

Crystallization of RNA–Protein Complexes I. Methods for the Large-scale Preparation of RNA Suitable for Crystallographic Studies

Stephen R. Price, Nobutoshi Ito, Chris Oubridge, Johanna M. Avis and Kiyoshi Nagai*

MRC Laboratory of Molecular Biology, Hills Road
Cambridge CB2 2QH, UK

In vitro transcription using bacteriophage RNA polymerases and linearised plasmid or oligodeoxynucleotide templates has been used extensively to produce RNA for biochemical studies. This method is, however, not ideal for generating RNA for crystallisation because efficient synthesis requires the RNA to have a purine rich sequence at the 5' terminus, also the subsequent RNA is heterogenous in length. We have developed two methods for the large scale production of homogeneous RNA of virtually any sequence for crystallization. In the first method RNA is transcribed together with two flanking intramolecularly-, (*cis*-), acting ribozymes which excise the desired RNA sequence from the primary transcript, eliminating the promoter sequence and heterogeneous 3' end generated by run-off transcription. We use a combination of two hammerhead ribozymes or a hammerhead and a hairpin ribozyme. The RNA–enzyme activity generates few sequence restrictions at the 3' terminus and none at the 5' terminus, a considerable improvement on current methodologies. In the second method the *Bsm*AI restriction endonuclease is used to linearize plasmid template DNA thereby allowing the generation of RNA with any 3' end. In combination with a 5' *cis*-acting hammerhead ribozyme any sequence of RNA may be generated by *in vitro* transcription. This has proven to be extremely useful for the synthesis of short RNAs.

Keywords: *in vitro* transcription; ribozymes; RNA–protein interactions; X-ray crystallography; oligoribonucleotide.

*Corresponding author

Introduction

The growing number of DNA–protein complex structures has provided considerable insight into the structural basis of sequence-specific interactions between protein domains and double-helical DNA (for a comprehensive review of the field see Harrison & Sauer, 1994). In contrast, protein–RNA interactions are so far poorly understood. Our lack of understanding of how proteins interact with their cognate RNA is due to the existence of only five published structures of a specialised type, the aspartyl-, glutaminy- and the seryl-tRNA synthetases with their cognate tRNAs (Rould *et al.*, 1989; Ruff *et al.*, 1991; Biou *et al.*, 1994) and the bacteriophage MS2 coat protein complexed to a small RNA hairpin (Valegård *et al.*, 1994) and the U1A

protein complexed with its cognate RNA hairpin (Oubridge *et al.*, 1994). RNA is capable of folding into complex structures through base-pairing of complementary stretches, base stacking and base-triplet formation (see Varani & Tinoco, 1991 for a comprehensive review of RNA secondary structure). These interactions create hairpin, bulge and internal loop structures within double-stranded regions, which are often binding sites for RNA binding proteins (Nagai, 1992). Structures of these RNAs with their binding proteins will reveal how unique nucleic acid structures and proteins interact. The lack of structural information on protein–RNA interactions is due in part to the difficulty of synthesising and purifying RNA of a defined length and sequence in large quantities, sufficient for structural work.

The chemical synthesis of RNA (Beaucage & Caruthers, 1981; Usman *et al.*, 1987) is less efficient than that of DNA due largely to the chemical protecting groups at the 2' oxygen of the ribose sugar. This inefficiency, at present, renders the synthesis of

Present address: N. Ito, Department of Biochemistry and Molecular Biology, University of Leeds, Leeds, LS2 9JT, U.K.

oligoribonucleotides of greater than about 30 nucleotides in length impractical at the preparative scale. RNA synthesis by *in vitro* transcription using bacteriophage RNA polymerases, which have been over-expressed and purified, is an alternative method (for a review see Francklyn & Schimmel, 1990). Linearised plasmid DNA or chemically synthesised oligodeoxynucleotides (Milligan *et al.*, 1987) are used as a template for run-off transcription, and allow the generation of milligram quantities of RNA. This method has proved to be powerful for preparing RNA for biochemical studies but has serious drawbacks as a method of RNA preparation for crystallisation. Firstly, the efficiency of transcription depends on the sequence of the initiation site as well as the promoter sequence. The yield of RNA is poor unless the 5' end of the transcribed RNA is purine-rich, thus there is significant restriction on the sequences that can be made efficiently. Secondly, the RNA polymerase often adds one or more extra nucleotides at the 3' end of the transcript in a template-independent manner (Milligan *et al.*, 1987).

In most DNA-protein complex crystals the packing of repeating units is mediated by end-to-end interactions of the DNA. RNA-protein complexes are likely to pack within crystals using similar interactions between the RNA components. Therefore, the homogeneity of oligonucleotide preparations and unrestricted choice of sequence, particularly at either end of the sequence, are likely to prove essential factors for success in RNA-protein co-crystallisation. In this paper we describe two methods of large scale RNA synthesis by *in vitro* transcription that overcome these difficulties. All methods use ribozymes in order to generate precise termini of RNA.

The hammerhead ribozyme (see inset of Figure 1a) was identified as a sequence sufficient for self-cleavage in satellite RNAs of certain plant viruses (Prody *et al.*, 1986; Hutchins *et al.*, 1986; Buzayan *et al.*, 1986a,c), and catalyses a transesterification reaction generating a cyclic 2',3'-phosphodiester and a free 5'-hydroxyl terminus of the product RNAs (Hutchins *et al.*, 1986; Buzayan *et al.*, 1986b; Epstein & Gall, 1987). Mutagenesis studies of the core of the hammerhead domain revealed that sequence requirements for cleavage immediately upstream of the cleavage site may be approximated to a NUX (N = rA, rC, rG or rU, X = rA, rC or rU) trinucleotide, although CAC, CGC and AAC trinucleotides at this site also appear to cleave efficiently when placed in the context of the domain shown in the inset of Figure 1a (Sheldon & Symons, 1989; Perriman *et al.*, 1992; Koizumi *et al.*, 1989). What is clear, however, is that the base-pairing potential of the stem immediately upstream of the cleavage site must be conserved with the nucleotide at the cleavage site unpaired (Koizumi *et al.*, 1988a,b; Sheldon & Symons, 1989). Divalent cations are believed to participate in the cleavage reaction with Mg²⁺ and Mn²⁺ being preferred (Uhlenbeck, 1987; Dahm & Uhlenbeck, 1991). Structural studies on the hammerhead ribozyme by NMR spectroscopy confirmed the domain structure

(Heus & Pardi, 1991), this being further refined through the three-dimensional structure by X-ray crystallography (Pley *et al.*, 1994).

Taira *et al.* (1991) have shown that a truncated RNA sequence can be generated when sequences at both 5' and 3' ends of RNA fold into hammerhead ribozyme structures. This method was originally designed to deliver hammerhead ribozymes into cells infected with HIV and to cleave the viral genomic RNA, but its possible use as a method for the large scale synthesis of RNA for crystallisation has not been investigated until now. We have further refined and optimised this method for large scale RNA preparation. The only sequence restriction of this method is the requirement for a UX (X ≠ G) (Sheldon & Symons, 1989; Perriman *et al.*, 1992; Koizumi *et al.*, 1989) di-nucleotide at the 3' end of the sequence. However, we are able to overcome this sequence restriction by the use of a hairpin ribozyme to generate the 3' terminus. The hairpin ribozyme was, like the hammerhead, identified as a self-cleavage domain from certain satellite RNAs of plant viruses (Feldstein *et al.*, 1989; Haseloff & Gerlach, 1989; Hampel & Tritz, 1989). The hairpin ribozyme requires a G, C or U nucleotide at the -2 position for efficient cleavage (Berzal-Herranz *et al.*, 1992) thus allowing the generation of any RNA terminating in the sequence (G/C/U)N. For instance, the generation of RNA terminating with a CCA trinucleotide, for example transfer-RNA, would be possible using a hairpin ribozyme to generate the 3' terminus.

The third method described here overcomes this restriction and involves transcription of a plasmid template in which the gene encoding the desired RNA sequence is preceded by a hammerhead ribozyme gene and followed by a restriction site for *Bsm*AI, an enzyme which cleaves outside its recognition sequence so that the plasmid template can be linearised for run-off transcription regardless of the nucleotide sequence at the 3' end. This third method complements, rather than supersedes, the others. Collectively, these methods allow the large scale preparation of virtually any sequence of RNA overcoming most of the current problems associated with *in vitro* transcription of RNA for crystallisation. RNA synthesised by this method has been used successfully for crystallisation of the U1A protein with its cognate RNA (see accompanying paper).

Results

The use of double *cis*-acting hammerhead ribozymes, 5' and 3' to the desired sequence yields large amounts of a homogeneous RNA product

A DNA fragment was designed containing the T7 phage ϕ 10 promoter and a 27 nucleotide sequence, which is flanked by 5' and 3' *cis*-acting hammerhead ribozyme sequences in order to make an RNA sequence representing part of the U2 small nuclear

RNA. This double-stranded fragment was produced by assembling four oligodeoxynucleotides and cloned into the *EcoRI-BamHI* site of pUC18. The plasmid DNA was linearised with *BamHI* and used for *in vitro* transcription. The RNA transcript was

designed to fold into a secondary structure shown in Figure 1a. If both ends of the transcript fold into the hammerhead ribozyme structure, the transcript should be cleaved at two sites following the GUC sequences, shown by the arrow, to liberate the

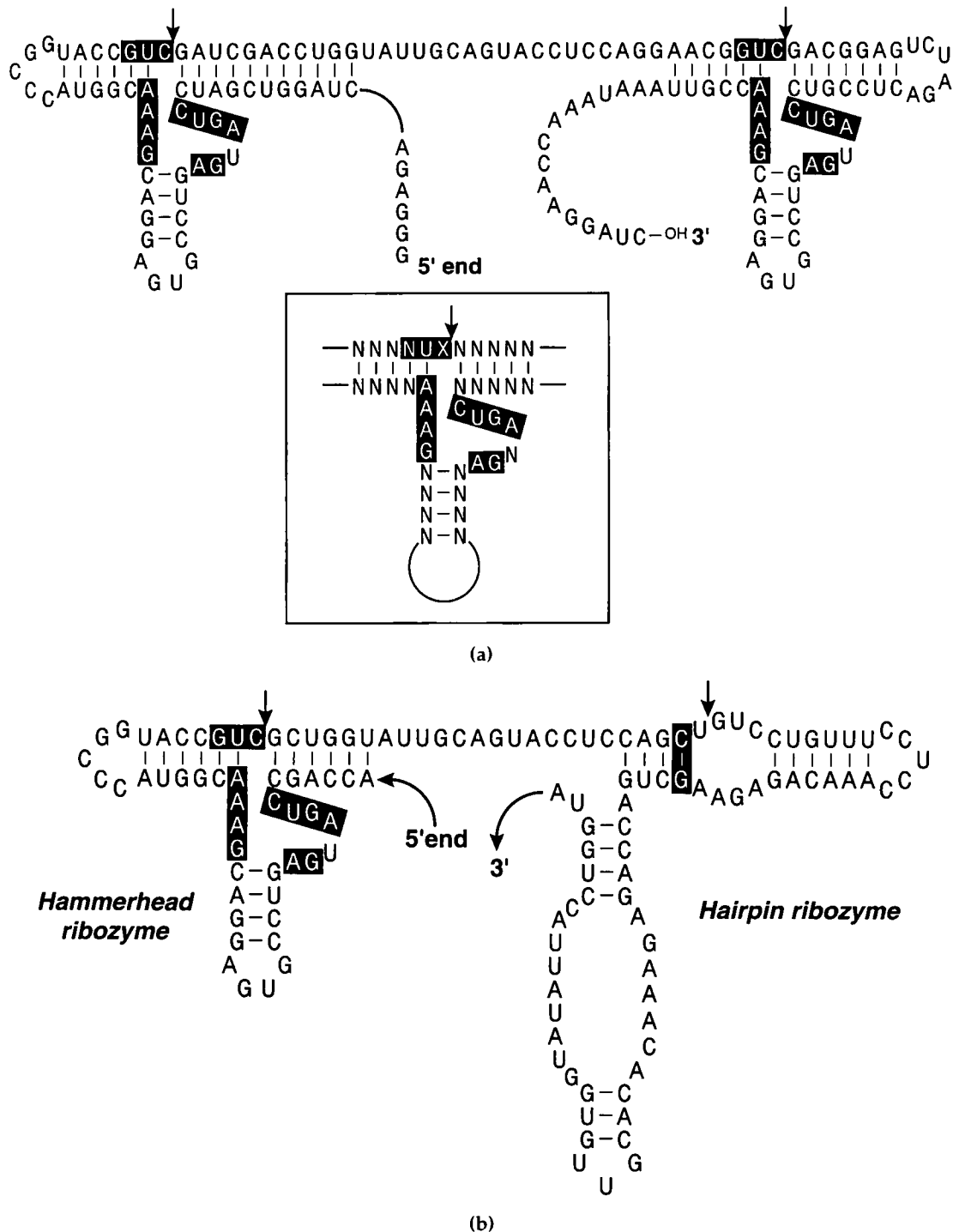


Figure 1. a, Sequence and designed secondary structure of a double *cis*-acting hammerhead ribozyme construct. The cleavage sites are marked with arrows, with cleavage releasing a 36 nucleotide RNA which constitutes part of the stem-loop four of human U2 snRNA. The inset shows a schematic of the hammerhead RNA structure with nucleotides conserved for cleavage highlighted, N = rA, rC, rG or rU, X = rA, rC or rU. b, The sequence and designed secondary structure of the 5' acting hammerhead ribozyme 3' acting hairpin ribozyme construct. The cleavage sites are marked with arrows. Positions required for efficient cleavage are highlighted, although their identity is explained in the text.

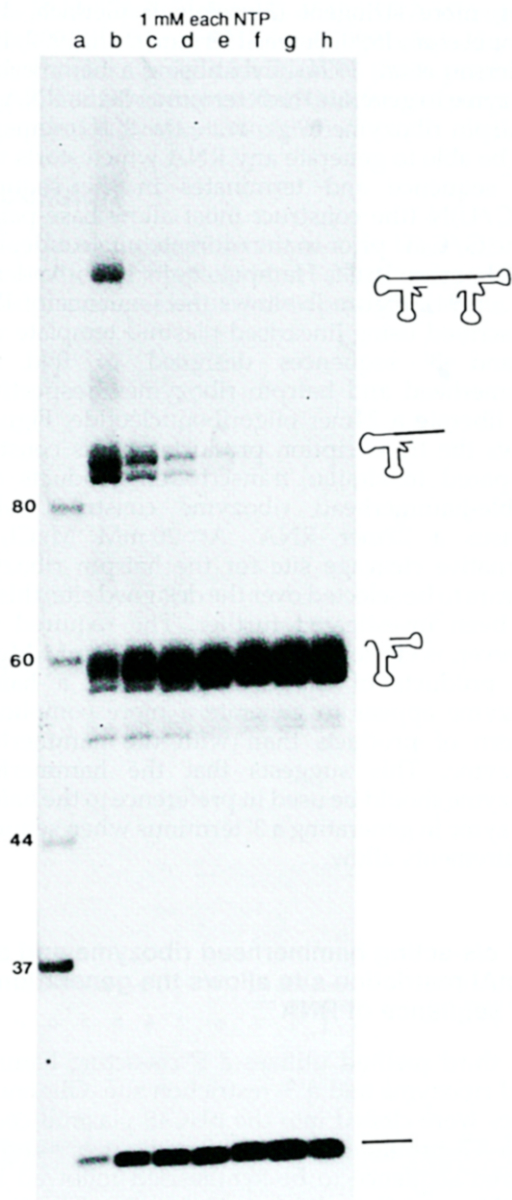


Figure 2. The effect of Mg^{2+} concentration on the efficiency of cleavage of the double hammerhead ribozyme construct with the sequence shown in Figure 1a. Lane a, single-stranded DNA markers, lanes b to h are titration points at 6, 10, 14, 20, 24, 28 and 32 mM added $MgCl_2$. All transcriptions include each nucleotide triphosphate at a concentration of 1 mM, and were continued for three hours at 37°C.

required fragment. the NUX ($X \neq G$) trinucleotide preceding both cleavage sites is essential for high efficiency cleavage, this restricts the RNA that can be made by this method to that which terminates in the sequence UX. Linearised plasmid template was transcribed in the presence of $[\alpha\text{-}^{32}\text{P}]\text{UTP}$ with varying concentrations of Mg^{2+} , in order to demonstrate the dependence on Mg^{2+} of cleavage of the ribozymes (see Figure 2). Several bands were produced at low concentrations of Mg^{2+} but above 24 mM Mg^{2+} only two major bands were seen. The

highest band from low Mg^{2+} conditions corresponds to the uncleaved, full length RNA and the second highest band is a product of a single cleavage at either the 5' or 3' cleavage site, the third highest band is a mixture of the 5' and 3' hammerhead fragments, which co-migrate in this gel system, and the lowest band is the 27 nucleotide U2 snRNA fragment. The nucleotide sequence at the transcription initiation site incorporated into this plasmid is the GGGAG sequence observed in the natural $\phi 10$ transcript and, therefore, the template is transcribed to high efficiency. Since the 27 nucleotide product migrates on a preparative polyacrylamide/urea gel away from any other bands, a large amount of transcription-cleavage product can be applied to the gel; this results in fewer gels being run in order to purify the RNA and a concomitant decrease in the risk of degrading the product before it can be used. In contrast, a simple run-off transcription reaction produces additional RNA fragments one or two nucleotides longer than the correct product, hence only a small amount of the transcription product can be applied to a preparative gel. Other ribozyme constructs have been tested which generate fragments ranging from 36 nucleotides to 140 nucleotides in length. All constructs show a similar Mg^{2+} dependence in their transcription and result in homogeneous RNA preparations. All further characterisation of the RNA produced by this method was carried out on the 36mer produced by a construct, the sequence of which is shown in Figure 1a.

In order to characterise RNA made by this method, unlabelled purified RNA was phosphorylated in the presence of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ using T4 polynucleotide kinase and analysed on a thin polyacrylamide/urea gel. The product of conventional *in vitro* transcription has a 5' triphosphate terminus and cannot be radio-labelled efficiently with ^{32}P by T4 polynucleotide kinase without prior treatment with calf intestinal alkaline phosphatase. In contrast, RNA produced by this method should have 5' hydroxyl and 2', 3' cyclic phosphodiester termini and should, therefore be phosphorylated by T4 polynucleotide kinase. As shown in Figure 3a, the phosphorylation reaction produces two closely running bands. It appears that one product is one nucleotide shorter than the other. However, this explanation may be discounted by the disappearance of the faster running product generating more of the slower running product during the time-course of the incubation with T4 polynucleotide kinase. In Figure 2, only one band is observed when product RNA is labelled using $[\alpha\text{-}^{32}\text{P}]\text{UTP}$ during the transcription reaction, and the analysis of this product on a polyacrylamide gel after elution under identical conditions produced a single band (data not shown). This indicates that the appearance of the double band is due to an additional activity present in the kinase reaction, and not impurity of the RNA itself.

Incubation of chemically synthesised RNA, which has hydroxyl groups at both 5' and 3' termini, with

polynucleotide kinase and [γ - 32 P]ATP generates only one product visible by autoradiography (data not shown). Therefore, it is most likely that the additional product is due to modification by the kinase, of the 2', 3' cyclic phosphodiester terminus (present due to the 3' ribozyme activity). Calf intestinal alkaline phosphatase has no phosphodiesterase activity and treatment of the RNA produced by ribozyme cleavages prior to phosphorylation with T4 polynucleotide kinase, had no effect on the time-course of the phosphorylation reaction (data not shown). Upon treatment of the non-phosphorylated RNA to decyclise the 2', 3' cyclic phosphate with maleic acid at pH 2 and subsequent incubation with calf intestinal alkaline phosphatase, to remove the mono-phosphate group, the lower band is substantially decreased (see Figure 3a, lanes g to l). Treatment of the RNA with maleic acid at pH 5.5, however, appears to leave the RNA unaffected, as assayed by the subsequent phosphorylation time-course (Figure 3a, lanes m to r). This observation is consistent with the existence of a 2', 3' cyclic phosphate terminus; treatment at pH 2 being sufficient for its decyclisation, in contrast to the treatment at pH 5.5 (Buzayan *et al.*, 1986b). The activity in the T4 polynucleotide kinase preparation that we have observed has been postulated to comprise part of a cyclic phosphodiesterase activity intrinsic to the enzyme itself (an unpublished result O. C. Uhlenbeck quoted Snyder, 1983). T4 phage polynucleotide kinase has an intrinsic 3' phosphatase activity, which is located independently to the kinase activity (Cameron & Uhlenbeck, 1977; Soltis & Uhlenbeck, 1982). The existence of this activity may be used to explain the time-course of reaction of T4 polynucleotide kinase with RNA generated by our ribozyme method, the upper band being the *cis*-diol terminated RNA, the lower being a phosphate mono-ester terminated RNA.

RNA produced by this method was also sequenced using the phosphorothioate footprinting method as described by Schatz *et al.* (1991). An analysis of the iodine cleavage products on a 20% thin polyacrylamide/urea gel demonstrated that the correct RNA sequence was produced, although the sequence of the first six nucleotides and the last was not verified (see Figure 3b). We have shown that the product is pure and strictly template-encoded and, therefore, suitable for structural investigation of RNA and RNA-protein complexes by NMR and crystallography. The only sequence requirement of this method is the NUX trinucleotide at the 3' end of the sequence which is essential for cleavage utilising the 3' *cis*-acting hammerhead ribozyme.

The use of a hairpin ribozyme to generate the 3' end of the desired RNA

The hairpin ribozyme has sequence requirements at the 5' side of the cleavage site that are less stringent than those of the hammerhead ribozyme, with sequence restrictions to the 3' side of the cleavage site

being more stringent (Haseloff & Gerlach, 1989; Hampel *et al.*, 1990; Berzal-Herranz *et al.*, 1992, 1993; Anderson *et al.*, 1994). By utilising a hammerhead ribozyme to generate the 5' terminus of the RNA and a hairpin ribozyme to generate the 3' terminus, we will be able to generate any RNA which starts with any sequence and terminates in the sequence (G/C/U)N (the construct must allow base-pairing of the G/C/U prior to the ultimate nucleotide, with the ribozyme itself, Hampel *et al.*, 1990; Anderson *et al.*, 1994). Figure 1b shows the sequence of RNA transcribed using linearised plasmid template with 5' and 3' sequences designed to fold into hammerhead and hairpin ribozymes, respectively, and liberate a 24mer oligoribonucleotide. Figure 4 shows the transcription products of this construct compared to similar transcription products of a double-hammerhead ribozyme construct which releases a 27mer RNA. At 20 mM MgCl₂ an alternative cleavage site for the hairpin ribozyme appears to be selected over the designed site; this has not been investigated further. The required site, however, is chosen over all others at 30 mM MgCl₂. The products of the cleavages using a hairpin ribozyme appear to generate a more complicated pattern of products than with the hammerhead ribozyme. This suggests that the hammerhead ribozyme should be used in preference to the hairpin ribozyme in generating a 3' terminus when sequence requirements allow.

A 5' *cis*-acting hammerhead ribozyme and a 3' *Bsm*AI restriction site allows the generation of any sequence of RNA

A third method utilizes a 5' *cis*-acting hammerhead ribozyme and a 3' restriction site. Oligonucleotides were cloned into the pUC18 plasmid coding for a T7 promoter, 5' *cis*-acting ribozyme and the desired sequence to be synthesised followed by a *Bsm*AI restriction site (see Figure 5a). The *Bsm*AI enzyme cleaves 5' to the sequence 5' GAGAC 3'. Therefore, the template DNA can be cleaved at any sequence if a *Bsm*AI site is placed downstream of the cleavage site. The 5' *cis*-acting ribozyme eliminates any sequence requirements of the RNA polymerase on the 5' side such that any RNA sequence can be synthesised at high levels by this method. Figure 5a shows the DNA sequence of one of the constructs, which generates a 24mer oligoribonucleotide. The analysis of the transcription product is shown in the left hand lane of Figure 5b. The upper band corresponds to the ribozyme fragment (labelled with a hammerhead ribozyme cartoon) and the lower band corresponds to the product together with additional bands arising from the addition of extra nucleotides to the 3' end which are often found with run-off transcripts. The right-hand lane of Figure 5b shows a transcription of a double hammerhead ribozyme construct. For RNA of 20 to 30 nucleotides these bands are well-resolved even at the preparative scale and a pure single product can be isolated.

The RNA binding domain of the U1A ribonucleoprotein with a 21mer RNA hairpin produced by this method yields crystals identical to those produced using chemically synthesized RNA (see accompanying paper, Oubridge *et al.*, 1995).

Discussion

The hammerhead ribozyme is one of the most-studied examples of a self-cleaving RNA. Although nucleotides at certain positions within this

folded domain are required for cleavage the identity of only two nucleotides (UX, X ≠ G) preceding the cleavage site is largely conserved. We have described three different methods for large scale RNA preparation by *in vitro* transcription all of which use ribozymes to generate one or both termini of the RNA. The double *cis*-acting ribozyme method originally described by Taira *et al.* (1989) was modified and conditions optimised for large scale preparation. The product to be isolated migrates on a preparative polyacrylamide gel away from the

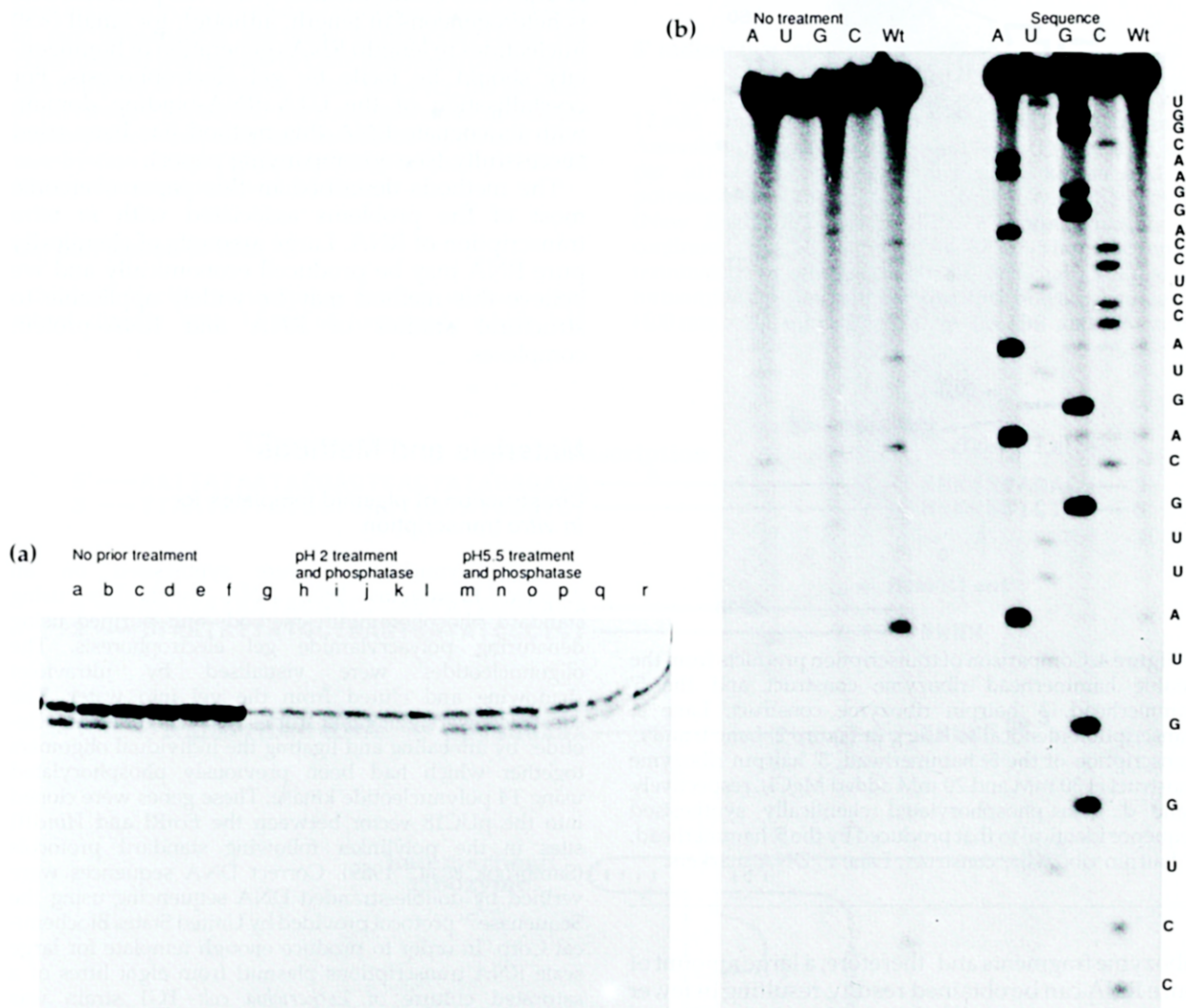


Figure 3. a, Characterisation of the termini of the RNA produced by a double hammerhead ribozyme construct. Lanes a to f show a time-course (5, 10, 20, 60, 90 and 120 minutes respectively) of the phosphorylation reaction with a 36mer RNA with no prior treatment. Lanes g to l show the same time-course after the RNA has been treated by maleic acid at pH 2 and subsequent incubation with calf intestinal alkaline phosphatase. Lanes m to r show the same time-course after the control incubation of the RNA with maleic acid at pH 5.5 and subsequent incubation with calf intestinal alkaline phosphatase (see the text for a full discussion). The origin of electrophoresis is not shown in the Figure. b, Sequencing of the U2 snRNA stem loop IV (36mer) produced by *cis*-acting ribozyme cleavage. Lanes A show reaction of RNA solution where adenosine phosphorothioate nucleotides have been incorporated into the sequence, with ethanol solution and with ethanolic iodine at 100 nM for left and right, respectively. Lanes U show reactions where uridine phosphorothioate nucleotides have been incorporated, similarly G stands for guanosine phosphorothioate nucleotide incorporation and C stands for cytosine phosphorothioate nucleotide incorporation. Lanes Wt show reaction of RNA in which no phosphorothioate analogues have been incorporated.

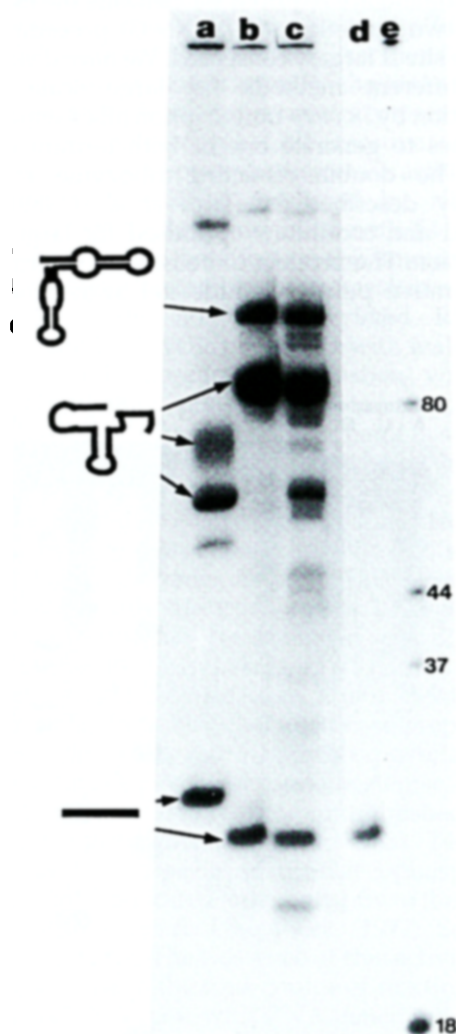


Figure 4. Comparison of transcription products from the double hammerhead ribozyme construct and the 5' hammerhead, 3' hairpin ribozyme construct. Lane a, transcription identical to lane g in Figure 2. Lane b and c, transcription of the 5' hammerhead, 3' hairpin ribozyme construct at 30 mM and 20 mM added $MgCl_2$, respectively. Lane d is a phosphorylated chemically synthesised sequence identical to that produced by the 5' hammerhead, 3' hairpin ribozyme construct. Lane e, DNA markers.

ribozyme fragments and, therefore, a large amount of pure RNA can be obtained readily, resulting in fewer gels being required for the purification and thus a reduced risk of ribonuclease contamination. For large RNA sequences it is quite difficult to separate the desired product from $N + 1$ and $N + 2$ products and the double *cis*-acting ribozyme method is essential in preparing RNA that is homogeneous in length. Although RNA made by this method requires a UX sequence at the 3' end, this restriction may not be critical for crystallisation of large RNAs and is certainly a great improvement on current technologies. A possible remedy for the sequence restrictions imposed by the hammerhead ribozyme at the 3' cleavage site is the use of other self-cleaving

sequences, such as the hairpin ribozyme at this position. The hairpin ribozyme has fewer sequence restrictions immediately 5' to the cleavage site and we show that it allows the synthesis of almost any sequence of RNA homogeneous in length. In addition, ribozymes with novel cleavage site specificities have been reported (Nakamaye & Eckstein, 1994), which may prove to be useful in the synthesis of RNAs terminating in different sequences.

For shorter RNAs the choice of sequence may be critical and therefore the third method using a 5' *cis*-acting ribozyme and 3' *Bsm*AI site is preferable. A drawback of this method is that the product RNA is heterogeneous in length, although for small (<30 nucleotides in length) RNAs generation of homogeneity should be facile by gel electrophoresis. For crystallisation of the U1A RNA-binding domain with its cognate RNA this method has been used successfully (see accompanying paper).

The methods described in this paper overcome most of the problems associated with *in vitro* transcription of RNA. Large amounts of chemically pure RNA may be produced economically and we believe this method may be widely applicable to structural studies of RNA and RNA-protein complexes.

Materials and Methods

Construction of plasmid templates for *in vitro* transcription

Oligodeoxynucleotides were synthesised on an Applied Biosystems 380B DNA Synthesizer using standard phosphoramidite methods and purified using denaturing polyacrylamide gel electrophoresis. The oligonucleotides were visualised by ultraviolet shadowing and eluted from the gel into water. The constructs were generated from several oligodeoxynucleotides by annealing and ligating the individual oligomers together which had been previously phosphorylated using T4 polynucleotide kinase. These genes were cloned into the pUC18 vector between the *Eco*RI and *Hind*III sites in the polylinker following standard protocols (Sambrook *et al.*, 1989). Correct DNA sequences were verified by double-stranded DNA sequencing using the SequenaseTM protocol provided by United States Biochemical Corp. In order to produce enough template for large scale RNA transcriptions plasmid from eight litres of a saturated culture of *Escherichia coli* TG1 strain was prepared by the alkaline lysis method (Sambrook *et al.*, 1989) and purified by caesium chloride density gradient centrifugation. We include RNase A treatment in the preparation of the plasmid template, even though this may appear to be foolhardy, as the plasmid will eventually come into contact with RNA. This plasmid DNA was linearised with a restriction endonuclease, phenol extracted three times and ethanol precipitated.

RNA synthesis by *in vitro* transcription

Adenosine-triphosphate was obtained from Calbiochem, cytidine-triphosphate, guanosine-triphosphate

and uridine-triphosphate were obtained from Sigma. All NTPs were acquired as the sodium salt with the exception of lithium guanosine triphosphate. Each nucleotide-triphosphate was stored as a 50 mM stock solution at pH 8.0 at -20°C with no loss of competence noted over a period of one year. T7 RNA polymerase was purified as described by Grodberg & Dunn (1988) with slight modifications (S.R.P., C.O. & K.N., unpublished). Transcription conditions use 40 mM Tris-HCl (pH 8), 30 mM MgCl₂, 200 µg linearized plasmid DNA, 4 mM ATP, 4 mM UTP, 4 mM GTP, 4 mM CTP, 20 µg T7 phage RNA polymerase, 2 mM spermidine in a 2 ml reaction. The enzyme reaction is continued for three hours at 37°C, the magnesium pyrophosphate precipitate is pelleted by centrifugation at 10,000 g for ten minutes and the supernatant removed. We have not used pyrophosphatase or ribonuclease inhibitors in our transcription reactions. Further, we have not experimented with the use of GMP to initiate transcription (see Sampson & Uhlenbeck, 1988). However, the amounts of polymerase added and the concentration of nucleotide triphosphates and template should be optimised in a case-by-case manner. Hence assaying the quality of the T7 RNA polymerase was not undertaken.

Gel purification of RNA

3 ml of *in vitro* transcription reaction mixture were mixed directly with 1.5 ml of formamide which had

been deionised by a 16 hour incubation with Amberlite MB-3 resin (BDH biochemicals) and applied to a 20 cm × 40 cm × 0.1 cm, 15% (w/v) polyacrylamide gel containing 8 M urea, 1 × Tris-borate EDTA buffer (TBE, Sambrook *et al.*, 1989). Bands containing the correct RNA product were visualised by ultraviolet shadowing and excised. RNA was removed from the gel slices by electroelution into 1 × TBE using a Schleicher & Schuell (Dassel, Germany) Biotrap. The RNA solution was dialysed extensively against water and concentrated using rotary evaporation under reduced pressure. Using these methods as much as 10 mg of pure RNA may be purified on as few as five gels. This amount represents one or two NMR samples and enough for several hundred crystallisation trials.

5' radioactive labelling of RNA

A phosphorylation reaction contains RNA (1 µM), 12 mM dithiothreitol, 12 µM [γ -³²P]ATP (3 Ci/mole, Amersham), 12 units per µl of RNasin (Promega), 1 unit per µl of T4 phage polynucleotide kinase and T4 polynucleotide kinase buffer supplied with the enzyme (New England Biolabs, NEB). A normal reaction is continued for 30 minutes at 37°C. The reaction is terminated by the addition of one half volume of formamide dye, and applied to a 15% polyacrylamide gel (400 mm × 200 mm × 0.5 mm) in 1 × TBE and 8 M urea.

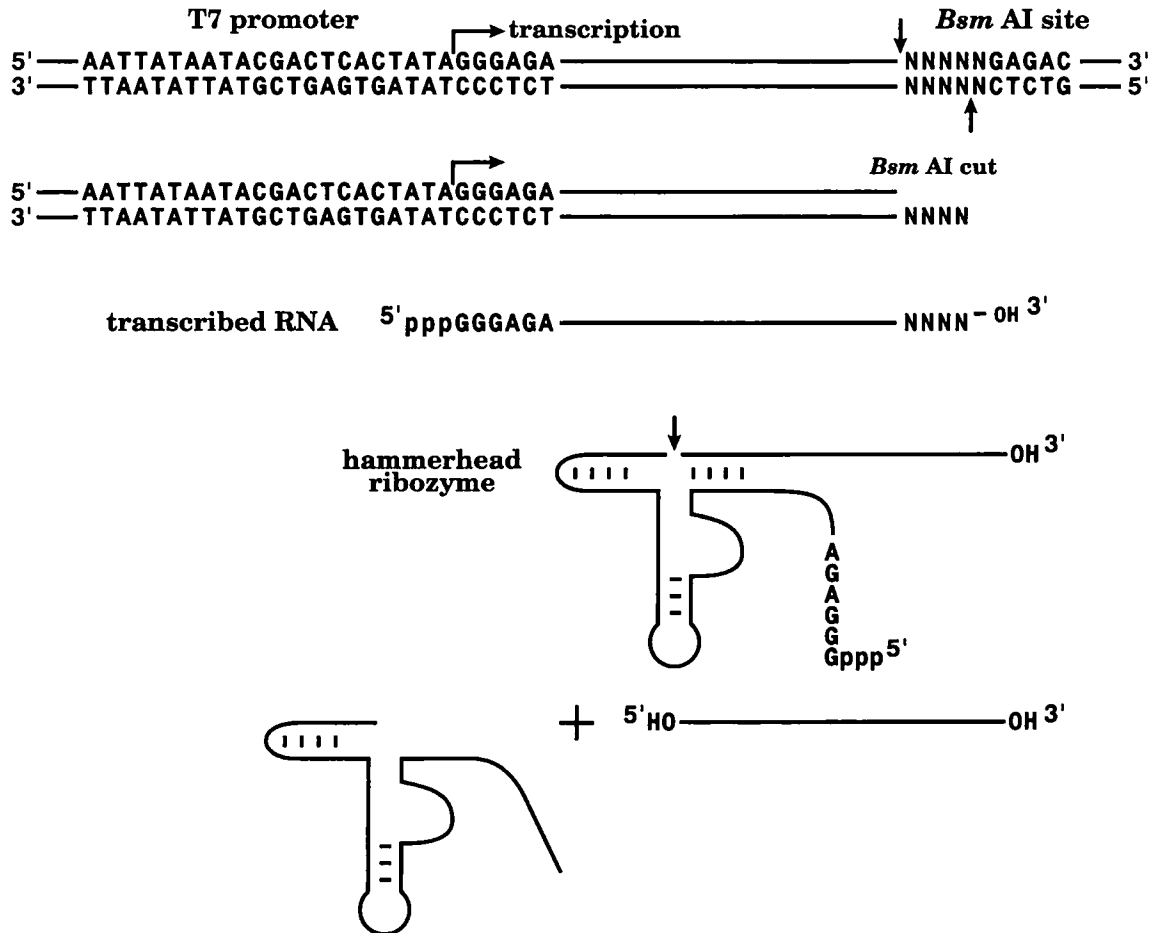


Figure 5(a) (caption opposite)

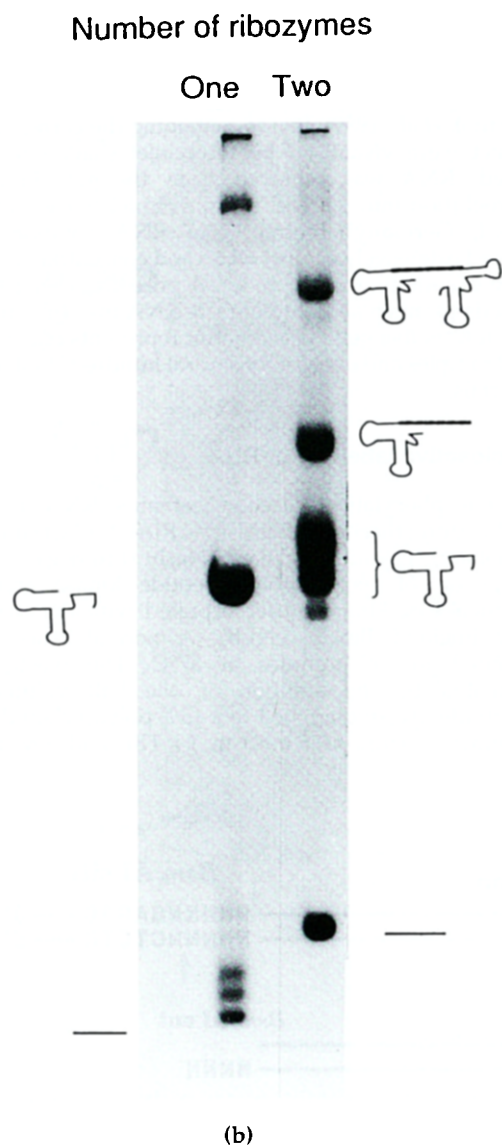


Figure 5. a, A strategy for producing RNA of any sequence by run-off transcription. A 5' *cis*-acting hammerhead ribozyme generates the 5' end of the RNA, and the restriction enzyme *Bsm*AI is used to linearise the plasmid template to generate the run-off transcription template. b, Transcription of two templates yielding different length RNAs. The left-hand lane shows production of a 24mer RNA by run-off transcription utilizing a 5' hammerhead ribozyme and 3' terminus generated by run-off transcription of a plasmid which had been linearized using the *Bsm*AI restriction enzyme. The right-hand lane shows transcription of a construct which yields a 27mer RNA through the action of 5' and 3' *cis*-acting hammerhead ribozymes. Note the homogeneity of the right-hand lane product compared to that produced by run-off transcription.

After electrophoresis and visualisation by autoradiography desired bands were excised by cutting the gel with a razor blade for further characterisation. Gel slices were cut into small pieces (around 1 mm cubes) and shaken in sterile water at 37°C in a sterile plastic tube, for four hours. The RNA solution is decanted and ethanol precipitated

by standard methods (Sambrook *et al.*, 1989). After centrifugation at 100,000 *g*, washing and drying, the RNA pellet is re-dissolved in diethyl pyrocarbonate (DEPC) treated water to a final concentration of 100 nM.

De-cyclisation of 2', 3' cyclic phosphate by sodium maleate treatment

RNA made by the double hammerhead ribozyme method was treated with sodium maleate at pH 2 in order to decyclise the 2', 3' cyclic phosphate which results from the cleavage at the 3' ribozyme. RNA at a concentration of 1 μ M was incubated with 0.8 M maleic acid/0.4 M NaOH (pH 2) for 80 minutes at 25°C after which the pH was increased to 5.5 by the addition of NaOH. A control incubation contained RNA at a concentration of 1 μ M in 0.8 M maleic acid/0.8 M NaOH (pH 5.5) for 80 minutes at 25°C. Tris buffer was added to a final concentration of 0.15 M and HCl was added to a final concentration of 10 mM, this brings the final pH to around 8.0 at which calf intestinal alkaline phosphatase may work. One unit of calf intestinal alkaline phosphatase was added and the mixture incubated at 55°C for 60 minutes. After phenol extraction and ethanol precipitation, the RNA was then phosphorylated by T4 phage polynucleotide kinase and [γ -³²P]ATP and a time-course of reaction performed (see Figure 3a).

Sequencing of RNA by iodine mediated phosphorothioate oxidation

RNA sequences are determined as described by Schatz *et al.* (1991). Sulphurs are incorporated at nucleotide specific positions by *in vitro* transcription reactions where one of the four ribose nucleotide solutions in the transcription is supplemented by its ribose nucleotide phosphorothioate analogue (Amersham) at 1/20 concentration. The concentration of iodine used in the sequencing reaction is 1 mM. The reactions are terminated by the addition of a dithiothreitol solution to a final concentration of 2 mM and loaded onto a 20% denaturing polyacrylamide gel after the addition of one-half volume of formamide dye. The construct which is shown in Figure 1a was transcribed in four separate reactions, each reaction incorporating a different nucleoside phosphorothioate analogue. The 36 nucleotide cleavage product of each reaction was purified by gel electrophoresis and phosphorylated at the 5' terminus in the presence of [γ -³²P]ATP and T4 polynucleotide kinase and the upper band of the product was gel purified for further characterisation. A small proportion of the RNA has a phosphorothioate analogue incorporated at each position. Iodine treatment of each of the four reactions produces nucleotide-specific partial cleavage (Schatz *et al.*, 1991).

Acknowledgements

It is a pleasure for the authors to acknowledge the help of Dr K. Taira for inspiration and helpful discussion, Dr W. Scott for communication of results prior to publication, C. McNally and C.-H. Teo for help and J. Fogg, T. Smith and R. Grenfell for oligonucleotide synthesis. One of the authors (S.R.P.) thanks Professor John Frink and Steve Bennett for many enlightened evenings. S.R.P. is supported by a Medical Research Council (MRC) studentship,

J.M.A. by an MRC training fellowship. This project was supported by the MRC and Human Frontier Science Programme.

References

- Anderson, P., Monforte, J., Tritz, R., Nesbitt, S., Hearst, J. & Hampel, A. (1994). Mutagenesis of the hairpin ribozyme. *Nucl. Acids Res.* **22**, 1096–1100.
- Beaucage, S. L. & Caruthers, M. H. (1981). Deoxynucleotide phosphoramidites—a new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Letters*, **22**, 1859–1862.
- Berzal-Herranz, A., Joseph, S. & Burke, J. M. (1992). In vitro selection of active hairpin ribozymes by sequential RNA-catalysed cleavage and ligation reactions. *Genes Dev.* **6**, 129–134.
- Berzal-Herranz, A., Joseph, S., Chowrira, B. M., Butcher, S. E. & Burke, J. M. (1993). Essential nucleotide sequences and secondary structure elements of the hairpin ribozyme. *EMBO J.* **12**, 2567–2574.
- Biou, V., Yaremchuk, A., Tkalalo, M. & Cusack, S. (1994). 2.9 Ångstrom crystal structure seryl-tRNA synthetase complexed with t-RNA^{Ser}. *Science*, **263**, 1404–1410.
- Buzayan, J. M., Gerlach, W. L. & Bruening, G. (1986a). Non-enzymatic cleavage and ligation of RNAs complementary to a plant-virus satellite RNA. *Nature*, **323**, 349–353.
- Buzayan, J. M., Gerlach, W. L., Bruening, G., Keese, P. & Gould, A. R. (1986b). Nucleotide sequence of satellite tobacco ringspot virus RNA and its relationship to multimeric forms. *Virology*, **151**, 186–199.
- Buzayan, J. M., Hampel, A. & Bruening, G. (1986c). Nucleotide-sequence and newly formed phosphodiester bond of spontaneously ligated satellite tobacco ringspot virus-RNA. *Nucl. Acids Res.* **14**, 9729–9743.
- Cameron, V. & Uhlenbeck, O. C. (1977). 3' Phosphatase activity in T4 polynucleotide kinase. *Biochemistry*, **16**, 5120–5126.
- Dahm, S. C. & Uhlenbeck, O. C. (1991). Role of divalent metal-ions in the hammerhead RNA cleavage reaction. *Biochemistry*, **30**, 9464–9469.
- Grodberg, J. & Dunn, J. J. (1988). OmpT encodes the *Escherichia coli* outer-membrane protease that cleaves T7-RNA polymerase during purification. *J. Bacteriol.* **170**, 1245–1253.
- Epstein, L. M. & Gall, J. G. (1987). Self-cleaving transcripts of satellite DNA from the newt. *Cell*, **48**, 535–543.
- Feldstein, P. A., Buzayan, J. W. & Bruening, G. (1989). Two sequences participating in the autolytic processing of satellite tobacco ringspot virus complementary RNA. *Gene*, **82**, 53–61.
- Francklyn, C. S. & Schimmel, P. (1990). Synthetic RNA molecules as substrates for enzymes that act on transfer-RNAs and transfer-RNA-like molecules. *Chem. Rev.* **90**, 1327–1342.
- Hampel, A. & Tritz, R. (1989). RNA catalytic properties of the minimum(–) sTRSV sequence. *Biochemistry*, **28**, 4929–4933.
- Hampel, H., Tritz, R., Hicks, M. & Cruz, P. (1990). Hairpin catalytic RNA model—evidence for helices and sequence requirement for substrate RNA. *Nucl. Acids Res.* **18**, 299–304.
- Harrison, S. C. & Sauer, R. T. (eds). (1994). Protein-nucleic acid interactions. *Curr. Opin. Struc. Biol.* **4**, 1–66.
- Haseloff, J. & Gerlach, W. L. (1988). Simple RNA enzymes with new and highly specific endoribonuclease activities. *Nature*, **334**, 585–591.
- Haseloff, J. & Gerlach, W. L. (1989). Sequences required for self-catalysed cleavage of the satellite RNA of tobacco ringspot virus. *Gene*, **82**, 43–52.
- Heus, H. A. & Pardi, A. (1991). Nuclear magnetic resonance studies of the hammerhead ribozyme domain, secondary structure formation and magnesium ion dependence. *J. Mol. Biol.* **217**, 113–124.
- Hutchins, C. J., Rathjen, P. D., Forster, A. C. & Symons, R. H. (1986). Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid. *Nucl. Acids Res.* **14**, 3627–3640.
- Koizumi, M., Iwai, S. & Ohtsuka, E. (1988a). Construction of a series of several self-cleaving RNA duplexes using synthetic 21-mers. *FEBS Letters*, **228**, 228–230.
- Koizumi, M., Iwai, S. & Ohtsuka, E. (1988b). Cleavage of specific sites of RNA by designed ribozymes. *FEBS Letters*, **239**, 285–288.
- Koizumi, M., Hayase, Y., Iwai, S., Kamiya, H., Inoue, H. & Ohtsuka, E. (1989). Design of RNA enzymes distinguishing a single base mutation in RNA. *Nucl. Acids Res.* **17**, 7059–7071.
- Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. (1987). Oligoribonucleotide synthesis using T7 RNA-polymerase and synthetic DNA templates. *Nucl. Acids Res.* **15**, 8783–8798.
- Nagai, K. (1992). RNA-protein interactions. *Curr. Opin. Struc. Biol.* **2**, 131–137.
- Nakamaye, K. L. & Eckstein, F. (1994). AUA cleaving hammerhead ribozymes—attempted selection for improved cleavage. *Biochemistry*, **33**, 1271–1277.
- Oubridge, C., Ito, N., Teo, C.-H., Fearnley, I. and Nagai, K. (1995). Crystallisation of RNA-protein complexes II. Crystallisation of the complex of U1A small nuclear ribonucleoprotein with its cognate RNA hairpin. *J. Mol. Biol.* **249**, 398–412.
- Oubridge, C., Ito, N., Evans, P. R., Teo, C.-H. & Nagai, K. (1994). Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature*, **372**, 432–438.
- Perriman, R., Delves, A. & Gerlach, W. L. (1992). Extended target site specificity for a hammerhead ribozyme. *Gene*, **113**, 157–163.
- Pley, H. W., Flaherty, K. M. & McKay, D. B. (1994). Three-dimensional structure of a hammerhead ribozyme. *Nature*, **372**, 68–74.
- Prody, G. A., Bakos, J. T., Buzayan, J. M., Schneider, I. R. & Bruening, G. (1986). Autolytic processing of dimeric plant satellite RNA. *Science*, **231**, 1577–1580.
- Rould, M. A., Perona, J. J., Soll, D. & Steitz, T. A. (1989). Structure of *Escherichia coli* glutaminyl t-RNA synthetase complexed with transfer-RNA^{Gln} and ATP at 2.8 Å resolution. *Science*, **246**, 1135–1142.
- Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjamy, A., Rees, B., Thierry, J. C. & Moras, D. (1991). Class II aminoacyl t-RNA synthetases—crystal structure of yeast aspartyl transfer RNA synthetase complexed with tRNA^{Asp}. *Science*, **252**, 1682–1689.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular cloning—A Laboratory Manual*, 2nd edit. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sampson, J. R. & Uhlenbeck, O. C. (1988). Biochemical and physical characterisation of an unmodified yeast phenylalanine transfer RNA transcribed *in vitro*. *Proc. Natl Acad. Sci. USA*, **85**, 1033–1037.

- Schatz, D., Leberman, R. & Eckstein, F. (1991). Interaction of *Escherichia coli* tRNA^{Ser} with its cognate aminoacyl-tRNA synthetase as determined by footprinting with phosphorothioate-containing tRNA transcripts. *Proc. Natl Acad. Sci. USA*, **88**, 6132–6136.
- Sheldon, C. C. & Symons, R. H. (1989). Mutagenesis analysis of a self-cleaving RNA. *Nucl. Acids Res.* **17**, 5679–5685.
- Snyder, L. (1983). In *Bacteriophage T4* (Mathews, C. K., Kutter, E. M., Mosig, G. & Berget, P. B., eds) pp. 351–355. American Society for Microbiology, Washington, DC.
- Soltis, D. A. & Uhlenbeck, O. C. (1982). Independent locations of kinase and 3'-phosphatase activities on T4 polynucleotide kinase. *J. Biol. Chem.* **257**, 11340–11345.
- Taira, K., Nakagawa, K., Nishikawa, S. & Furukawa, K. (1991). Constructing a novel RNA-transcript-trimming plasmid which can be used both *in vitro* in place of run-off and (G). free transcription and *in vivo* as multi-sequence transcription vectors. *Nucl. Acids Res.* **19**, 5125–5130.
- Uhlenbeck, O. C. (1987). A small catalytic oligoribonucleotide. *Nature*, **328**, 596–600.
- Usman, N., Ogilvie, K. K., Tiang, M.-Y. & Cedergren, R. G. (1987). Automated chemical synthesis of long oligoribonucleotides using 2'-O-silylated ribonucleoside 3'-O-phosphoramidites on a controlled glass support: synthesis of a 43 nucleotide sequence similar to the 3'-half molecule of an *Escherichia coli* formylmethionyl tRNA. *J. Am. Chem. Soc.* **109**, 7845–7854.
- Valegård, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J. & Liljas, L. (1994). Crystal structure of an RNA bacteriophage coat protein-operator complex. *Nature*, **371**, 623–626.
- Varani, G. & Tinoco, I. Jr (1991). RNA structure and NMR spectroscopy. *Quart. Rev. Biophys.* **24**, 479–532.

Edited by D. Draper

(Received 25 November 1994; accepted 8 March 1995)