-21 template, 1.25 units *Pfu* DNA polymerase (Stratagene), 0.2 mM of each dNTP and 1 μ M of each primer in 1X *Pfu* Buffer (Stratagene). The PCR conditions were: 1 minute at 94°C, 2 minutes at 60°C, 2 minutes at 72°C for 30 cycles followed by incubation for 10 minutes at 72°C. 10 μ l from each reaction were run on a two percent agarose gel to verify that the main products were of the expected length. Product bands were excised and transferred to 0.5 mL siliconized microcentrifuge tubes using extreme caution not to contaminate the samples with wild type

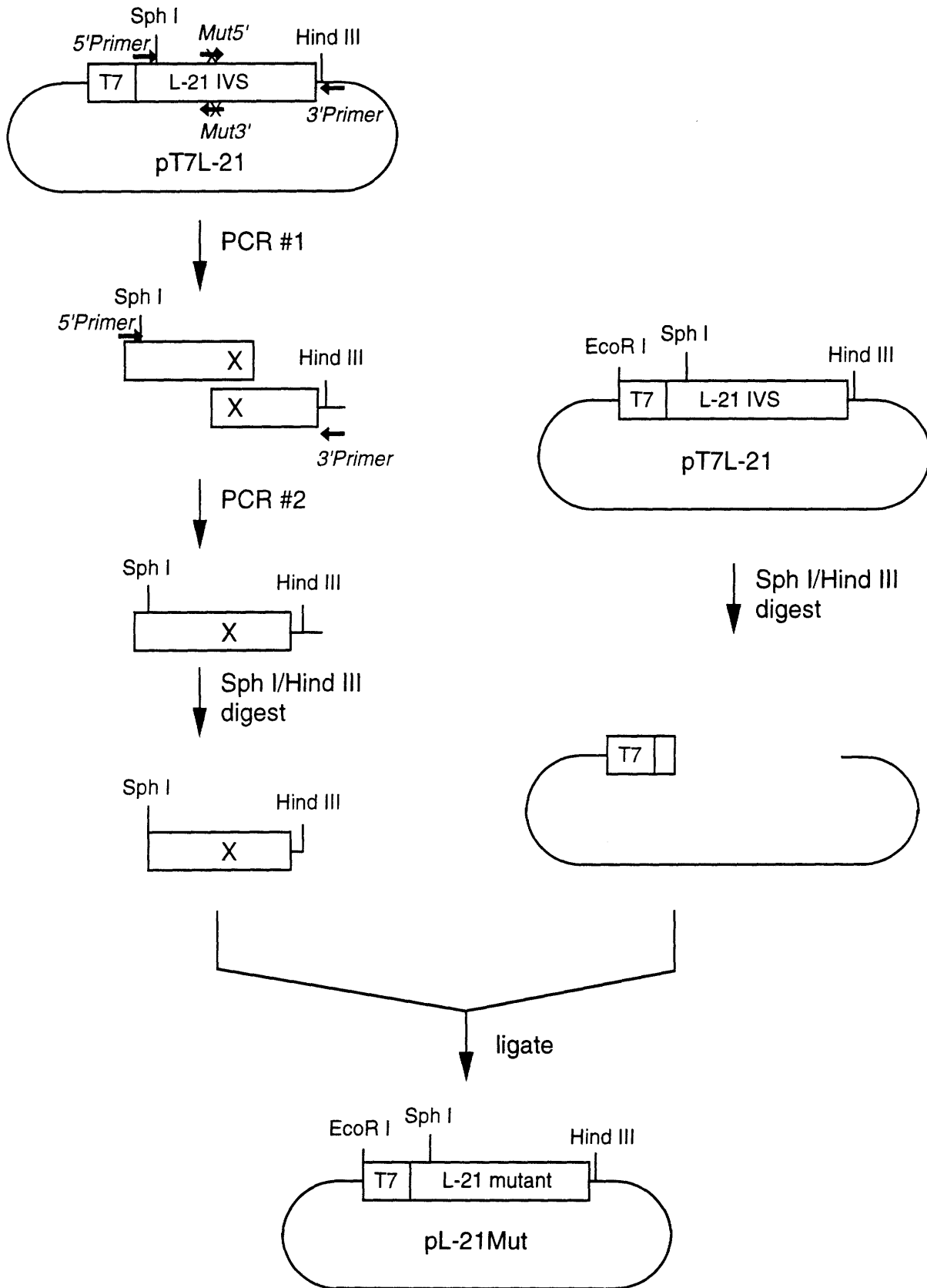


Figure 36. Recombinant PCR strategy for construction of ribozyme mutants.

plasmid. To elute the products from these gel slices, 100 μ l H₂O was added to each slice, and the tubes allowed to sit at room temperature with periodic vortexing for two hours or overnight at 4°C. The products from the two PCR reactions in the first step are partially complementary to each other. When they are added together during the second PCR step, the separated strands from the two products will anneal to each other, and will be elongated to form full length fragment, which can then be amplified in the same reaction using the 5' *Primer* and 3' *Primer* from the first step. The second PCR step accordingly consisted of a single 100 μ l reaction containing 15 μ l of each of the elution solutions containing products from the first step, 0.2 mM of each dNTP, 1 μ M each 5' *Primer* and 3' *Primer* and 2.5 units *Pfu* polymerase in 1X *Pfu* buffer. The temperatures and times were the same as for the first step reactions. The product from this second step was ethanol precipitated, gel purified, digested with Sph I and Hind III to produce the desired mutant fragment, and gel purified again (best results for gel purifications were obtained using Promega Wizard PCR Preps and two percent low melting point agarose gels). It was then ligated to pT7L-21 that had been digested with Sph I and Hind III, gel purified to remove the wild type fragment and dephosphorylated. Ligation reactions contained 300 ng pure, digested pT7L-21, 60 pg PCR product from step two and 400 units T4 DNA ligase (New England Biolabs) in 20 μ l 1X ligase buffer (New England Biolabs), and were allowed to proceed for 2.5 hours at 16°C. The reaction mixtures were used directly in the chemical transformation of JM109 *E. coli* and mutant constructs identified by sequencing. Primer sequences were: 5' *Primer*: 5'-GGA GGG AAA AGT TAT CAG GC-3'; 3' *Primer*: 5'-AAA CGA CGG CCA GTG CCA AG-3'.

All sequencing was performed using Sequenase kits (United States Biochemical/Amersham) and recommended protocols. Three primers were used to provide unambiguous information for the region containing the template sequence and the T7

promoter: 5'-GGA ATT GTG AGC GGA-3' (complementary to the sequence 40 base pairs upstream from the T7 promoter); 5'-GTC TGT GAA CTG CAT-3' (complementary to the noncoding strand in the region coding for ribozyme nucleotides 252-266); 5'-GCT GAC GGA CAT GGT-3' (complementary to the coding strand in the region of ribozyme nucleotides 188-202).

Once individual transformants containing the desired mutations were identified, a large amount of the mutant plasmid was prepared using Wizard Megaprep plasmid prep kits (Promega). Each final plasmid preparation, including those of the wild type, was sequenced to confirm the absence of undesired and the presence of desired mutations.

Ribozyme preparation

L-21 RNA was transcribed from plasmid pT7L-21, or mutated derivatives of this plasmid, linearized with Sca I or Nhe I. Reactions (100 μ l) were performed for 3.5 hours at 37°C in 40 mM Tris•HCl (pH 7.5), 2 mM spermidine, 10 mM dithiothreitol, 25 mM MgCl₂, 1 mM each GTP, CTP, UTP, 0.1 mM ATP and 250 μ Ci of [α -³²P] ATP (New England Nuclear) (for unlabeled RNA the concentration of each NTP was 4 mM), with 500 units of T7 RNA polymerase (New England Biolabs) and 15 μ g of template. The reaction was quenched with stop solution (90 mM EDTA and gel running dyes in 82 percent formamide) and the product purified on a six percent polyacrylamide gel and soaked into buffer containing 10 mM Tris•HCl (pH 7.5), 1 mM EDTA, and 0.3 M sodium acetate. After ethanol precipitation and resuspension in 10 mM Tris•HCl (pH 7.5)/0.1 mM EDTA the RNA was quantitated by Cerenkov counting. Unlabeled RNA was quantitated either by measuring the OD₂₆₀, using the published extinction coefficient (Zaug *et al.*, 1988), or kinetically by measuring the stoichiometric burst of product formation in a multiple turnover reaction (Herschlag and Cech, 1990a).

RNase P RNA was prepared similarly from plasmids pDW66 (*B. subtilis*) and pDW98 (*E. coli*), generously provided by Prof. Norman Pace (Indiana University), except that only 5 mM MgCl₂ was used, and nucleotide concentrations were changed to 1 mM each ATP, CTP, UTP, 0.1 mM GTP and 400 μCi of [α-³²P] GTP (New England Nuclear).

Oligodeoxynucleotide probe preparation

Oligodeoxynucleotide probes were synthesized on a 1 μmol scale on an Applied Biosystems DNA synthesizer, deprotected overnight at 65°C in 2 ml concentrated ammonium hydroxide, and purified on 20 percent denaturing polyacrylamide gels. One full synthesis was run on each gel, using a single well across the width of the gel (9-10 inches). Full length bands were excised and eluted from the gel overnight into 40 ml water at 4°C, followed by desalting on C₁₈ Sep-Paks® (Waters). The Sep-Paks® (two for each synthesis) were pre-equilibrated with 6 mL acetonitrile and 9 mL H₂O, before the soaking solution was loaded (1-2 ml/min⁻¹). Each pak was washed with 10 ml H₂O before the probes were eluted in 2 mL 50 percent acetonitrile/H₂O (per pak). Samples were dried under vacuum (in a Savant speed-vac) and each synthesis resuspended in 150 μl 10 mM Tris•HCl (pH 7.5)/0.1 mM EDTA. Probe concentrations were determined from the absorbance at 260 nm (extinction coefficients were calculated for each probe from the extinction coefficients of nucleotides at 260 nm).

Kinetic oligonucleotide hybridization assay

Ribozyme (final concentration 1 nM) in 60 μl of buffer containing 1 mM Tris•HCl (pH 7.5) and 0.01 mM EDTA was annealed by heating to 95°C for 45 seconds followed by equilibration at 37°C for 3 minutes. Folding was initiated by addition of an equal volume of 2X folding buffer (1X: 50 mM Tris•HCl (pH 7.5), 10 mM NaCl, 1 mM dithiothreitol;

for RNase P folding experiments, 100 mM NH_4Cl was substituted for the NaCl) which also contained MgCl_2 to give the desired concentration. Aliquots (10 μl) were taken at the times indicated and added to 10 μl of 1X folding buffer containing oligonucleotide probe and RNase H (United States Biochemical, final concentration 0.1 U/ μl) and enough MgCl_2 to bring the final concentration to 10 mM (20 mM for RNase P). Oligonucleotide binding and RNase H cleavage were allowed to proceed for 30 seconds before the reaction was quenched with 14 μl of stop solution. The zero time points were obtained by adding oligonucleotide probe and RNase H in 2X folding buffer and MgCl_2 (final concentration 10 mM for the group I intron, 20 mM for RNase P) to the RNA immediately after annealing in a separate reaction. Products were separated on six percent denaturing polyacrylamide gels and quantitated using a Molecular Dynamics PhosphorImager. The data were fit to single exponentials [$f = f_{\text{equ.}} + (f_0 - f_{\text{equ.}})\exp(-k_{\text{obs.}}t)$], where f is the fraction cleaved at time t , $f_{\text{equ.}}$ is the fraction cleaved at equilibrium, f_0 is the fraction cleaved at time $t = 0$, and $k_{\text{obs.}}$ is the observed rate constant. This fitting procedure allows the endpoints (f_0 and $f_{\text{equ.}}$) as well as the rate ($k_{\text{obs.}}$) to vary, and independent values for the rate constant and equilibrium endpoint are obtained.

Mg^{2+} concentration dependence of folding

To obtain the Mg^{2+} concentration dependence of folding at equilibrium for probes targeting P3 or P7 in the *Tetrahymena* ribozyme, and for probes targeting P7 in RNase P, the Mg^{2+} concentration in the folding buffer was varied, and kinetic experiments as described above were performed at each Mg^{2+} concentration. During the probe binding/ RNase H cleavage step the Mg^{2+} concentration was always adjusted to 10 mM (20 mM for RNase P). The equilibrium endpoint from the fit of the data to single exponentials yielded the fraction of RNA folded at equilibrium for each Mg^{2+} concentration.

For probes targeting P4 or P6 in the *Tetrahymena* ribozyme this procedure did not produce accurate results due to the small amplitude of the change upon addition of Mg^{2+} , and an alternative strategy was applied where oligonucleotide probe was added after incubation of the RNA in the $MgCl_2$ concentration to be tested, but before the total $MgCl_2$ concentration was brought to 10 mM for RNase H cleavage. For each Mg^{2+} concentration, RNA was annealed in 10 μ l of buffer containing 1 mM Tris•HCl (pH 7.5)/0.01 mM EDTA as described, before 10 μ l 2X folding buffer containing enough $MgCl_2$ to give the desired concentration was added. The resulting solution was incubated at 37°C for 10 minutes, and then enough oligonucleotide to give the indicated concentrations (usually 60-70 μ M) added in 20 μ l 1X folding buffer and $MgCl_2$ to keep the Mg^{2+} concentration constant. After 0.5, 1 and 2 minutes 10 μ l aliquots were added to 10 μ l 1X folding buffer containing RNase H (final concentration 0.1 U/ μ l) and enough $MgCl_2$ to bring the final concentration to 10 mM, and allowed to react for 30 seconds before being quenched with 14 μ l stop solution. The average of the fraction cleaved for the three time points taken at each Mg^{2+} concentration was taken as the fraction of RNA unfolded at equilibrium at that concentration.

The data from both strategies were fit to an expression of two state binding of n Mg^{2+} ions [$f = 1/\{([Mg^{2+}]/[Mg^{2+}]_{1/2})^n + 1\}$], where f is the fraction cleaved at equilibrium, n the number of Mg^{2+} ions bound per RNA molecule, and $[Mg^{2+}]_{1/2}$ the midpoint of the transition, to obtain independent values for n and $[Mg^{2+}]_{1/2}$.

Oligonucleotide screens

The zero time point in the initial oligonucleotide screens (Figs. 10 and 31) for both the *Tetrahymena* ribozyme and for RNase P was performed exactly as described above, except that higher probe concentrations were used (300-400 μ M). For the ten minute time point, the RNA was annealed as described in 5 μ l buffer and equilibrated for ten minutes after addition of one volume 2X folding buffer containing 10 mM (for the group I intron)

or 20 mM (for RNase P) MgCl_2 (final concentration), before oligonucleotide probe (final concentration the same as that in the zero time point) and RNase H (0.1 U/ μl final concentration) were added in 10 μl of 1X folding buffer (maintaining the MgCl_2 concentration). The reaction was allowed to proceed for 30 seconds and quenched with 0.7 volumes stop solution. Products were separated on six percent polyacrylamide gels.

Ribozyme activity assays

To measure k_c for wild type and mutant introns, ribozyme (final concentration 50 nM) in 30 μl 0.1X TE buffer was annealed as described, added to 30 μl 2X reaction buffer (1X: 50 mM MES (2-[N-Morpholino]ethanesulfonic acid) (pH 5.5), 10 mM MgCl_2 , 10 mM NaCl, 1 mM dithiothreitol) and allowed to fold at 37°C for 10 minutes (wild type, P3 mutant, P7 mutant) or 25 minutes (triple-scaffold mutant) before the reaction was initiated by adding 5'- ^{32}P labeled CCCUCUAAAAA substrate (final concentration 0.25 nM) and GTP (final concentration 0.5 mM) in 60 μl 1X reaction buffer. Aliquots (10 μl) were taken at increasing times and added to one volume stop solution to quench the reaction. Products were separated on 20 percent denaturing polyacrylamide gels, quantitated as described, and the data were fit to single exponentials. Saturating conditions were ensured by independently increasing the concentrations of ribozyme (to 100 nM) and GTP (to 2 mM) for each RNA and confirming that this did not result in a change of the observed rate.

To measure the rate of gain of full catalytic activity, ribozyme (final concentration 7 nM) was annealed in 30 μl buffer, as described, and an equal volume of 2X folding buffer (1X: 50 mM Tris•HCl (pH 7.5), 10 mM MgCl_2 , 10 mM NaCl, 1 mM dithiothreitol) was added. The RNA was allowed to fold at 37°C for the specified amount of time before the reaction was initiated by addition of 5'- ^{32}P -labeled CCCUCUAAAAA substrate (final concentration 16 nM) in 60 μl 1X folding buffer containing GTP (final concentration 0.5 mM). Samples (10 μl) were taken at increasing times and added to 8 μl stop solution. For

the zero time point, substrate and GTP were added together with folding buffer immediately after annealing. Products were separated on 20 percent denaturing polyacrylamide gels, and the results quantitated as above. The folding rate was determined by plotting the extent of substrate cleavage after 30 seconds as a function of the time allowed for the ribozyme to fold before substrate was added. Near stoichiometric concentrations of ribozyme and substrate were used, instead of more standard single or multiple turnover conditions, in order to be able to follow the time course of gain of full activity. In large ribozyme excess, fully active ribozyme accumulates too rapidly to observe changes in initial reaction rate, whereas in large substrate excess the fraction of product formed in the very early stages of the reaction is too small for accurate quantitation.

RNA sequencing by iodine cleavage

To allow direct sequencing of RNA, ribozymes containing a low level of phosphorothioates were produced. Transcription conditions were as described above, except that the concentration of each NTP was 2 mM. Each of four separate reactions was supplemented with 0.1 mM of a different [α -S] NTP. RNA was purified as described and 5'- or 3'-³²P labeled using standard protocols. To generate sequencing ladders, iodine cleavage was carried out by adding 1 μ l of I₂-saturated H₂O to 5 μ l labeled RNA, allowing cleavage to proceed at room temperature for 5 minutes, followed by addition of 5 μ l stop solution. Reactions can also be quenched using thiosulfate. The I₂-saturated H₂O was prepared fresh each day by adding one or two I₂ crystals to 0.5 mL H₂O and vortexing vigorously every 30 seconds for several minutes. After addition of stop solution reactions were loaded directly onto sequencing gels.

REFERENCES

- Allain, F. H.-T., and Varani, G. (1995). Structure of the P1 helix from group I self-splicing introns. *J. Mol. Biol.* **250**, 333-353.
- Altman, S., Kirsebom, L., and Talbot, S. (1995). Recent studies of RNase P. In *tRNA: Structure, Biosynthesis and Function*, Söll, D. and RajBhandary, U., eds. (Washington, D.C.: ASM Press), pp. 67-78.
- Bai, Y., Sosnick, T., Mayne, L., and Englander, S. W. (1995). Protein folding intermediates: native-state hydrogen exchange. *Science* **269**, 192-197.
- Baldwin, R. L. (1996). On-pathway versus off-pathway folding intermediates. *Folding & Design* **1**, R1-R8.
- Banerjee, A. R., Jaeger, J. A., and Turner, D. H. (1993). Thermal unfolding of a group I ribozyme: the low-temperature transition is primarily disruption of tertiary structure. *Biochemistry* **32**, 153-163.
- Banerjee, A. R., and Turner, D. H. (1995). The time dependence of chemical modification reveals slow steps in the folding of a group I ribozyme. *Biochemistry* **34**, 6504-6512.
- Barford, E. T., and Cech, T. R. (1989). The conserved U•G pair in the 5' splice site duplex of a group I intron is required in the first but not the second step of self-splicing. *Mol. Cell. Biol.* **9**, 3657-3666.
- Bartel, D. P., and Szostak, J. W. (1993). Isolation of new ribozymes from a large pool of random sequences. *Science* **261**, 1411-1418.
- Bassi, G. S., Møllegaard, N.-E., Murchie, A. I. H., von Kitzing, E., and Lilley, D. M. J. (1995). Ionic interactions and the global conformations of the hammerhead ribozyme. *Nature Struct. Biol.* **2**, 45-55.
- Beaudry, A. A., and Joyce, G. F. (1990). Minimum secondary structure requirements for catalytic activity of a self-splicing group I intron. *Biochemistry* **29**, 6534-6539.
- Beaudry, A. A., and Joyce, G. F. (1992). Directed evolution of an RNA enzyme. *Science* **257**, 635-641.
- Beebe, J. A., and Fierke, C. A. (1994). A kinetic mechanism for cleavage of precursor tRNA^{Asp} catalyzed by the RNA component of *Bacillus subtilis* ribonuclease P. *Biochemistry* **33**, 10294-10304.
- Been, M. D., and Cech, T. R. (1986). One binding site determines sequence specificity of *Tetrahymena* pre-rRNA self-splicing, *trans*-splicing and RNA enzyme activity. *Cell* **47**, 207-216.

- Bertrand, E., and Rossi, J. J. (1994). Facilitation of hammerhead ribozyme catalysis by the nucleocapsid protein of HIV-1 and the heterogeneous nuclear ribonucleoprotein A1. *EMBO J.* 13, 2904-2912.
- Bevilacqua, P. C., Kierzek, R., Johnson, K. A., and Turner, D. H. (1992). Dynamics of ribozyme binding of substrate revealed by fluorescence-detected stopped-flow methods. *Science* 258, 1355-1358.
- Bevilacqua, P. C., and Turner, D. H. (1991). Comparison of binding of mixed ribose-deoxyribose analogues of CUCU to a ribozyme and to GGAGAA by equilibrium dialysis: evidence for ribozyme specific interactions with 2' OH groups. *Biochemistry* 30, 10632-10640.
- Biou, V., Yaremchuk, A., Tukalo, M., and Cusack, S. (1994). The 2.9 Å crystal structure of *T. thermophilus* seryl-tRNA synthetase complexed with tRNA^{Ser}. *Science* 263, 1404-1410.
- Brehm, S. L., and Cech, T. R. (1983). Fate of an intervening sequence ribonucleic acid: excision and cyclization of the *Tetrahymena* ribosomal ribonucleic acid intervening sequence *in vivo*. *Biochemistry* 22, 2390-2397.
- Brierley, I., Digard, P., and Inglis, S. C. (1989). Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell* 57, 537-547.
- Brown, J. W., Haas, E. S., and Pace, N. R. (1993). Characterization of ribonuclease P RNAs from thermophilic bacteria. *Nucleic Acids Res.* 21, 671-679.
- 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, 7217-7221.
- Burke, J. M., Esherick, J. S., Burfeind, W. R., and King, J. C. (1990). A 3' splice site-binding sequence in the catalytic core of a group I intron. *Nature* 344, 80-82.
- 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, 167-176.
- Caponigro, G., and Parker, R. (1995). Multiple functions for the poly(A)-binding protein in mRNA decapping and deadenylation in yeast. *Genes Dev.* 9, 2421-2432.
- Cavarelli, J., Rees, B., Ruff, M., Thierry, J.-C., and Moras, D. (1993). Yeast tRNA^{Asp} recognition by its cognate class II aminoacyl-tRNA synthetase. *Nature* 362, 181-184.

- Cech, T. R. (1986). The generality of self-splicing RNA: relationship to nuclear mRNA splicing. *Cell* **44**, 207-210.
- Cech, T. R. (1993). Structure and mechanism of the large catalytic RNAs: group I and group II introns and ribonuclease P. In *The RNA World*, Gesteland, R. F. and Atkins, J. F., eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 239-269.
- Cech, T. R., Damberger, S. H., and Gutell, R. R. (1994). Representation of the secondary and tertiary structure of group I introns. *Nature Struct. Biol.* **1**, 273-280.
- Cech, T. R., and Rio, D. C. (1979). Localization of transcribed regions on extrachromosomal ribosomal RNA genes of *Tetrahymena thermophila* by R-loop mapping. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 5051-5055.
- 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**, 487-496.
- Celander, D. W., and Cech, T. R. (1991). Visualizing the higher order folding of a catalytic RNA molecule. *Science* **251**, 401-407.
- Chastain, M., and Tinoco, I., Jr. (1992). A base-triple structural domain in RNA. *Biochemistry* **31**, 12733-12741.
- Chastain, M., and Tinoco, I., Jr. (1993). Nucleoside triples from the group I intron. *Biochemistry* **32**, 14220-14228.
- Chen, X., Woodson, S. A., Burrows, C. J., and Rokita, S. E. (1993). A highly sensitive probe for guanosine N7 in folded structures of RNA: application to tRNA^{Phe} and *Tetrahymena* group I intron. *Biochemistry* **32**, 7610-7616.
- Christensen, H., and Pain, R. H. (1994). The contribution of the molten globule model. In *Mechanisms of Protein Folding*, Pain, R. H., ed. (New York: Oxford University Press), pp. 55-79.
- Christian, E. L., and Yarus, M. (1992). Analysis of the role of phosphate oxygens in the group I intron from *Tetrahymena*. *J. Mol. Biol.* **228**, 743-758.
- Christian, E. L., and Yarus, M. (1993). Metal coordination sites that contribute to structure and catalysis in the group I intron from *Tetrahymena*. *Biochemistry* **32**, 4475-4480.
- Cload, S. T., Richardson, P. L., Huang, Y.-H., and Schepartz, A. (1993). Kinetic and thermodynamic analysis of RNA binding by tethered oligonucleotide probes: alternative structures and conformational changes. *J. Am. Chem. Soc.* **115**, 5005-5014.

- Coetzee, T., Herschlag, D., and Belfort, M. (1994). *Escherichia coli* proteins, including ribosomal protein S12, facilitate *in vitro* splicing of phage T4 intron by acting as RNA chaperones. *Genes Dev.* 8, 1575-1588.
- Cole, P. E., and Crothers, D. M. (1972). Conformational changes of transfer ribonucleic acid. Relaxation kinetics of the early melting transition of methionine transfer ribonucleic acid. *Biochemistry* 11, 4368-4374.
- Cole, P. E., Yang, S. K., and Crothers, D. M. (1972). Conformational changes of transfer ribonucleic acid. Equilibrium phase diagrams. *Biochemistry* 11, 4358-4368.
- Couture, S., Ellington, A. D., Gerber, A. S., Cherry, J. M., Doudna, J. A., Green, R., Hanna, M., Pace, U., Rajagopal, J., and Szostak, J. W. (1990). Mutational analysis of conserved nucleotides in a self-splicing group I intron. *J. Mol. Biol.* 215, 345-358.
- Crothers, D. M., Cole, P. E., Hilbers, C. W., and Shulman, R. G. (1974). The molecular mechanism of thermal unfolding of *Escherichia coli* formylmethionine transfer RNA. *J. Mol. Biol.* 87, 63-88.
- Dai, X., De Mesmaeker, A., and Joyce, G. F. (1995). Cleavage of an amide bond by a ribozyme. *Science* 267, 237-240.
- Darr, S. C., Brown, J. W., and Pace, N. R. (1992). The varieties of ribonuclease P. *TIBS* 17, 178-182.
- Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A., and Scazzocchio, C. (1982). Making ends meet: a model for RNA splicing in fungal mitochondria. *Nature* 300, 719-724.
- De Rijk, P., Van de Peer, Y., Chapelle, S., and De Wachter, R. (1994). Database on the structure of large ribosomal subunit RNA. *Nucleic Acids Res.* 22, 3495-3501.
- Decker, C. J., and Parker, R. (1993). A turnover pathway for both stable and unstable mRNAs in yeast: evidence for a requirement for deadenylation. *Genes Dev.* 7, 1632-1643.
- Decker, C. J., and Parker, R. (1995). Diversity of cytoplasmic functions for the 3' untranslated region of eukaryotic transcripts. *Curr. Opin. Cell Biol.* 7, 386-392.
- Dobson, C. M. (1995). Finding the right fold. *Nature Struct. Biol.* 2, 513-517.
- Dobson, C. M., Evans, P. A., and Radford, S. E. (1994). Understanding how proteins fold: the lysozyme story so far. *TIBS* 19, 31-37.
- Doudna, J. A., and Cech, T. R. (1995). Self-assembly of a group I intron active site from its component tertiary structural domains. *RNA* 1, 36-45.

- Doudna, J. A., and Szostak, J. W. (1989). Miniribozymes, small derivatives of the *sunY* intron, are catalytically active. *Mol. Cell. Biol.* **9**, 5480-5483.
- Downs, W. D., and Cech, T. R. (1990). An ultraviolet-inducible adenosine-adenosine cross-link reflects the catalytic structure of the *Tetrahymena* ribozyme. *Biochemistry* **29**, 5605-5613.
- Eaton, B. E., Gold, L., and Zichi, D. A. (1995). Let's get specific: the relationship between specificity and affinity. *Chemistry & Biology* **2**, 633-638.
- Emerick, V. L., and Woodson, S. A. (1993). Self-splicing of the *Tetrahymena* pre-rRNA is decreased by misfolding during transcription. *Biochemistry* **32**, 14062-14067.
- Flor, P. J., Flanagan, J. B., and Cech, T. R. (1989). A conserved base pair within helix P4 of the *Tetrahymena* ribozyme helps to form the tertiary structure required for self-splicing. *EMBO J.* **8**, 3391-3399.
- Frankel, A. D. (1992). Activation of HIV transcription by Tat. *Curr. Opin. Gen. Dev.* **2**, 293-298.
- Gesteland, R. F., and Atkins, J. F. (1993). *The RNA World* (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
- Giegé, R., Puglisi, J. D., and Florentz, C. (1993). tRNA structure and aminoacylation efficiency. *Prog. Nuc. Acids Res. Mol. Biol.* **45**, 129-206.
- Gilley, D., Lee, M. S., and Blackburn, E. H. (1995). Altering specific telomerase RNA template residues affects active site function. *Genes Dev.* **9**, 2214-2226.
- Goguel, V., and Rosbash, M. (1993). Splice-site choice and splicing efficiency are positively influenced by pre-mRNA intramolecular base pairing in yeast. *Cell* **72**, 893-901.
- Gold, L., Allen, P., Binkley, J., Brown, D., Schneider, D., Eddy, S., Tuerk, C., Green, L., MacDougal, S., and Tasset, D. (1993). RNA: the shape of things to come. In *The RNA World*, Gesteland, R. F. and Atkins, J. F., eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 497-507.
- Gold, L., Polisky, B., Uhlenbeck, O., and Yarus, M. (1995). Diversity of oligonucleotide functions. *Annu. Rev. Biochem.* **64**, 763-797.
- Gralla, J., and Crothers, D. M. (1973). Free energy of imperfect nucleic acid helices. II. Small hairpin loops. *J. Mol. Biol.* **73**, 497-511.
- Green, R., and Szostak, J. W. (1994). *In vitro* genetic analysis of the hinge region between helical elements P5-P4-P6 and P7-P3-P8 in the *sunY* group I self-splicing intron. *J. Mol. Biol.* **235**, 140-155.
- Greenblatt, J., Nodwell, J. R., and Mason, S. W. (1993). Transcriptional antitermination. *Nature* **364**, 401-406.

- Griffin, E. A., Jr., Qin, Z., Michels, W. J., Jr., and Pyle, A. M. (1995). Group II intron ribozymes that cleave DNA and RNA linkages with similar efficiency, and lack contacts with substrate 2'-hydroxyl groups. *Chemistry & Biology* 2, 761-770.
- Grosshans, C. A., and Cech, T. R. (1989). Metal ion requirements for sequence specific endoribonuclease activity of the *Tetrahymena* ribozyme. *Biochemistry* 28, 6888-6894.
- Guerrier-Takada, C., and Altman, S. (1984). Catalytic activity of an RNA molecule prepared by transcription *in vitro*. *Science* 223, 285-286.
- Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35, 849-857.
- Guo, Q., and Lambowitz, A. M. (1992). A tyrosyl-tRNA synthetase binds specifically to the group I intron catalytic core. *Genes Dev.* 6, 1357-1372.
- Haas, E. S., Brown, J. W., Pitulle, C., and Pace, N. R. (1994). Further perspective on the catalytic core and secondary structure of ribonuclease P RNA. *Proc. Natl. Acad. Sci. U.S.A.* 91, 2527-2531.
- Harris, M. E., Nolan, J. M., Malhatra, A., Brown, J. W., Harvey, S. C., and Pace, N. R. (1994). Use of photoaffinity crosslinking and molecular modeling to analyze the global architecture of ribonuclease P RNA. *EMBO J.* 13, 3953-3963.
- Harris, M. E., and Pace, N. R. (1995). Identification of phosphates involved in catalysis by the ribozyme RNase P RNA. *RNA* 1, 210-218.
- Held, W. A., and Nomura, M. (1973). Rate-determining step in the reconstitution of *Escherichia coli* 30S ribosomal subunits. *Biochemistry* 12, 3273-3281.
- Herschlag, D. (1991). Implications of ribozyme kinetics for targeting the cleavage of specific RNA molecules *in vivo*: more isn't always better. *Proc. Natl. Acad. Sci. U.S.A.* 88, 6921-6925.
- Herschlag, D. (1992). Evidence for processivity and two-step binding of the RNA substrate from studies of J1/2 mutants of the *Tetrahymena* ribozyme. *Biochemistry* 31, 1386-1399.
- Herschlag, D. (1995). RNA chaperones and the RNA folding problem. *J. Biol. Chem.* 270, 20871-20874.
- Herschlag, D., and Cech, T. R. (1990a). Catalysis of RNA cleavage by the *Tetrahymena thermophila* ribozyme. 1. Kinetic description of the reaction of an RNA substrate complementary to the active site. *Biochemistry* 29, 10159-10171.

- Herschlag, D., and Cech, T. R. (1990b). Catalysis of RNA cleavage by the *Tetrahymena thermophila* ribozyme. 2. Kinetic description of the reaction of an RNA substrate that forms a mismatch at the active site. *Biochemistry* 29, 10172-10180.
- Herschlag, D., Eckstein, F., and Cech, T. R. (1993). Contributions of 2'-hydroxyl groups of the RNA substrate to binding and catalysis by the *Tetrahymena* ribozyme. An energetic picture of an active site composed of RNA. *Biochemistry* 32, 8299-8311.
- Herschlag, D., and Khosla, M. (1994). Comparison of pH dependencies of the *Tetrahymena* ribozyme reaction with RNA 2'-substituted and phosphorothioate substrates reveals a rate-limiting conformational step. *Biochemistry* 33, 5291-5297.
- Herschlag, D., Khosla, M., Tsuchihashi, Z., and Karpel, R. L. (1994). An RNA chaperone activity of non-specific RNA binding proteins in hammerhead ribozyme catalysis. *EMBO J.* 13, 2913-2924.
- 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. U.S.A.* 88, 11105-11109.
- Higuchi, R. (1990). Recombinant PCR. In *PCR Protocols: A Guide to Methods and Applications*, Innis, M. A. et al., eds. (New York: Academic Press Inc.), pp. 177-183.
- Illangasekare, M., Sanchez, G., Nickles, T., and Yarus, M. (1995). Aminoacyl-RNA synthesis catalyzed by an RNA. *Science* 267, 643-647.
- Jaeger, J. A., Zuker, M., and Turner, D. H. (1990). Melting and chemical modification of a cyclized self-splicing group I intron: similarity of structures in 1M Na⁺, in 10mM Mg²⁺, and in the presence of substrate. *Biochemistry* 29, 10147-10158.
- Jaeger, L., Westhof, E., and Michel, F. (1991). Function of P11, a tertiary base pairing in self-splicing introns of subgroup IA. *J. Mol. Biol.* 221, 1153-1164.
- Jaeger, L., Westhof, E., and Michel, F. (1993). Monitoring of the cooperative unfolding of the *sunY* group I intron of bacteriophage T4. The active form of the *sunY* ribozyme is stabilized by multiple interactions with 3' terminal intron components. *J. Mol. Biol.* 234, 331-346.
- Jennings, P. A., and Wright, P. E. (1993). Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science* 262, 892-896.
- Joyce, G. F. (1994). *In vitro* evolution of nucleic acids. *Curr. Opin. Struct. Biol.* 4, 331-336.
- Kiefhaber, T. (1995). Kinetic traps in lysozyme folding. *Proc. Natl. Acad. Sci. U.S.A.* 92, 9029-9033.

- Knitt, D. S., Narlikar, G. I., and Herschlag, D. (1994). Dissection of the role of the conserved G•U pair in group I RNA self-splicing. *Biochemistry* 33, 13864-13879.
- Kotik, M., Radford, S. E., and Dobson, C. M. (1995). Comparison of the refolding of hen lysozyme from dimethyl sulfoxide and guanidinium chloride. *Biochemistry* 34, 1714-1724.
- Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* 31, 147-157.
- Kufel, J., and Kirsebom, L. A. (1994). Cleavage site selection by M1 RNA, the catalytic subunit of *Escherichia coli* RNase P, is influenced by pH. *J. Mol. Biol.* 244, 511-521.
- Kuge, H., and Richter, J. D. (1995). Cytoplasmic 3' poly(A) addition induces 5' cap ribose methylation: implications for translational control of maternal mRNA. *EMBO J.* 14, 6301-6310.
- Laggerbauer, B., Murphy, F. L., and Cech, T. R. (1994). Two major tertiary folding transitions of the *Tetrahymena* catalytic RNA. *EMBO J.* 13, 2669-2676.
- Latham, J. A., and Cech, T. R. (1989). Defining the inside and outside of a catalytic RNA molecule. *Science* 245, 276-282.
- Lazinski, D. W., and Taylor, J. M. (1995). Regulation of the hepatitis delta virus ribozymes: to cleave or not to cleave? *RNA* 1, 225-233.
- LeCuyer, K. A., and Crothers, D. M. (1994). Kinetics of an RNA conformational switch. *Proc. Natl. Acad. Sci. U.S.A.* 91, 3373-3377.
- Lehman, N., and Joyce, G. F. (1993). Evolution *in vitro* of an RNA enzyme with altered metal dependence. *Nature* 361, 182-185.
- Lewin, A. S., Thomas, J., Jr., and Tirupati, H. K. (1995). Cotranscriptional splicing of a group I intron is facilitated by the CBP2 protein. *Mol. Cell. Biol.* 15, 6971-6978.
- Libri, D., Stutz, F., McCarthy, T., and Rosbash, M. (1995). RNA structural patterns and splicing: molecular basis for an RNA-based enhancer. *RNA* 1, 425-436.
- Litvak, S., Sarih-Cottin, L., Fournier, M., Andreola, M., and Tarrayo-Litvak, L. (1994). Priming of HIV replication by tRNA^{Lys3}: role of reverse transcriptase. *TIBS* 19, 114-118.
- Loh, S. N., Kay, M. S., and Baldwin, R. L. (1995). Structure and stability of a second molten globule intermediate in the apomyoglobin folding pathway. *Proc. Natl. Acad. Sci. U.S.A.* 92, 5446-5450.

- Lorsch, J. R., and Szostak, J. W. (1994). *In vitro* evolution of new ribozymes with polynucleotide kinase activity. *Nature* 371, 31-36.
- Lynch, D. C., and Schimmel, P. R. (1974). Cooperative binding of magnesium to transfer ribonucleic acid studied by a fluorescent probe. *Biochemistry* 13, 1841-1852.
- Madhani, H. D., and Guthrie, C. (1994). Dynamic RNA-RNA interactions in the spliceosome. *Annu. Rev. Genet.* 28, 1-26.
- Malhotra, A., and Harvey, S. C. (1994). A quantitative model of the *Escherichia coli* 16S RNA in the 30S ribosomal subunit. *J. Mol. Biol.* 240, 308-340.
- Marino, J. P., Gregorian, R. S., Jr., Csankovszki, G., and Crothers, D. M. (1995). Bent helix formation between RNA hairpins with complementary loops. *Science* 268, 1448-1454.
- McCarthy, J. E. G., and Kollmus, H. (1995). Cytoplasmic mRNA-protein interactions in eukaryotic gene expression. *TIBS* 20, 191-197.
- McConnell, T. S., Cech, T. R., and Herschlag, D. (1993). Guanosine binding to the *Tetrahymena* ribozyme: thermodynamic coupling with oligonucleotide binding. *Proc. Natl. Acad. Sci. U.S.A.* 90, 8362-8366.
- McSwiggen, J. A., and Cech, T. R. (1989). Stereochemistry of RNA cleavage by the *Tetrahymena* ribozyme and evidence that the chemical step is not rate-limiting. *Science* 244, 679-683.
- 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, 578-580.
- Michel, F., Hanna, M., Green, R., Bartel, D. P., and Szostak, J. W. (1989a). The guanosine binding site of the *Tetrahymena* ribozyme. *Nature* 342, 391-395.
- Michel, F., Umesono, K., and Ozeki, H. (1989b). Comparative and functional anatomy of group II catalytic introns-a review. *Gene* 82, 5-30.
- 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, 585-610.
- Moazed, D., Stern, S., and Noller, H. F. (1986). Rapid chemical probing of conformation in 16S ribosomal RNA and 30S ribosomal subunits using primer extension. *J. Mol. Biol.* 187, 399-416.
- 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, 147-150.

- 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, 483-494.
- Moore, M. J., Query, C. C., and Sharp, P. A. (1993). Splicing of precursors to mRNA by the spliceosome. In *The RNA World*, Gesteland, R. F. and Atkins, J. F., eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 303-357.
- Moore, P. B. (1993). Ribosomes and the RNA world. In *The RNA World*, Gesteland, R. F. and Atkins, J. F., eds. (Cold Spring Harbor, NY: Cold Spring Harbor Press), pp. 119-135.
- Murphy, F. L., and Cech, T. R. (1989). Alteration of substrate specificity for the endoribonucleolytic cleavage of RNA by the *Tetrahymena* ribozyme. *Proc. Natl. Acad. Sci. U.S.A.* 86, 9218-9222.
- Murphy, F. L., and Cech, T. R. (1993). An independently folding domain of RNA tertiary structure within the *Tetrahymena* ribozyme. *Biochemistry* 32, 5291-5300.
- Murphy, F. L., and Cech, T. R. (1994). GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. *J. Mol. Biol.* 236, 49-63.
- Murphy, F. L., Wang, Y.-H., Griffith, J. D., and Cech, T. R. (1994). Coaxially stacked RNA helices in the catalytic center of the *Tetrahymena* ribozyme. *Science* 265, 1709-1712.
- Nagai, K., Oubridge, C., Ito, N., Avis, J., and Evans, P. (1995). The RNP domain: a sequence-specific RNA-binding domain involved in processing and transport of RNA. *TIBS* 20, 235-240.
- Nakamura, T. M., Wang, Y.-H., Zaug, A. J., Griffith, J. D., and Cech, T. R. (1995). Relative orientation of RNA helices in a group I ribozyme determined by helix extension electron microscopy. *EMBO J.* 14, 4849-4859.
- Narlikar, G., Gopalakrishnan, V., McConnell, T. S., Usman, N., and Herschlag, D. (1995). Use of binding energy by an RNA enzyme for catalysis by positioning and substrate destabilization. *Proc. Natl. Acad. Sci. U.S.A.* 92, 3668-3672.
- Neefs, J.-M., Van de Peer, Y., De Rijk, P., Chapelle, S., and De Wachter, R. (1993). Compilation of small ribosomal subunit RNA structures. *Nucleic Acids Res.* 21, 3025-3049.
- Nissen, P., Kjeldgaard, M., Thirup, S., Polekhina, G., Reshetnikova, L., Clark, B. F. C., and Nyborg, J. (1995). Crystal structure of the ternary complex Phe-tRNA^{Phe}, EF-Tu, and a GTP analog. *Science* 270, 1464-1472.
- Noller, H. F. (1991). Ribosomal RNA and translation. *Annu. Rev. Biochem.* 60, 191-227.

- Noller, H. F. (1993). On the origin of the ribosome: coevolution of subdomains of tRNA and rRNA. In *The RNA World*, Gesteland, R. F. and Atkins, J. F., eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 137-156.
- Noller, H. F., Hoffarth, V., and Zimniak, L. (1992). Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* 256, 1416-1419.
- Padgett, R. A., Podar, M., Boulanger, S. C., and Perlman, P. S. (1994). The stereochemical course of group II intron self-splicing. *Science* 266, 1685-1688.
- Pan, T. (1995). Higher order folding and domain analysis of the ribozyme from *Bacillus subtilis* ribonuclease P. *Biochemistry* 34, 902-909.
- Pan, T., Long, D. M., and Uhlenbeck, O. C. (1993). Divalent metal ions in RNA folding and catalysis. In *The RNA World*, Gesteland, R. F. and Atkins, J. F., eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 271-302.
- Parker, R., and Siliciano, P. G. (1993). Evidence for an essential non-Watson-Crick interaction between the first and last nucleotide of a nuclear pre-mRNA intron. *Nature* 361, 660-662.
- Peebles, C. L., Perlman, P. S., Mecklenburg, K. L., Petrillo, M. L., Tabor, J. H., Jarrell, K. A., and Cheng, H.-L. (1986). A self-splicing RNA excises an intron lariat. *Cell* 44, 213-223.
- Piccirilli, J. A., McConnell, T. S., Zaug, A. J., Noller, H. F., and Cech, T. R. (1992). Aminoacyl esterase activity of the *Tetrahymena* ribozyme. *Science* 256, 1420-1424.
- Piccirilli, J. A., Vyle, J. S., Caruthers, M. H., and Cech, T. R. (1993). Metal ion catalysis in the *Tetrahymena* ribozyme reaction. *Nature* 361, 85-88.
- Pley, H., Flaherty, K. M., and McKay, D. B. (1994). Three-dimensional structure of a hammerhead ribozyme. *Nature* 372, 68-74.
- Pörschke, D., and Eigen, M. (1971). Cooperative non-enzymic base recognition III. Kinetics of the helix-coil transition of the oligoribouridylic-oligoriboadenylic acid system and of oligoriboadenylic acid alone at acidic pH. *J. Mol. Biol.* 62, 361-381.
- Portman, D. S., and Dreyfuss, G. (1994). RNA annealing activities in HeLa nuclei. *EMBO J.* 13, 213-221.
- Powers, T., Daubresse, G., and Noller, H. F. (1993). Dynamics of *in vitro* assembly of 16S rRNA into 30S ribosomal subunits. *J. Mol. Biol.* 232, 362-374.
- Pyle, A. M. (1993). Ribozymes: a distinct class of metalloenzymes. *Science* 261, 709-714.

- Pyle, A. M., and Cech, T. R. (1991). Ribozyme recognition of RNA by tertiary interactions with specific ribose 2'-OH groups. *Nature* 350, 628-631.
- Pyle, A. M., and Green, J. B. (1995). RNA folding. *Curr. Opin. Struct. Biol.* 5, 303-310.
- Pyle, A. M., McSwiggen, J. A., and Cech, T. R. (1990). Direct measurement of oligonucleotide substrate binding to wild-type and mutant ribozymes from *Tetrahymena*. *Proc. Natl. Acad. Sci. U.S.A.* 87, 8187-8191.
- Pyle, A. M., Moran, S., Strobel, S. A., Chapman, T., Turner, D. H., and Cech, T. R. (1994). Replacement of the conserved G•U with a G•C pair at the cleavage site of the *Tetrahymena* ribozyme decreases binding, reactivity and fidelity. *Biochemistry* 33, 13856-13863.
- Pyle, A. M., Murphy, F. L., and Cech, T. R. (1992). RNA substrate binding site in the catalytic core of the *Tetrahymena* ribozyme. *Nature* 358, 123-128.
- Reich, C., Olson, G. J., Pace, B., and Pace, N. R. (1988). Role of the protein moiety of ribonuclease P, a ribonucleoprotein enzyme. *Science* 239, 178-181.
- Ringquist, S., Schneider, D., Gibson, T., Baron, C., Böck, A., and Gold, L. (1993). Recognition of the mRNA selenocysteine insertion sequence by the specialized translational elongation factor SELB. *Genes Dev.* 8, 376-385.
- Rould, M. A., Perona, J. J., and Steitz, T. A. (1991). Structural basis of anticodon loop recognition by glutamyl-tRNA synthetase. *Nature* 352, 213-217.
- Rould, M. A., Perona, J. S., Söll, D., and Steitz, T. A. (1989). Structure of *E. coli* glutamyl-tRNA synthetase complexed with tRNA^{Glu} and ATP at 2.8 Å resolution. *Science* 246, 1135-1142.
- Rudinger, J., Puglisi, J. D., Pütz, J., Schatz, D., Eckstein, F., Florentz, C., and Giegé, R. (1992). Determinant nucleotides of yeast tRNA^{Asp} interact directly with aspartyl-tRNA synthetase. *Proc. Natl. Acad. Sci. U.S.A.* 89, 5882-5886.
- 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, 1275-1287.
- Scheffler, I. E., Elson, E. L., and Baldwin, R. L. (1968). Helix formation by dAT oligomers I. Hairpin and straight-chain helices. *J. Mol. Biol.* 36, 291-304.
- Schimmel, P. R., and Redfield, A. G. (1980). Transfer RNA in solution: selected topics. *Annu. Rev. Biophys. Bioeng.* 9, 181-221.

- Schmelzer, C., and Schweyen, R. J. (1986). Self-splicing of group II introns *in vitro*: mapping of the branch point and mutational inhibition of lariat formation. *Cell* **46**, 557-565.
- Scott, W. J., Finch, J. T., and Klug, A. (1995). The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell* **81**, 991-1002.
- Serrano, L., Matouschek, A., and Fersht, A. R. (1992). The folding of an enzyme VI. The folding pathway of barnase: comparison with theoretical models. *J. Mol. Biol.* **224**, 847-859.
- Shen, L. X., Cai, Z., and Tinoco, I., Jr. (1995). RNA structure at high resolution. *FASEB J.* **9**, 1023-1033.
- Sieber, G., and Nierhaus, K. H. (1978). Kinetic and thermodynamic parameters of the assembly *in vitro* of the large subunit from *Escherichia coli* ribosomes. *Biochemistry* **17**, 3505-3511.
- Söll, D. (1993). Transfer RNA: an RNA for all seasons. In *The RNA World*, Gesteland, R. F. and Atkins, J. F., eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 157-184.
- Sontheimer, E. J., and Steitz, J. A. (1993). The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science* **262**, 1989-1996.
- Sosnick, T. R., Mayne, L., Hiller, R., and Englander, S. W. (1994). The barriers in protein folding. *Nature Struct. Biol.* **1**, 149-156.
- Stein, A., and Crothers, D. M. (1976a). Conformational changes of transfer RNA. The role of magnesium(II). *Biochemistry* **15**, 160-168.
- Stein, A., and Crothers, D. M. (1976b). Equilibrium binding of magnesium(II) by *Escherichia coli* tRNA^{fMet}. *Biochemistry* **15**, 157-160.
- Steitz, T. A., and Steitz, J. A. (1993). A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6498-6502.
- Stern, S., Powers, T., Changchien, L.-M., and Noller, H. F. (1989). RNA-protein interactions in 30S ribosomal subunits: folding and function of 16S rRNA. *Science* **244**, 783-790.
- Stern, S., Weiser, B., and Noller, H. F. (1988). Model for the three-dimensional folding of 16S ribosomal RNA. *J. Mol. Biol.* **204**, 447-481.
- Strobel, S. A., and Cech, T. R. (1993). Tertiary interactions with the internal guide sequence mediate docking of the P1 helix into the catalytic core of the *Tetrahymena* ribozyme. *Biochemistry* **32**, 13593-13604.

- Strobel, S. A., and Cech, T. R. (1994). Translocation of an RNA duplex on a ribozyme. *Nature Struct. Biol.* **1**, 13-17.
- Strobel, S. A., and Cech, T. R. (1995). Minor groove recognition of the conserved G•U pair at the *Tetrahymena* ribozyme reaction site. *Science* **267**, 675-679.
- Strobel, S. A., and Cech, T. R. (1996). Exocyclic amine of the conserved G•U pair at the cleavage site of the *Tetrahymena* ribozyme contributes to 5'-splice site selection and transition state stabilization. *Biochemistry* **35**, 1201-1211.
- Symons, R. H. (1992). Small catalytic RNAs. *Annu. Rev. Biochem.* **61**, 641-671.
- Szewczak, A. A., and Moore, P. B. (1995). The sarcin/ricin loop, a modular RNA. *J. Mol. Biol.* **247**, 81-98.
- Szostak, J. W., and Ellington, A. D. (1993). *In vitro* selection of functional RNA sequences. In *The RNA World*, Gesteland, R. F. and Atkins, J. F., eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 511-533.
- Tallsjö, A., and Kirsebom, L. (1993). Product release is a rate-limiting step during cleavage by the catalytic RNA subunit of *Escherichia coli* RNase P. *Nucleic Acids Res.* **21**, 51-57.
- Tranguch, A. J., Kindelberger, D. W., Rohlman, C. E., Lee, J.-Y., and Engelke, D. R. (1994). Structure-sensitive RNA footprinting of yeast nuclear ribonuclease P. *Biochemistry* **33**, 1778-1787.
- Traub, P., and Nomura, M. (1969). Structure and function of *Escherichia coli* ribosomes VI. Mechanism of assembly of 30S ribosomes studied *in vitro*. *J. Mol. Biol.* **40**, 391-413.
- Tsuchihashi, Z., Khosla, M., and Herschlag, D. (1993). Protein enhancement of hammerhead ribozyme catalysis. *Science* **262**, 99-102.
- Uhlenbeck, O. C., Baller, J., and Doty, P. (1970). Complementary oligonucleotide binding to the anticodon loop of fMet-transfer RNA. *Nature* **225**, 508-510.
- Valegård, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J., and Liljas, L. (1994). Crystal structure of an RNA bacteriophage coat protein-operator complex. *Nature* **371**, 623-626.
- van der Horst, G., Christian, A., and Inoue, T. (1991). Reconstitution of a group I intron self-splicing reaction with an activator RNA. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 184-188.
- van der Veen, R., Arnberg, A. C., van der Horst, G., Bonen, L., Tabak, H. F., and Grivell, L. A. (1986). Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing *in vitro*. *Cell* **44**, 225-234.

- Varani, G. (1995). Exceptionally stable nucleic acid hairpins. *Annu. Rev. Biophys. Biomol. Struct.* 24, 379-404.
- von Hippel, P. H., Bear, D. G., Morgan, W. D., and McSwiggen, J. A. (1984). Protein-nucleic acid interactions in transcription: a molecular analysis. *Annu. Rev. Biochem.* 53, 389-446.
- Wang, J.-F., and Cech, T. R. (1992). Tertiary structure around the guanosine-binding site of the *Tetrahymena* ribozyme. *Science* 256, 526-529.
- Wang, J.-F., and Cech, T. R. (1994). Metal ion dependence of active-site structure of the *Tetrahymena* ribozyme revealed by site-specific photo-cross-linking. *J. Am. Chem. Soc.* 116, 4178-4182.
- Wang, J.-F., Downs, W. D., and Cech, T. R. (1993). Movement of the guide sequence during RNA catalysis by a group I ribozyme. *Science* 260, 504-508.
- Wang, Y.-H., Murphy, F. L., Cech, T. R., and Griffith, J. D. (1994). Visualization of a tertiary structural domain of the *Tetrahymena* group I intron by electron microscopy. *J. Mol. Biol.* 236, 64-71.
- Waring, R. B., Towner, P., Minter, S. J., and Davies, R. W. (1986). Splice-site selection by a self-splicing RNA of *Tetrahymena*. *Nature* 321, 133-139.
- 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, 221-230.
- Weeks, K. M., and Cech, T. R. (1996). Assembly of a ribonucleoprotein catalyst by tertiary structure capture. *Science* 271, 345-348.
- Weller, J. W., and Hill, W. E. (1992). Probing dynamic changes in rRNA conformation in the 30S subunit of the *Escherichia coli* ribosome. *Biochemistry* 31, 2748-2757.
- Westhof, E., and Altman, S. (1994). Three-dimensional working model of M1 RNA, the catalytic RNA subunit of ribonuclease P from *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 91, 5133-5137.
- Westhof, E., and Michel, F. (1992). Some tertiary motifs of RNA foldings. In *Structural Tools for the Analysis of Protein-Nucleic Acid Complexes*, Lilley, D. J. M., Heumann, H. and Suck, D., eds. (Basel: Birkhäuser Verlag), pp. 255-267.
- Williams, K. P., Imahori, H., Fujimoto, D. N., and Inoue, T. (1994). Selection of novel forms of a functional domain within the *Tetrahymena* ribozyme. *Nucleic Acids Res.* 22, 2003-2009.
- 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, 14672-14682.

- Wilson, C., and Szostak, J. W. (1995). *In vitro* evolution of a self-alkylating ribozyme. *Nature* 374, 777-782.
- Wimberly, B., Varani, G., and Tinoco, I., Jr. (1993). The conformation of loop E of eukaryotic 5S ribosomal RNA. *Biochemistry* 32, 1078-1087.
- Woodson, S. A., and Cech, T. R. (1991). Alternative secondary structures in the 5' exon affect both forward and reverse self-splicing of the *Tetrahymena* intervening sequence RNA. *Biochemistry* 30, 2042-2050.
- Woodward, C. K. (1994). Hydrogen exchange rates and protein folding. *Curr. Opin. Struct. Biol.* 4, 112-116.
- Wu, L. C., Peng, Z.-y., and Kim, P. S. (1995). Bipartite structure of the α -lactalbumin molten globule. *Nature Struct. Biol.* 2, 281-286.
- Yang, S. K., and Crothers, D. M. (1972). Conformational changes of transfer ribonucleic acid. Comparison of the early melting transition of two tyrosine-specific transfer ribonucleic acids. *Biochemistry* 11, 4375-4381.
- Yarus, M., Illangsekare, M., and Christian, E. (1991). An axial binding site in the *Tetrahymena* precursor RNA. *J. Mol. Biol.* 222, 995-1012.
- Young, B., Herschlag, D., and Cech, T. R. (1991). Mutations in a nonconserved sequence of the *Tetrahymena* ribozyme increase activity and specificity. *Cell* 67, 1007-1019.
- Yu, Y.-T., Maroney, P. A., Darzynkiewicz, E., and Nilsen, T. W. (1995). U6 snRNA function in nuclear pre-mRNA splicing: a phosphorothioate interference analysis of the U6 phosphate backbone. *RNA* 1, 46-54.
- Zakian, V. A. (1995). Telomeres: beginning to understand the end. *Science* 270, 1601-1607.
- Zaug, A. J., Been, M. D., and Cech, T. R. (1986). The *Tetrahymena* ribozyme acts like an RNA restriction endonuclease. *Nature* 324, 429-433.
- Zaug, A. J., and Cech, T. R. (1982). The intervening sequence excised from the ribosomal RNA precursor of *Tetrahymena* contains a 5'-terminal guanosine residue not encoded by the DNA. *Nucleic Acids Res.* 10, 2823-2838.
- Zaug, A. J., Grosshans, C. A., and Cech, T. R. (1988). Sequence-specific endoribonuclease activity of the *Tetrahymena* ribozyme: enhanced cleavage of certain oligonucleotide substrates that form mismatched ribozyme-substrate complexes. *Biochemistry* 27, 8924-8931.
- Zhang, A., Derbyshire, V., Galloway Salvo, J. L., and Belfort, M. (1995a). *Escherichia coli* protein StpA stimulates self-splicing by promoting RNA assembly *in vitro*. *RNA* 1, 783-793.

References

- Zhang, F., Ramsay, E. S., and Woodson, S. A. (1995b). *In vivo* facilitation of *Tetrahymena* group I intron splicing in *Escherichia coli* pre-ribosomal RNA. *RNA* 1, 284-292.
- Zimmerly, S., Guo, H., Eskes, R., Yang, J., Perlman, P. S., and Lambowitz, A. M. (1995). A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. *Cell* 83, 529-538.

BIOGRAPHICAL NOTE

Patrick P. Zarrinkar

Education:

- Sept. 1990-June 1996 Ph.D. in Biochemistry
Massachusetts Institute of Technology
Department of Chemistry
Advisor: Prof. James R. Williamson
- Jan. 1986-May 1990 B.S. in Chemistry
with Honors from S.C. College
Magna Cum Laude
University of South Carolina

Honors/Awards:

- Howard Hughes Medical Institute Predoctoral Fellowship in the Biological Sciences (1991-1996)
- Phi Beta Kappa
- Outstanding Chemistry Major Award, American Chemical Society, S.C. Section (1988)
- James A. Hicks President's Scholarship, University of South Carolina (1987-1990)

Experience:

- July 1991-June 1996 Predocctoral Fellow
MIT, Dept. of Chemistry
Prof. James R. Williamson
Kinetics of RNA folding
- Sept. 1990-Dec. 1990 Teaching Assistant
MIT, Dept. of Chemistry
General Chem. Lab.
- June 1990-Aug. 1990 Research Assistant
MIT, Division of Toxicology
Prof. John M. Essigmann
Purification and characterization of modified oligonucleotides
- June 1988-May 1990 Undergraduate Research Assistant
University of S.C., Dept. of Chemistry
Prof. Roy E. Wuthier
Fatty acid analysis of biological tissue
- Jan. 1987-May 1988 Undergraduate Research Assistant
University of S.C., Dept. of Chemistry
Prof. Paul E. Peterson
Synthetic organic chemistry

Publications:

Research Articles

- Zarrinkar, P. P., Wang, J. and Williamson, J. R. Slow folding kinetics of RNase P RNA. (submitted).
- Zarrinkar, P. P. and Williamson, J. R. (1996). The kinetic folding pathway of the *Tetrahymena* ribozyme reveals possible similarities between RNA and protein folding. *Nature Struct. Biol.*, (in press).
- Zarrinkar, P. P. and Williamson, J. R. (1996). The P9.1-P9.2 peripheral extension helps guide folding of the *Tetrahymena* ribozyme. *Nucleic Acids Res.* **24**, 854-858.
- Zarrinkar, P. P. and Williamson, J. R. (1994). Kinetic intermediates in RNA folding. *Science* **265**, 918-924.
- Adkisson, H. D. IV., Risener, F. S. Jr., Zarrinkar, P. P., Walla, M. D., Christie, W. W., and Wuthier, R. E. (1991). Unique fatty acid composition of normal cartilage: discovery of high levels of n-9 eicosatrienoic acid and low levels of n-6 polyunsaturated fatty acids. *FASEB J.* **5**, 344-353.

Abstracts

- Zarrinkar, P. P. and Williamson, J. R. (1994). Kinetic intermediates on the folding pathway of a group I ribozyme. *RNA Processing Meeting of the RNA Society*, 509.
- Zarrinkar, P. P. and Williamson, J. R. (1993). A kinetic approach to the study of RNA folding. *Young Scientist's View of Molecular Biotechnology, 2nd International Meeting*.
- Adkisson, H. D. IV., Trancik, T. M., Zarrinkar, P. P., and Wuthier, R. E. (1990). Accumulation of essential fatty acids in hyaline cartilage with aging and osteoarthritis: possible implications in development of arthritis. *Orthop. Trans.* **14**, 423-424.

Personal Data:

- born July 24, 1967 in Zweibrücken, W. Germany
- U.S. citizenship obtained July 7, 1994