ENZYMATIC CLEAVAGE OF RNA BY RNA

Nobel Lecture, December 8, 1989

SIDNEY ALTMAN

Department of Biology, Yale University, New Haven, CT 06520, USA

Introduction

The transfer of genetic information from nucleic acid to protein inside cells can be represented as shown in Fig. 1. This simple scheme reflects accurately the fact that the information contained in the linear arrangement of the subunits of DNA is copied accurately into the linear arrangement of subunits of RNA which, in turn, is translated by machinery inside the cell into proteins, the macromolecules responsible for governing many of the important biochemical processes *in vivo*. The function of the straightforward transfer of information is carried out by a class of molecules called messenger RNAs (mRNAs). The diagram shown does not elaborate on the properties of other RNA molecules that are transcribed from DNA, namely transfer RNA (tRNA) and ribosomal RNA (rRNA) and many other minor species of RNA found *in vivo* that had no identifiable function prior to 1976, nor does it indicate that the information in DNA and RNA can be replicated as daughter DNA and RNA molecules, respectively (see Crick, 1970, for further discussion).

Ribosomes are complexes, which in *Escherichia coli,* are made of about 50 proteins and three RNA molecules. It is on these particles that mRNA directs the synthesis of protein from free amino acids. tRNA molecules (Fig. 2) perform an adaptor function in the sense that they match particular amino acids to a group of three specific nucleotides in the mRNA to be translated and ensure that the growing polypeptide (protein) chain contains the right linear sequence of amino acid subunits. Thus, rRNA and tRNA participate in the process of information transfer inside cells but they clearly do so in a comparatively complex manner. mRNA, understandably,

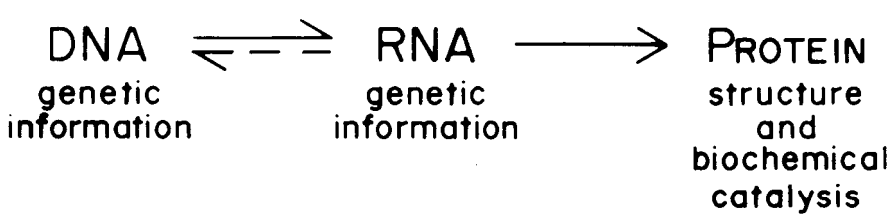

Figure 1. A representation of the flow of information inside cells from DNA to protein. This diagram is not a complete representation of the central dogma (see Crick, 1970).

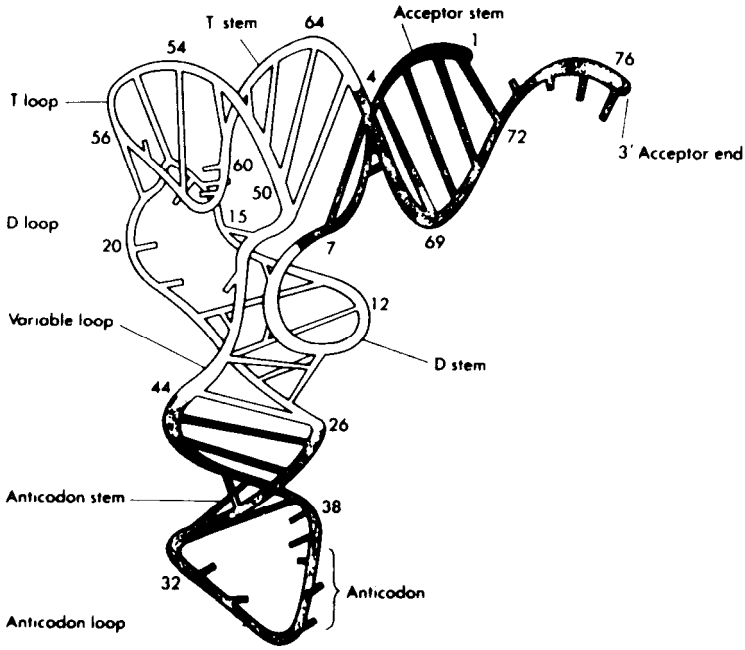

Figure 2. A diagram illustrating the folding of the yeast $tRNA^{Phe}$ molecule. The ribose-phosphate backbone is drawn as a continuous ribbon and internal hydrogen-bonding is indicated by crossbars. Positions of single bases are indicated by bars that are intentionally shortened. The anticodon and acceptor arms are shaded. (Reprinted with permission from Watson, J.D., Hopkins, N.H., Roberts, J.W., Steitz, J.A. and Weiner, A.M., Molecular Biology of the Gene, 4th ed., Benjamin/Cummings, Menlo Park, 1987).

has a very short half-life inside cells but rRNA and tRNA are relatively stable molecules since they are part of the translational machinery that must be used over and over.

My work on RNA began as a study of certain mutants that disrupted the ability of tRNA molecules to function normally during translation (Altman, 1971). This research, in turn, led to the identification of another stable RNA molecule that had, unexpectedly, all the properties of an enzyme (Guerrier-Takada et al., 1983). Aside from its intrinsic interests to students of catalysis and enzymology, our finding of enzymatic activity associated with RNA has stimulated reconsideration of the role of RNA in biochemical systems today (see Cech, 1987, and Altman, 1989, for reviews) as illustrated in Fig. 1, and of the nature of complex biochemical systems eons ago (Darnell and Doolittle, 1986; Westheimer, 1986; Weiner and Maizels, 1987; Joyce, 1989).

As was first pointed out over twenty years ago by Woese (1967), Crick (1968) and Orgel (1968), if RNA can act as a catalyst the origin of the genetic code plays a much less critical role in the early stages of evolution of the first biochemical systems that were capable of replicating themselves. Indeed, the variety of biochemical reactions now known to be governed by RNA, as outlined in Table 1 (Altman, 1989), allows one to consider the

Table 1. *Some Properties of Catalytic RNAs*

RNA	End Groups[a]	Cofactor[b]	Mechanism
1. Group I introns	5'-P, 3′-OH	Yes	Transesterification
2. Group II introns	5′-P, 3′-OH	No	Transesteritication
3. Ml RNA	5′-P, 3′-OH	No	Hydrolysis
4. Viroid/satellite	5′-OH, 2′,3′-cyclic phosphate	No	Transesterification
5. Lead ion/tRNA	5′-OH, 2′3′-cyclic phosphate	NO	Similar to RNAse A

[a]The end groups are those produced during the initial cleavage step of self-splicing reactions or during the usual cleavage reactions of other RNA species.
[b]This column refers to the use of a nucleotide cofactor.

possibility that a large number of diverse enzymatic reactions took place in the absence of protein. To add further substance to these ideas about life on earth over a billion years ago, it is important to understand exactly how catalytic RNA, as we know it, works and what role it plays *in vivo* today. This discussion deals primarily with the catalytic RNA subunit of the enzyme ribonuclease P from *Escherichia coli.*

A BRIEF ACCOUNT OF STUDIES OF RIBONUCLEASE P

Finding the substrate

In October, 1969 I arrived at the MRC laboratory of Molecular Biology in Cambridge, England ostensibly to study the three-dimensional structure of tRNA through the use of physical-chemical methods. On my arrival, Sydney Brenner and Francis Crick informed me that the crystal structure of yeast $tRNA^{Phe}$ had recently been solved (Rim et al., 1974; Robertus et al., 1974) and that there was no further need to engage in the studies originally outlined for me. I was further instructed to get settled, to think about a new problem for a week or two, and then to return for another discussion. Although some of my colleagues remember me as being upset by that conversation with Brenner and Crick, the feeling must have passed quickly because I only recall being presented with a marvelous opportunity to follow my own ideas.

I proposed to make acridine-induced mutants of $tRNA^{Tyr}$ from *E. coli* to determine if altering spatial relationships in tRNA, by deleting or adding a nucleotide to its sequence, would drastically alter the function of the molecule. Since Brenner and John D. Smith and their colleagues (Abelson et al., 1970; Russell et al., 1970; Smith et al., 1971) had just completed a classic series of studies of base-substitution mutants of $tRNA^{Tyr}$, they were not overly excited by the prospect of someone simply producing more mutants. Nevertheless, Brenner and Crick did not prevent me from pushing ahead and John Smith, in time, provided valuable advice about the genetics of the system in use in the laboratory.

The mutants I made lacked the usual function of suppressor tRNAs and

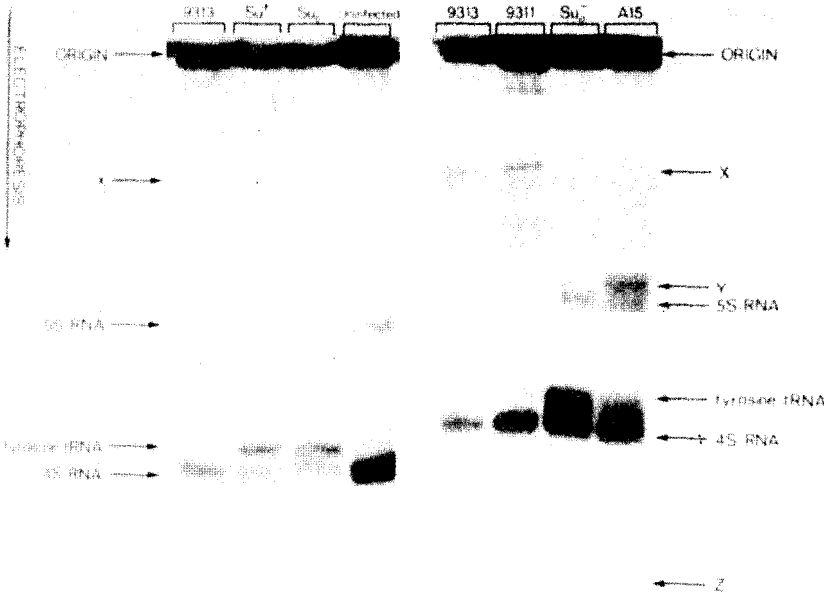

Figure 3. Separation by clectrophoresis of labeled RNA from *E. coli* infected with derivatives of bacteriophage $\phi 80$ carrying various genes for $tRNA^{Tyr}$. The figure shows an autoradiogram of polyacrylamide gels. Experimental details are given in Altman (1971). Each column in the gel patterns is titled according to the $tRNA^{Tyr}$ gene carried by the infecting phage. 9313 and 9311 are acridine-induced mutants of the suppressor $tRNA^{Tyr}su_3^+$. Al5 is a mutant derivative carrying the G15-A15 mutation and su_0- is the wild type $tRNA^{Tyr}$ gene. (Reprinted with permission).

made no mature tRNA *in vivo,* but they reverted at a very high rate (about 1%) to wild type. These properties indicated that there might be an unstable duplication or partial duplication of the gene for tRNA in the DNA that contained the information for the tRNA. Furthermore, it seemed likely that RNA would be transcribed from this mutant gene. I reasoned that if I could isolate the RNA transcript, which had to be unstable since no mature tRNA was made, I might be able to understand the nature of the duplication event.

The simple expedient of quickly pouring an equal volume of phenol into a growing culture of *E. coli* labeled with $^{32}PO_4^{3-}$ enabled me to isolate and characterize not only the transcript of the gene for $tRNA^{Tyr}$ mutated by acridines, but also transcripts of the gene for $tRNA^{Tyr}$ (Fig. 3; Altman, 1971) which had been previously mutated by other means by Brenner, Smith and their colleagues. The ability to isolate these gene transcripts, which contained sequences in addition to the mature tRNA sequences at both ends of the molecules (Fig. 4; Altman and Smith, 1971) and were, therefore, tRNA precursor molecules, depended on the rapid phenol extraction technique and the fact that the mutated molecules were less susceptible to attack by

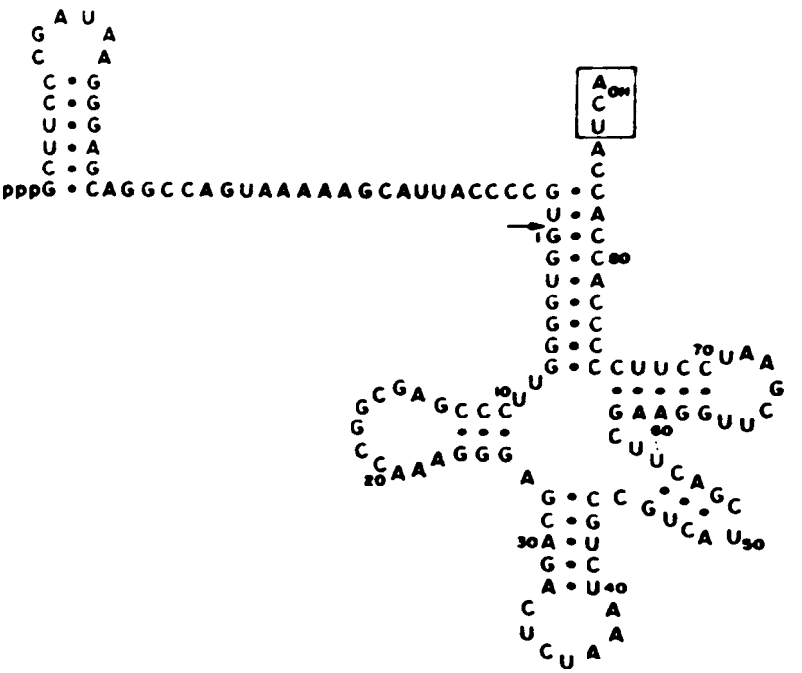

Figure 4. Nucleotide sequence of the precursor to $tRNA^{Tyr}su_3^+$. The arrow pointing toward the sequence indicates the site of cleavage by RNase P on the 5′ side of nucleotide 1 of the mature tRNA sequence. The boxed nucleotides are extra nucleotides at the 3′ terminus (after Altman and Smith, 1971).

intra-cellular ribonucleases than the transcripts of the wild-type gene. The "extra" sequences, themselves, though of interest because no such segments of gene transcripts had been characterized at that time, proved not to be particularly revealing *per se.*

Although the earlier work of Darnell (Bernhardt and Darnell, 1967), Burdon (Burdon, 1971; 1974) and their coworkers had shown that tRNAs were probably made from precursor molecules in eucaryotic cells, further characterization of the enzymes involved in the biosynthesis of tRNA, or tRNA processing events, could not proceed without a radiochemically pure, homogeneous substrate of the kind that I had isolated.

When the the precursor to $tRNA^{Tyr}$ was mixed with an extract of *E. coli,* it was immediately apparent that enzymatic activities were present in the cell extract that could remove the "extra" nucleotides from both the 5′ and 3′ ends of the mature tRNA sequence (Fig. 5; Altman and Smith, 1971; Robertson et al., 1972). The activity that processed the 5′ end of the tRNA precursor, which we named Ribonuclease P, did so by one endonucleolytic cleavage event in contrast to what appeared to be non-specific exonucleolytic degradation at the 3′ end of the molecule. In fact, no ribonucleases with such limited specificity with respect to the site of cleavage as that exhibited by RNase P were known at that time, so the novelty of this reaction assured our continuing interest in it. Some characterization of the reaction was

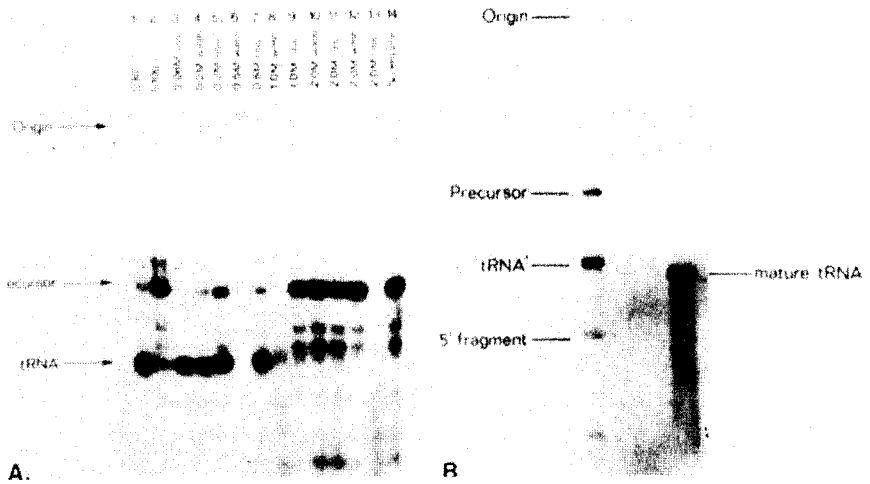

Figure 5. A. RNase P activity in extracts of *E. coli.* **Extracts of** *E. coli,* **partial purification of RNase P, and cleavage reactions were carried out as generally described in Robertson et al., (1972). The substrate used was the precursor to** *E. coli* $tRNA^{Tyr}$. **B. Separation by gel electrophoresis of the products of cleavage in vitro of the precursor to $tRNA^{Tyr}A25$ by a partially purified preparation of RNase P. The 3′ end fragment ('tRNA') includes the additional nucleotides of the precursor. (Reprinted with permission from Robertson et al., 1972).**

immediately carried out in collaboration with Hugh Robertson and John Smith (Robertson et al., 1972).

Characterization of Ribonuclease P from Escherichia coli

At the MRC laboratory, we showed that RNase P produced 5′ phosphate and 3′ hydroxyl groups at its site of cleavage (Robertson et al., 1972), unlike most non-specific nucleases which produce 5′ hydroxyl and 3′ phosphate groups. This observation fitted with the fact that mature tRNAs have a 5′ phosphate at their 5′ termini. While some progress was made in terms of chromatographic purification of the enzyme, in retrospect the most striking observation made in the early studies was that "it is possible that the active form of RNase P, which must have a strong negative charge, could be associated with some nucleic acid." The next important step was taken a few years later by Benjamin Stark, a graduate student in my laboratory, who showed that an RNA of high molecular weight copurified with the enzymatic activity and, in a classic experiment, he demonstrated that this RNA molecule was essential for enzymatic activity (Stark et al., 1978). (The RNA was named Ml RNA and was later shown to be similar to a stable RNA species [band IX] of unknown function that had been described by Ikemura and Dahlberg (19′73) as one of a series of minor RNA species found in *E. coli;* the protein subunit of RNase P from *E. coli* was named C5 protein.)

The essential role of the RNA component was established by first treating RNase P with micrococcal nuclease, an enzyme that destroys RNA, and subsequently assaying the treated enzyme for RNase P activity: there was none after treatment with micrococcal nuclease (Fig. 6) or, for that matter,

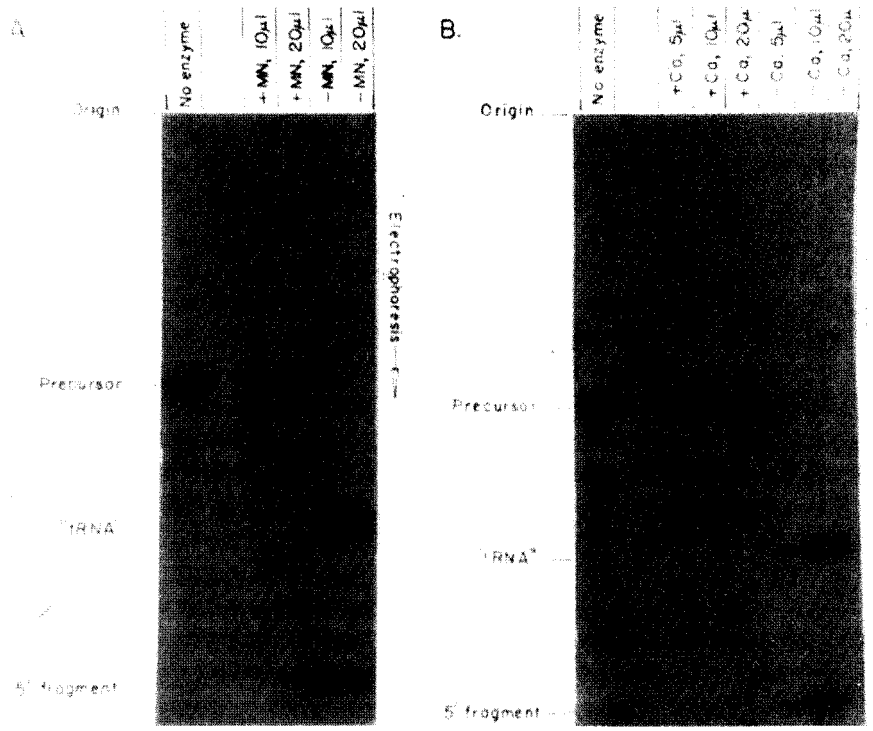

Figure 6. Inactivation of RNase P with RNase pretreatment. Control reactions were performed without micrococcal nuclease (MN) in the pretreatment mixture (A) or without CaCl, (R). RNase P pretreated with MN had less that 5% the activity of control reaction RNase P. The extent of inactivation tan be varied by changing the reaction conditions as shown. (Reprinted with permission from Stark et al., 1978).

after treatment with various proteinases. Thus, under the conditions we were then using (that is, buffers that contained 10 mM $MgCl_2$), both protein and RNA components were shown to be essential for enzymatic activity. Concurrently we showed that the enzyme had a buoyant density in CsCl of 1.72 g/ml (Stark et al., 1978), characteristic of an RNA-protein complex that consists predominantly of RNA. Velocity sedimentation experiments had previously determined the sedimentation coefficient to be 12.5 S (Robertson et al., 1972).

While the biochemical purification was proceeding, the study of mutants of *E. coli* made by Schedl and Primakoff (1973; Schedl et al., 1974), Shimura, Ozeki and their coworkers (Ozeki et al., 1974; Sakano et al., 1974) showed that RNase P was an essential enzyme in *E. coli* for the biosynthesis of all tRNAs and that both RNA and protein subunits were required *in vivo.* Furthermore, additional work from the laboratories of William McClain (reviewed in McClain, 1977) and John Carbon (Carbon et al., 1974) added to the evidence that RNase P was responsible for the processing of many different tRNA precursor molecules. Although appropriate genetic analyses could not be performed, we also showed that RNase P-like activities

existed in the extracts of cells from many other organisms, including humans (Altman and Robertson, 1973; Garber et al., 1978). These early studies showed that RNase P was capable of cleaving many different tRNA precursor molecules and that there was no identifiable similarity in terms of nucleotide sequence around the sites of cleavage. The manner in which the enzyme recognized its sites of cleavage in different substrates with such selectivity seemed worthy of study, and recognition of some feature of the structure in solution, common to all tRNA precursor molecules, was suspected.

When Stark's experiments were published we did not have the temerity to suggest, nor did we suspect, that the RNA component alone of RNase P could be responsible for its catalytic activity. The fact that an enzyme had an essential RNA subunit, in itself, seemed heretical enough. Shortly thereafter, however, when Ryszard Kole demonstrated that the enzyme consisted of an RNA (Ml RNA) and a protein subunit (C5 protein; Mr ~ 14,000), which were not covalently linked and which could be separated into inactive subunits and then reconstituted to form an active enzyme (Kole and Altman, 1979), the similarities in chemical composition and properties of assembly of this system to that of the ribosome were sufficiently striking that we could not escape thinking about the possibility that the RNA, at the very least, participated in the formation of the active site of the enzyme. Indeed, making the comparison with ribosomes proved to be important in overcoming some resistance to the idea that an enzyme could have an RNA subunit. From a purely chemical point of view, there was no reason why RNA could not participate in formation of the active site (Kole and Altman, 1981) or even in catalysis itself.

The advent of recombinant DNA technology and powerful systems for the transcription *in vitro* of isolated pieces of DNA enabled us to characterize in some detail the RNA subunit of RNase P (377 nucleotides in length; Reed et al., 1982) and to prepare large quantities for biochemical experiments. Concurrent progress in our purification of the protein subunit prepared us for a series of experiments, conducted in collaboration with Norman Pace's group, in which we made hybrid enzymes with subunits from *E. coli* (prepared in our laboratory) and from *B. subtilis* (prepared in Pace's laboratory). As an offshoot of these experiments, Cecilia Guerrier-Takada in my laboratory was testing reconstituted enzymes under ionic conditions optimal for the activity of the holoenzyme from *B. subtilis* and different from the ones we had previously usually employed. She found, in control experiments, that the RNA subunit from *E. coli,* exhibited catalytic activity of its own in buffers that contained 60 mM $MgCl_2$ (An example of such reactions is shown in Fig. 7; in fact the catalytic activity of Ml RNA is evident when the concentration of Mg^{2+} is greater than 20 mM; Guerrier-Takada et al., 1983). The protein subunit of the enzyme increased the $\mathbf{k_{cat}}$ by a factor of 10-20 but had little effect on the K_m. These observations were possible because of the purity of our preparations of Ml RNA and the use of a natural substrate, the precursor to $tRNA^{Tyr}$ from *E. coli.*

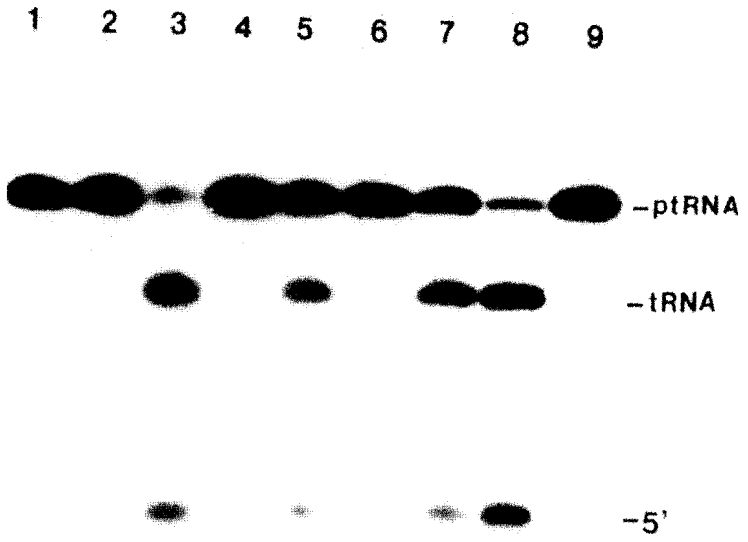

Figure 7. Dependence of the catalytic activity of Ml RNA on the concentration of Mg^{2+}. The precursor to $tRNA^{Tyr}$, abbreviated as pTyr, is the substrate. Lane 1: pTyr alone; 10 mM $MgCl_2$. Iane 2: Ml RNA added; 10 mM $MgCl_2$. Lane 3: Ml RNA and C5 protein added: 10 mM $MgCl_2$. Lane 4: pTyr alone; 100 mM $MgCl_2$. Lane 5: Ml RNA added; 100 mM $MgCl_2$.Lane 6:: Ml RNA and C5 protein added, 100 mM $MgCl_2$. Lane 7: Ml RNA added; 100 mM $MgCl_2$/4% polyethylene glycol. Lane 8: Ml RNA and C5 protein (20-fold excess) added; 100 mM $MgCl_2$. Lane 9: C5 protein added; 10 mM Mg Cl_2. Reactions were carried out as described by Guerrier-Takada et al., (1983).

We quickly determined that Ml RNA had all the properties of a true enzyme as defined in biochemistry textbooks (Fruton and Simmonds, 1958; p. 211): it was unchanged (in size) during the course of the reaction; it had a true turnover number as measured by Michaelis-Menten analysis of the kinetics (Fig. 8) and, therefore, it was a catalyst; it was needed in only small amounts and it was stable. Soon thereafter we proposed a model of the secondary structure of Ml RNA based on its susceptibility to nucleases in solution and some simple notions of the stability of RNA structures. We also rapidly outlined the general ionic requirements of the reaction (Table 2; Guerrier-Takada et al., 1986)). The curve of the dependence of the rate of the reaction on the pH is flat between 5 and 9 and is suggestive of the involvement of more than one group with a pK_a not characteristic of those found on nucleotides alone in solution. It is reasonable to expect, therefore, that the active site of Ml RNA is embedded in a folded structure and that the local environment of the active site will not be precisely identical to that of the aqueous buffer in which the whole molecule is dissolved.

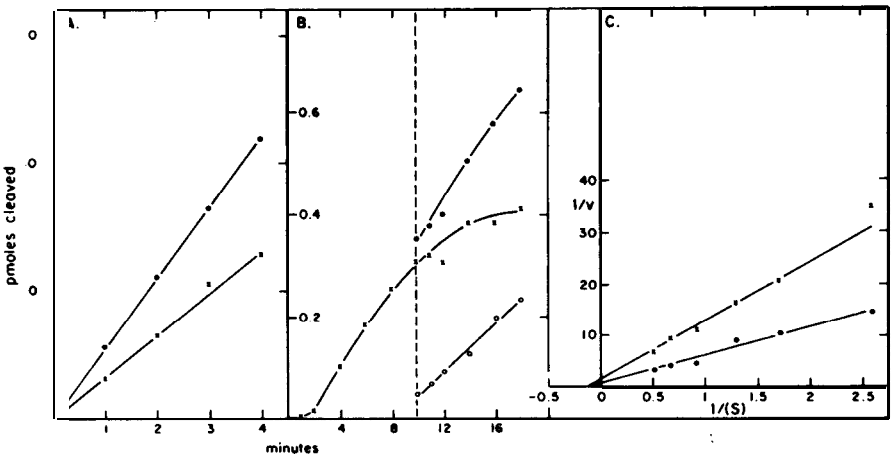

Figure 8. Kinetic analysis of the M1 RNA and RNase P reactions with the precursor to $tRNA^{TYR}$ (pTyr) as substrate. A. Comparison of the kinetics of reconstituted (by dialysis) *E.* coli RNase P in buffer that contained 5 mM $MgCl_2$ and M1 RNA in buffer that contained 60 mM $MgCl_2$ that had been treated in the same way. (●) RNase P activity; (x) M1 RNA activity. B. Kinetics of M1 RNA action in buffer that contained 60 mM $MgCl_2$. M1 RNA was incubated with a five-fold excess of pTyr. 10 min after the start of incubation a further three-fold excess of pTyr, or buffer alone, were added to the reaction mixture. (●) pTyr added after 10 min; (x) buffer alone added after 10 min; (o) net added pTyr cleaved after 10 min (all in pmoles). C. Determination of K_m and V_{max} for the reactions shown in A. A Lineweaver-Burk double reciprocal plot was constructed from the appropriate kinetic data. (●) RNase P in buffer that contained 5 mM $MgCl_2$; :(x) M1 RNA in buffer that contained 60 mM $MgCl_2$. Units: 1/S (pmoles x $5x10^{-4}$)-'; 1/v [(pmoles substrate cleaved/min)-1], (Reprinted from Guerrier-Takada et al., 1983).

Table 2. *Catalytic Activity of M1 RNA.*

M1 RNA active[a]:
≥ 20 mM $MgCl_2$
10 mM $MgCl_2$ plus C5 protein
10 mM $MgCl_2$ plus 5 mM polyamine
M1 RNA not active:
10 mM $MgCl_2$

[a] The table summarizes data presented in Guerrier-Takada et al., (1983). The complete composition of reaction mixtures is given in the reference.

These findings complemented those of Cech's group (Cech et al., 1981; Cech and Bass, 1986, for review) on self-splicing RNA and started intense speculation about the role RNA may have played in the origin of life. However, our immediate interest was in determining precisely how the enzyme works, what its role is *in vivo,* and how it manages to recognize 60 or so different substrates in *E. coli* with no apparent sequence specificity around the site of cleavage.

RECENT WORK

Structure

The original model of the secondary structure of M1 RNA has been extensively refined by phylogenetic analysis (Fig. 9) carried out primarily by Pace and coworkers (James et al., 1988). However, this analysis has not yet yielded a satisfactory correlation between the phenotypes of mutants (Lumelsky and Altman, 1988) and features of the secondary structure of M1 RNA or its analogs from other bacteria (see below). It does provide the basis for hypotheses about the regions of M1 RNA that are essential for function (Waugh et al., 1989), as indicated by evolutionary conservation, and it

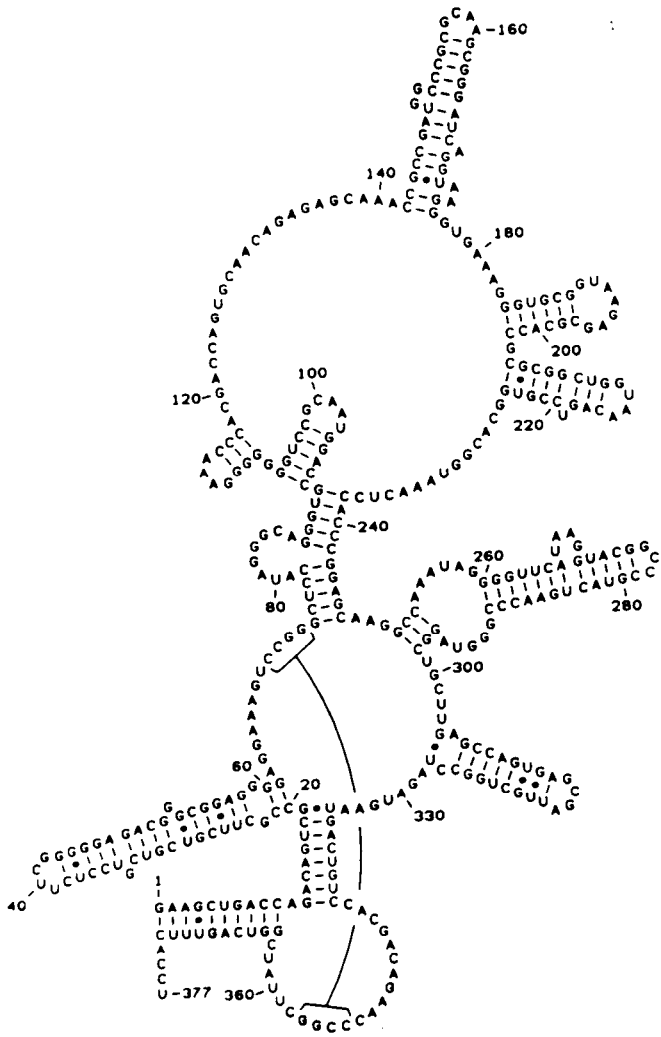

Figure 9. A model for the secondary structure of M1 RNA based on extensive phylogenetic analysis of the nucleotide sequence of the RNA subunit of RNase P from several eubacteria. (Reprinted with permission from James et al., 1988).

highlights the necessity of determining the three-dimensional features of the structure. To this end, additional phylogenetic comparisons, utilizing the data concerning the homolog of M1 RNA from several eucaryotic species (Miller and Martin, 1983; Krupp et al., 1986; Lee and Engelke, 1989), and crystallographic studies are in progress. One observation of continuing interest from these studies is that the evolutionary clock for both the RNA and protein subunits of RNase P seems to be a very fast one in comparison with that for rRNAs (Lawrence et al., 1987; Gold, 1988). Although the function of RNase P, as judged by the antigenic properties of the protein (Gold et al., 1988; Mamula et al., 1989), its ability to cleave various substrates and to reconstitute active enzyme with subunits from different organisms (Guerrier-Takada et al., 1983; Gold and Altman, 1986; Lawrence et al., 1987) has been highly conserved, the nucleotide sequences of the genes for the subunits of the enzyme have drifted extremely rapidly (Gold, 1988; Bartkiewicz et al., 1989).

Mechanism

The detailed mechanism of the reaction catalyzed by RNase P is not known but two proposals have been made. In one case (Guerrier-Takada et al., 1986), a variation of the S_N2 in-line displacement mechanism has been suggested in which a complex between one Mg ion and six water molecules facilitates the nucleophilic action of a water molecule in solution (Fig. 10). Investigations of the rRNA self-splicing reaction in *Tetrahymenu* in Tom Cech's laboratory indicate that the original proposal for the mechanism of the RNase P reaction may also be relevant in the self-splicing reaction (Cech, 1987; McSwiggen and Cech, 1989). In the other proposal for the mechanism of the RNase P reaction (Reich et al., 1988), the nucleophile is derived from groups on the surface of the enzyme and the role of the magnesium ion is not as clearly specified. Attempts are underway in our laboratory to test the first model, by the insertion of a phosphothioate bond at the cleavage site and analysis of the stereochemistry of the cleaved product.

While many aspects of structure-function relationships and clues to the mechanism of the reaction may be revealed if the crystal structure of the enzyme becomes available, the determination of a crystal structure may prove to be elusive. We have, therefore, embarked on an attempt to identify regions of Ml RNA that are critical for the reaction by cross-linking the substrate to the enzyme by irradiation with ultraviolet light. Such experiments have revealed that a cross-link is formed between a nucleotide close to the site of cleavage in the substrate (C -3) and residue C92 in M1 RNA (Guerrier-Takada et al., 1989). If C92 is deleted from M1 RNA, the kinetics of the enzymatic reaction, and its site of cleavage with particular substrates, are significantly altered. Furthermore, the region of secondary structure around C92 in M1 RNA resembles that of the tRNA E site in 23S rRNA (Fig. 11). Additional studies have shown that this site is important in the binding of the aminoacyl stem of a tRNA precursor to the enzyme and that,

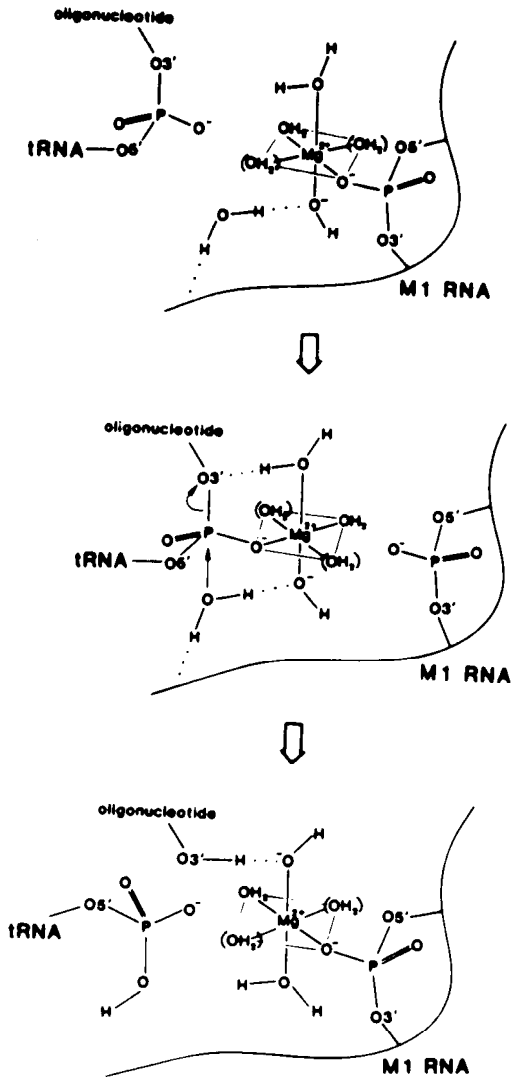

Figure 10. Hypothetical electronic mechanism of tRNA precursor hydrolysis by M1 RNA of RNase P. The reaction is catalyzed by an Mg-H_2O complex that is initially bound to a phosphate of M1 RNA. Mg^{2+} is formally shown as hexacoordinated, but it may well be tetracoordinated as indicated by the parentheses around the two equatorial water ligands. In the top panel, a water molecule from the solvent that will participate in hydrolysis is positioned by a hydrogen bond to an O or N atom in M1 RNA. In the middle and bottom panels, the tRNA precursor substrate is bound by the water molecule attached to M1 RNA and passes through a transition state prior to cleavage of the "extra" oligonucleotide and prior to the addition of OH to its O5′ terminal phosphate. After the reaction steps shown here, a solvent water chain between the axial ligands of Mg^{2+} recocks the enzyme for the next cycle. (Reprinted with permission from Guerrier-Takada et al., 1986. Copyright 1986 American Chemical Society).

as in the binding of tRNA to the E site of 23S rRNA (Moazed and Noller, 1989), the 3′ terminal CCA sequence plays a critical role in the interaction of the enzyme with the substrate. These results, in addition to allowing the

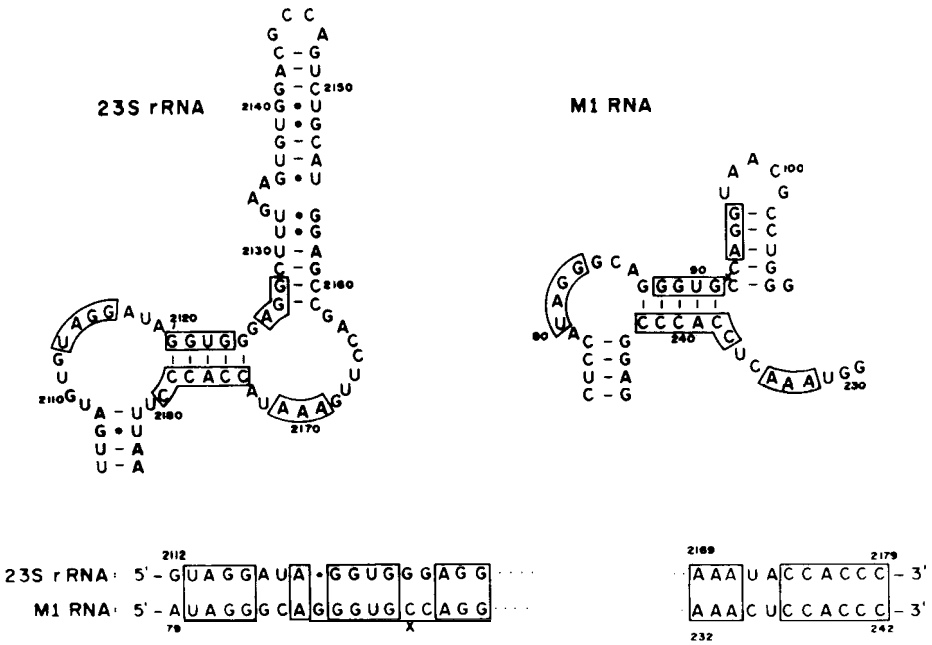

Figure 11. Comparison of part of the E site of 23S rRNA with a region in M1 RNA that surrounds the crosslink with the substrate. The secondary structures are taken from Moazed and Noller (1989) and James et al. (1988). The "x" marks C92, the nucleotide in M1 RNA that is crosslinked to the substrate. Nucleotides shown in boxes are found in approximately the same relative positions in the structures shown (see Guerrier-Takada et al., 1989).

identification of domains with similar structural and functional properties in RNA molecules with very different cellular functions, delineate a region of importance for function in M1 RNA and suggest further experiments for a more detailed definition of the interactions between enzyme and substrate and of the particular steps in the enzymatic reaction that involve this region.

Recognition of the substrate

Early notions of the features important for the recognition by RNase P of its substrate focussed either on the possibility of Watson-Crick pairing of nucleotide sequences common to all tRNAs (e.g. CCA and UUCG) with M1 RNA (Guerrier-Takada and Altman, 1986) or some other, incompletely specified, measuring mechanism that recognizes the three-dimensional structure of the tRNA moiety of the precursor (Bothwell et al., 1976). Results of several experiments indicated that extensive pairing between enzyme and substrate was not essential for the enzymatic reaction (Guerrier-Takada and Altman, 1986; Baer et al., 1987) and attention was focussed on the conformation in solution of the substrate. It had been demonstrated early on that RNase P from any one source can cleave tRNA precursors from any other source. Thus, when an unusual tRNA which lacked the D stem and loop was found, namely, $tRNA^{Ser}$ from bovine mitochondria (de

Bruijn and Klug, 1983), we examined whether M1 RNA could cleave an analog of a precursor to that tRNA.

In collaboration with Bill McClain (McClain et al., 1987), we showed not only that the D stem and loop were not essential for recognition, but also that the anticodon stem and loop were also dispensible (Fig. 12). A substrate that consisted merely of a single-stranded region at the 5′ end of a mini-tRNA, containing only the T stem and loop stacked on the aminoacyl acceptor stem, was cleaved almost as efficiently as the parent tRNA precursor from which it was derived. This minimal substrate contains an RNA helix that is analogous to one part of the intact three-dimensional structure of a normal tRNA (Fig. 2; Rim et al., 1974; Robertus et al., 1974)). Recently, we have also shown that neither the loop segment of the structure, nor more than six base pairs in the helical region, and, in separate experiments, that no more than one "extra" nucleotide at the 5′ terminus are needed for cleavage by M1 RNA. While substrates with these features are not cleaved as efficiently as either a normal precursor tRNA or pAT1 (the substrate shown in Fig. 12 that resembles a hairpin), they must, nevertheless, contain sufficient recognition elements to allow the reaction to proceed. All these model substrates have the 3′ terminal CCA sequence and none are cleaved at detectable levels when the CCA sequence is altered. Therefore, it appears that the CCA sequence and at least one half turn of an RNA helix play an essential role in substrate recognition.

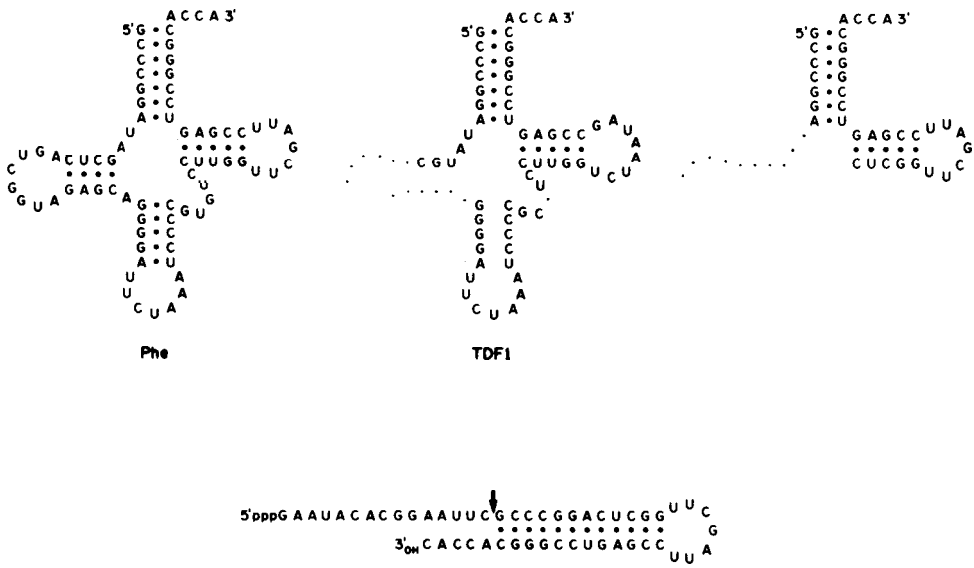

Figure 12. A. Structure of mature tRNA^{Phe}, and derivatives of it, encoded by synthetic genes. The sequences are shown without the modified nucleotides characteristic of mature tRNA^{Phe} because they are not present in the transcripts made *in vitro.* The transcripts contain "extra" nucleotides at both ends of the molecules (after McClain et al., 1987). B. Structure of the AT-l precursor drawn in a stem and loop structure similar to the corresponding region in tRNA^{Phe}. The arrows mark the site of cleavage by RNase P or M1 RNA. The sequence of mature AT-l is shown in A. (Reprinted with permission from McClain et al., 1987).

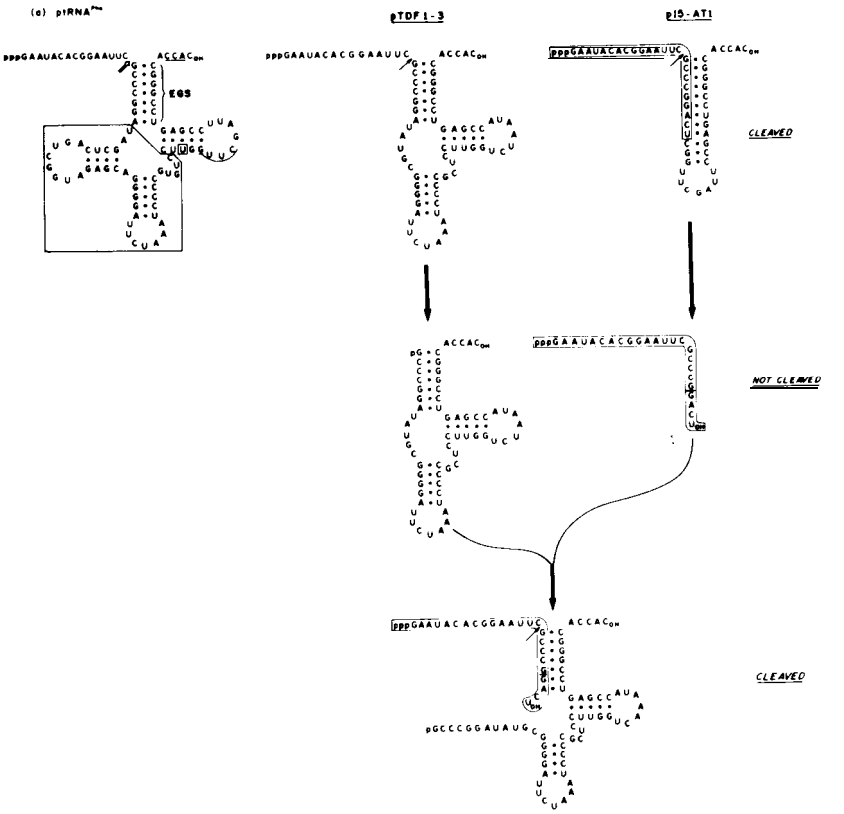

Figure 13. Scheme for formation of substrates for RNase P by hybridization of two oligoribonurleotides. TDF-I (see Fig. 3) was prepared by RNAse P cleavage *in vitro* of its precursor molecule that had been transcribed *in vitro* (see McClain et al., 1987). A portion (boxed sequence) of the precursor to AT-1 (Fig. 3) was prepared by transcription *in vitro* of a restriction fragment of the DNA encoding the AT-I synthetic gene (A.C. Forster, personal communication).

Conclusions from experiments with model substrates have to be tempered by the knowledge that some recognition elements, which appear to play a prominent role in these examples, may not play as important a role and may be supplemented by other elements in the normal tRNA precursors found in cells. It is certainly the case that a change in the D or anticodon stems of a normal tRNA precursor can have a dramatic effect on the rate of cleavage by RNase P even though these entire regions of the substrate are absent in the model substrates.

Through the hybridization of two oligoribonucleotides as shown in Fig. 13, we can create ad manipulate novel substrates. An "external guide sequence" which can guide RNase P to its target, can be hybridized to any other RNA of known sequence and will form the downstream, or 3′ part, of the substrate. RNase P should then cleave the hybrid target at the junction between the single- and double-stranded region at the 5′ side of the double-stranded region. This new method presents opportunities to investigate more precisely the details of the recognition mechanism and it also pro-

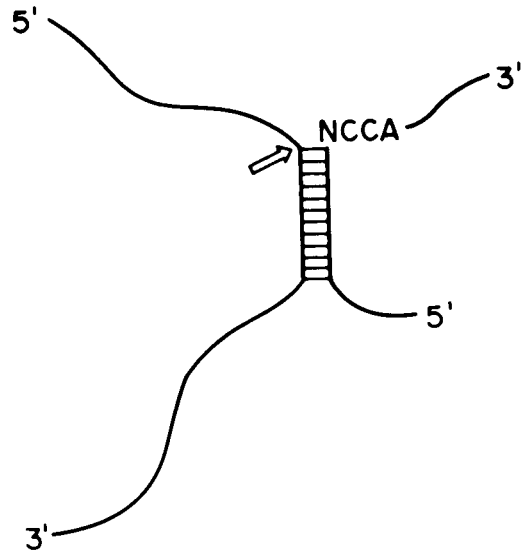

Figure 14. Targeting of RNA for cleavage by RNase P. An external guide sequence (EGS) is shown by the shorter line ending in NCCA. N is most frequently found to be A in tRNA molecules. The region of the EGS shown as hydrogen-bonded is designed to be complementary to a region of known sequence in the RNA to be targeted.

vides, in principle, a means to inactivate any mRNA of known sequence *in vivo.* Aside from the problem of the expression of the external guide sequence *in vivo,* the method does have the advantage that RNase P is already present in cells of all types. Providing that the hybrid can be designed to be compatible with the cleavage-site specificity of the enzyme in the particular host organism of choice, the target RNA should be inactivated.

In this example of the use of one oligoribonucleotide to target another RNA that is to be cleaved by RNase P, substrate recognition by the enzyme resembles, in a formal sense, selection of the site of cleavage by some of the other known RNA catalysts. Group I introns and the satellite and similar RNAs use guide sequences (Cech, 1987; Altman, 1989, for reviews) in the selection of cleavage sites or to form structures in which a cleavage site becomes defined. In virtually all other respects, these reactions are quite distinct from that carried out by RNase P.

The Past, Present and Future of RNase P

The discovery of RNA catalysis has led to new hypotheses about the origin of the earliest self-replicating biochemical systems from which the question of the origin of the genetic code can be excluded. Models of these early systems rely entirely on RNA as the genetic material and as the source of catalytic activity (Fig. 15; Darnell and Doolittle, 1986; Weiner and Maizels, 1987; Joyce, 1989). All this speculation clearly presupposes that what we see in present-day systems reflects, in some manner the properties of RNA over

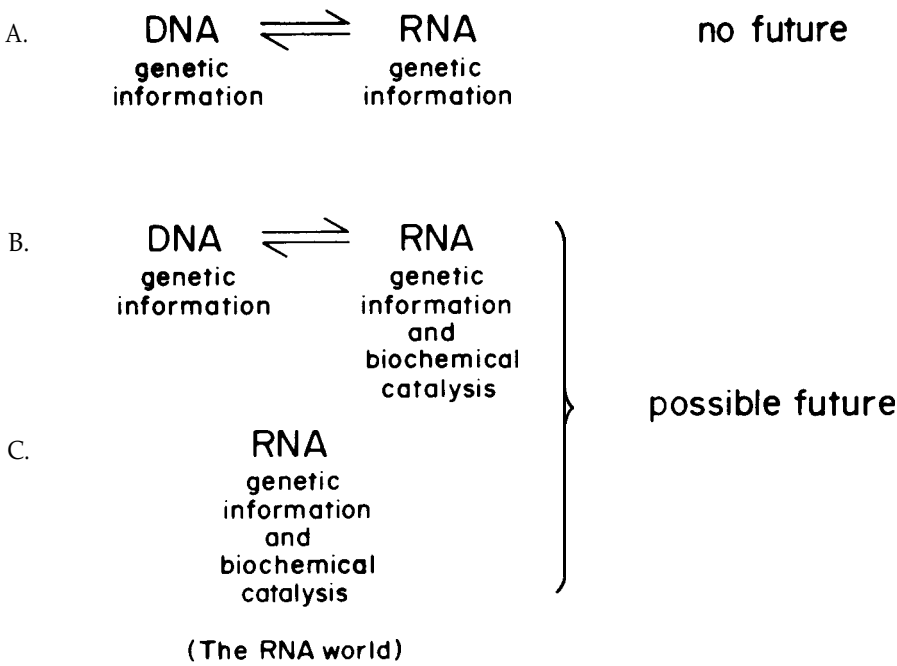

Figure 15 . A representation of three possible schemes of information transfer before proteins were part of the scheme.

a billion years ago. Should that indeed be the case, the richness of biochemical mechanisms exhibited by RNA (Table 1) is impressive and can allow for rather complex systems to develop in the absence of protein and DNA. In this limited context, we shall consider some aspects of the reactions governed by RNase P *in vitro.*

Although Ml RNA can cleave very simple substrates, it is apparent that these particular cieavage reactions cannot occur *in vivo* today because such cleavages would occur too frequently in the population of RNA in any cell: that is, the entire population of RNA molecules would be too susceptible to degradatian by RNase P. However, one can imagine that in an RNA world, there was considerable advantage to having an RNA molecule that could identify many sites in very long molecules generated by enzymatic or nonenzymatic mechanisms. The proliferation of many smaller molecules from larger ones would give rise to the possibility of a great variety of conformations of RNA in solution, some of which may have endowed RNA with catalytic activity or other useful functions very long transcripts did not have.

Setting aside for the moment the details of the origin of the genetic code and the appearance of proteins, one can ask, however, why a protein subunit became associated with Ml RNA. We recently showed that the protein subunit of RNase P can alter the site of cleavage and affect the rate of the reaction in a manner sensitive to the nature of the particular substrate being used (Guerrier-Takada et al., 1988; 1989). Thus, it is possible

that proteins may have fine-tuned the site specificity of RNA enzymes by enhancing the rates of reaction at particular sites and with particular substrates. What we see today as the "normal" cleavage sites of RNA enzymes may have been selected for over the eons, in conjunction with the appearance of protein cofactors, as physiological conditions changed during evolutionary time. The "unselected" reactions, for example those with very small hairpin substrates, became in consequence second- or lower-order reactions and are no longer relevant to events *in vivo.*

Finally, why do RNA enzymes only cleave phosphodiester bonds? There are three answers that come readily to mind. First and most trivially, RNA enzymes may cleave other classes of bonds and we just have not yet made the critical observations or found the right reaction conditions (the last part of the answer is a generic response to questions about the lack of success in performing certain reactions *in vitro).* Second, it is possible that, in the RNA world, perhaps RNA molecules could only cleave phosphodiester bonds: it was a primitive world and no other reactions were governed by enzymes. Once proteins appeared on the scene there was no further need to diversify RNA enzymes. Lastly, and most important, the chemistry of RNA enzymes, when sufficiently well-understood, may indicate to us that there is a compelling reason why RNA molecules cleave only phosphoester bonds. The validity, or lack of it, of this last answer can be tested by direct experimentation, and therein lies the work of the next several years.

ACKNOWLEDGEMENTS

My indebtedness to so many people makes it impractical to list them all here. Nevertheless, I wish to express my gratitude to my parents, my family, my teachers (especially Leonard Lerman, Mathew Meselson, Sydney Brenner and Lee Grodzins), my professional colleagues and collaborators (especially, Hugh Robertson and Bill McClain) and my students and coworkers in my laboratory (especially Cecilia Guerrier-Takada, Ben Stark, Ryszard Kole, Robin Reed and Madeline Baer). They have all tolerated my bouts of obsessiveness and have shared the moments of discovery and pleasure. The taxpayers of the United States, through the agencies of the National Institutes of Health and the National Science Foundation, have generously supported my work.

REFERENCES

Abelson, J. N., Gefter, M. L., Barnett, L., Landy, A. and Russell, R. L. (1970) J. Mol. Biol. 47 15-28.

Attman, S. (1971) Nature New Biology 229 19-21.

Attman, S. (1989) Adv. Enzymol., ed A. Meister (J. Wiley, NY) Vol. 62, pp. 1-36.

Attman, S. and J. D. Smith (1971) Nature New Biology 233 35-39.

Attman, S. and Robertson, H. D. (1973) Molec. Cell. Biochem. *1* 83-93.

Baer, M. F., Reilly, R. M., McCorkle, G. M., Hai, T-Y, Altman, S. and RajBhandary, U. I.. (1989) J. Biol. Chem. 26.3 2344-2351.

Bartkiewicz, M., Gold, H. and Altman, S. (1989) Genes and Development 3 488-499.

Bernhardt, D. and Darnell, J. E. (1967) J. Mol. Biol. 42 43-56

Bothwell, A. I.. M., Stark, B. C. and Altman, S. (1976) Proc. Nat. Acad. Sci. USA 73 1912-1916.

Burdon, R. H. (1971) Prog. Nucl. Acids Res. Mol. Biol. *11 33 - 79.*

Burdon, R. H. (1974) Brookhaven Symp. Biol. 26 138-153.

Carbon, J., Chang, S. and Kirk, L. L. (1974) Brookhaven Symp. Biol. 26 26-36.

Cech, T. R. (1987) Science 236 1532-1539.

Cech, T. R. and Bass, B.L. (1986) Annu. Rev. Biochem. 55 599-629.

Cech, T. R., Zaug, A. J. and Grabowski, P. J. (1981) Cell 27 487-496.

Chang, D.D. and Clayton, D.A. (1987) Science 235 1178-1184.

Crick, F. (1968) J. Mol. Biol. *38* 367-379.

Crick, F. (1970) Nature 277 561-563.

Darnell, J. E. and Doolittle, W. F. (1986) Proc. Natl. Acad. Sci. USA 83 1271-1275.

de Bruijn, M. H. L. and Klug, A. (1983) EMBO J. 2 1309-1321.

Fruton, J., and Simmonds, S. (1958) General Biochemistry, 2nd ed. (New York: J. Wiley and Sons).

Garber, R. I.., Siddiqui, M. A. Q. and Altman, S. (1978) Proc. Nat. Acad. Sci. USA. 75 635-639.

(Gilbert, W. (1986) Nature 319 618.

Gold, H. A. (1988) Ph.D. Thesis, Yale University, New Haven, CT USA.

Gold, H. A. and Altman, S. (1986) Cell 44 243-249.

Gold, H. A., Craft, J., Hardin, J. A., Bartkiewicz, M. and Altman, S. (1988) Proc. Nat. Acad. Sci. USA 85 5483-5487.

Guerrier--Takada, C. and Altman, S. (1984) Biochemistry 23 6327-6334.

Guerrier-Takada, C. and Altman, S. (1986) Cell 45 l77-183.

Guerrier-Takada, C. Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) Cell 35 849-857.

Guerrier-Takada, C., Haydock, K., Allen, L. and Altman, S. (1986) Biochemistry 25 1509-1515.

Guerrier-Takada, C., Knap, A. K., Lumetsky, N. and Altman, S. (1989) Science, 246 1578-1584.

Guerrier-Takada, C., van Betkum, A., Pleij, C. W. A. and Altman, S. (1988) Cell 53 267-272.

Guthrie, C., Seidman, J. G., Altman, S., Barrett, B. G., Smith, J. D. and McClain, W. H. (1973) Nature New Biol. 246 6-11.

tkemura, T. and Dahlberg, J. E. (1973) J. Biol. Chem. 248 5024-5032.

James, B., Olsen, G. J., Lin, J. and Pace, N. (1988) Cell 52 19-26.

Joyce, G. F. (1989) Nature 338 217-223.

Kim, S. H., Suddath, F. L., Quigtey, G. J., McPherson, A., Sussman, J. L., Wang, A. H. J., Seeman, N. C. and Rich, A. (1974) Science 185 435-440.

Kote, R. and Altman, S. (1979) Proc. Natl. Acad. Sci. USA 76 3795-3799.

Kole, R. and Altman, S. (1981) Biochem. 20 1902-1906.

Krupp, G., Cherayil, B., Frendeway, D., Nishikawa, S. and Soll. D. (1986) EMBO J. 5 1697-1703.
Lawrence, N. P., Richman, A., Amini, R. and Altman, S. (1987) Proc. Natl. Acad. Sci. USA 84 6825 - 6829.
Lee, J-Y., and Engelke, D. R. (1989) Molec. Cell. Biol. 9 2536-2543.
Lumelsky, N. and Altman, S. (1988) J. Mol. Biol. 202 443-454.
Mamula, M.J., Baer, M., Craft, J. and Altman, S. (1989) Proc. Nat. Acad. Sci. USA, 86 8717-8721.
McClain, W. H. (1977) Accts. Chem. Res. *10* 418-425.
McClain, W. H., Guerrier-Takada, C. and Altman, S. (1987) Science 238 527-530.
McSwiggen, J. A. and Cech, T. R. (1989) Science 244 679-683.
Miller, D. L. and Martin, N. C. (1983) Cell 34 91 1-917.
Moazed, D. and Noller, H. F. (1989) Cell 57 585-597.
Orgel, L. (1968) J. Mol. Biol. 38 381-393.
Ozeki, H., Sakano, H., Yamada, S., Ikemura, T. and Shimura, Y. (1974) Brookhaven Symp. Biol. 26 89-105.
Reed, R., Baer, M., Guerrier-Takada, C., Donis-Keller, H. and Altman, S. (1982) Cell 30 627-636.
Reich, C. I., Olsen, G. J., Pace, B. and Pace, N. R. (1988) Science 239 178-181.
Robertson, H. D. (1986) Nature 322 16-17.
Robertson, H. D., Altman, S. and Smith, J. D. (1972) J. Biol. Chem. 247 5243-5251.
Robertus, J. D., Ladner, J. E., Finch, J. T., Rhodes, D., Brown, R. S., Clark, B. F. C. and Klug, A. (1974) Nature 250 546-55 1.
Russell, R. L., Abelson, J. N., Landy, A., Gefter, M. L., Brenner, S. and Smith, J. D. (1970) J. Mol. Biol. 47 1-13.
Sakano, H., Yamada, S., Ikemura, T., Shimura, Y. and Ozeki, H. (1974) Nucleic Acids Res. *1 355*-371.
Schedl, P., and Primakoff, P. (1973) Proc. Natl. Acad. Sci. USA 70 2091-2095.
Schedl, P., Primakoff, P. and Roberts, J. (1974) Brookhaven Symp. Biol. 26 53-76.
Smith, J. D., Barnett, L., Brenner, S. and Russell, R.L. (1971) J. Mol. Biol. *54* 1-14.
Stark, B. C., Kole, R., Bowman, E. J. and Altman, S. (1978) Proc. Natl. Acad. Sci. USA 75 3717-3721.
Waugh, D. S., Green, C. J. and Pace, N. R. (1989) Science 244 1569-1571.
Weiner, A. M. and Maizels, N. (1987) Proc. Nat. Acad. Sci. USA 84 7383-7387.
Westheimer, F. H. (1986) Nature *319 534-535.*
Woese, C. R. (1967) The Origins of the Genetic Code (Harper and Row, New York)