RNA-catalysed synthesis of complementary-strand RNA

Jennifer A. Doudna & Jack W. Szostak

Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

The *Tetrahymena* ribozyme can splice together multiple oligonucleotides aligned on a template strand to yield a fully complementary product strand. This reaction demonstrates the feasibility of RNA-catalysed RNA replications.

THE recent identification of several catalytically active RNA molecules[1-6] has led to extensive speculation concerning the role of RNA in the origin of life[7-11]. A self-replicating RNA or related polynucleotide is thought to be the key intermediate in the evolution of living systems from prebiotic chemicals. The identification or design of an RNA replicase would therefore be helpful in establishing the feasibility of this pathway, and might also allow the construction of simple proto-cells in the laboratory. The protozoan *Tetrahymena* self-splicing intron[1,2,6] is of particular interest in this respect because of the variety of phosphoester transfer reactions that it can catalyse. The activities of this ribozyme, demonstrated in a series of experiments by Cech and coworkers, include that of a ribonuclease, phosphotransferase, acid phosphatase and RNA restriction endonuclease[12-14].

An enzyme capable of catalysing transesterification reactions on RNA substrates is potentially capable of catalysing RNA polymerization. Indeed, Zaug and Cech[12] and Been and Cech[15] have shown that the intron will catalyse limited polymerization of ribonucleotides onto a short primer annealed to a sequence within the intron. But the primer could be extended only to a maximum of ~15 nucleotides, and the nucleotides were added beyond the end of the template[15]. These experiments, although demonstrating the ability of the *Tetrahymena* ribozyme to polymerize RNA, also pointed out the problems to be overcome before replicase activity could be approached: first, the template must be on a separate molecule from the replicase; second, the polymerase activity must be template directed; and third, the nucleotide specificity that ensures accurate splicing must be overcome to allow copying of an arbitrary template sequence. Here we show that these three problems can be overcome, and that a modified version of the *Tetrahymena* ribozyme can catalyse the formation of an RNA molecule complementary to a template strand.

Catalysis on an independent template

The initial step in self-splicing (Fig. 1) is the attack of the 3′ hydroxyl of a free guanosine on a specific phosphate in a stem-loop of the intron referred to as P1. A phosphoester transfer reaction occurs in which the exon–intron junction is cleaved and the G becomes attached to the 5′ end of the intron. P1 is the first of a series of base-paired stems in the intron; the 3′ strand of P1 (the internal guide sequence) has been used as a primer-binding site for primer elongation experiments[12,15]. To overcome the requirement that a replicase act on a template that is a separate molecule, we synthesized the isolated P1 stem-loop (Fig. 1*a*) by transcription of a synthetic oligonucleotide, and designed a modified version of the ribozyme extending from stem P2 to P9 (Fig. 1*b*). We found that this enzyme RNA catalyses the site-specific attack of guanosine on the isolated P1 stem (Fig. 2), but that the K_m for free P1 was very high (>0.1 mM). This weak interaction probably reflects the fact that

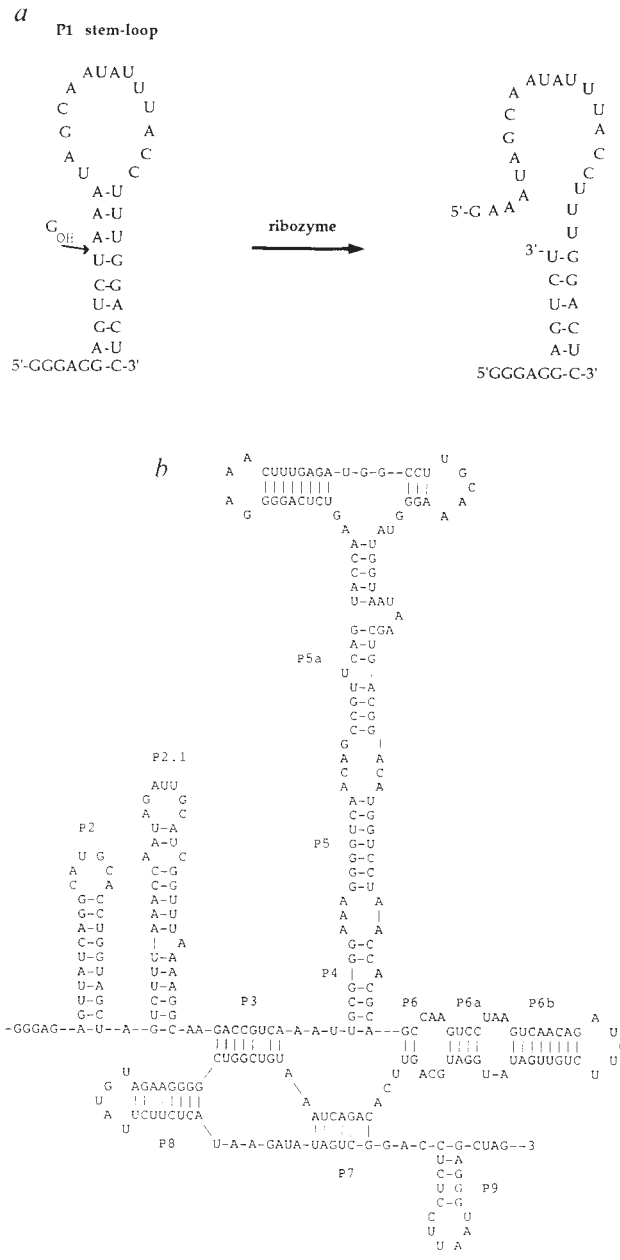

FIG. 1 *a*, Cleavage of an independent P1 substrate. The diagram illustrates the structure of the substrate used in this experiment and the guanosine cleavage reaction catalysed by the ribozyme. P1 is normally the first stem–loop of the intron, containing the 5′ exon-intron junction and the internal guide sequence; in our experiments it is a separate molecule. *b*, Sequence and secondary structure of the modified *Tetrahymena* intron used in this paper. This molecule is missing the P1 stem-loop, the 3′ stem–loops P9.1 and P9.2, and the 3′ intron–exon junction. RNA was synthesized by T7 RNA polymerase runoff transcription of plasmid pJD1100 digested with the restriction enzyme *Nhe*I, and purified either by electrophoresis on a 6% polyacrylamide, 7M urea gel, or by chromatography on a Sephadex G-75 spin column. In either case, the RNA was extracted with phenol, precipitated with ethanol, and stored in distilled water at −80 °C.

there are few sequence or size requirements for recognition of P1 by the core intron[16-18].

As a first step towards the assembly of oligonucleotides on an external template, we synthesized two RNA oligonucleotides that correspond to the products of guanosine attack on P1 (Fig. 3*a*). The sequences are different from that of wild-type P1, apart from the U·G base pair at the site of the guanosine cleavage reaction. The two oligonucleotides anneal to form a complex that consists of an RNA primer annealed to a partial hairpin RNA, with a guanosine residue extruded from the helix at the gap. Incubation of this complex with the modified ribozyme resulted in the extremely efficient regeneration of intact P1, with the release of free guanosine (Fig. 3*b*). This is the equivalent of the reverse of the reaction shown in Fig. 1*a*. The oligonucleotide ligation reaction was essentially complete after one hour, with about 250 turnovers of substrate per enzyme.

Spermidine overcomes sequence specificity

We have previously shown, in a different system, that only the wobble base pairs U·G and C·A in the intact P1 stem allow efficient guanosine attack[18]. To evaluate the base-pair requirements for the reverse reaction (oligonucleotide ligation), we synthesized four primers ending in either A, C, G or U, and four partial hairpins with either A, C, G or U opposite the last base of the primer (Fig. 4*a*). The 16 primer–hairpin combinations were then tested for ligation by the ribozyme (Fig. 4*b*). In contrast to the cleavage reaction, only the U·G and C·G base combinations allow efficient ligation. Ligation occurs to a lesser extent with the A·G combination.

The same set of primer–template combinations was tested for ligation by the intron under various reaction conditions. Because we have hypothesized that base-pair geometry at the reaction site is critical[18], it seemed possible that subtle changes in the structure of either P1 or the ribozyme might allow ligation to occur with other base-pair combinations. One of the conditions tested was the addition of 5 mM spermidine to the reaction (Fig. 4*c*). This led to the efficient ligation of all the substrate complexes with either Watson–Crick or wobble (U·G, C·A or A·G) base pairs at the ligation junction. A set of 10 related polyamines was tested for similar effects, and none was as effective as spermidine, although putrescine and spermine work to a lesser extent. Spermidine overcomes the inherent sequence specificity of the ribozyme, and should, in principle, allow the production of a faithful (that is, entirely Watson–Crick) copy of a template.

Template-directed oligonucleotide ligation

Previous work had shown that the loop of the P1 stem did not seem to be important in the guanosine-attack reaction. This suggested that we could simply dispense with the P1 loop in the reverse reaction, and try to ligate together two short oligonucleotides aligned on a longer third oligonucleotide which would act as a template. Of the two shorter oligonucleotides, the 5′ one is referred to as the primer, and the 3′ one as the ligator. The ligator oligonucleotide begins with a guanosine residue that is not paired with the template, just as in the P1 regeneration experiments described above. Several different primer, ligator and template combinations were designed: three substrates contained a U·G at the ligation junction, whereas four others contained the Watson–Crick base pairs at the junction. In each case, the ribozyme catalysed the ligation of the primer and the ligator in a template-dependent manner (Fig. 5). The number of turnovers per enzyme molecule ranged from 100 in a 15-min reaction for the U·G substrates, to 30–100 in 60 min for the C·G, G·C and A·U substrates, to a low of 5 in 60 min for a U·A substrate. The ligated products were characterized by T1 digestion; in all cases the expected fragments were observed. These experiments show that the ribozyme can catalyse oligonucleotide ligation which is independent of sequence.

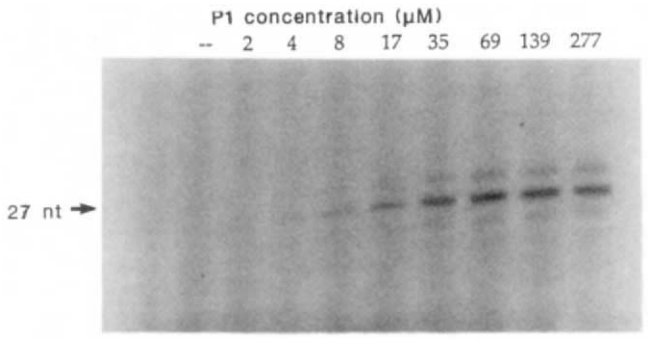

FIG. 2 Cleavage of a P1-like substrate by the modified *Tetrahymena* ribozyme. P1 RNA was prepared by T7 transcription of a synthetic oligodeoxynucleotide[19], followed by purification on a 20% acrylamide, 7 M urea gel. Enzyme reactions contained 10 mM NH_4Cl, 20 mM $MgCl_2$, 30 mM Tris-HCl pH 7.4, 1 mM aurin trichloroacetic acid, 0.2 μM enzyme (prepared as in Fig. 1*b*), 200 μM [α-^{32}P]GTP, and the indicated concentration of P1 RNA, in a 5-μl reaction volume. Reactions were incubated at 58 °C for 20 min, then stopped by the addition of an equal volume of 90% formamide, 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.2% bromphenol blue and xylene cyanol. Aliquots were electrophoresed on a 15% acrylamide, 7 M urea gel; the gel was then dried and autoradiographed. Lanes (left to right): no enzyme, no substrate, increasing substrate concentration. The major band is the expected product of the reaction (see text and Fig. 1*a*). The minor bands are due to 3′ heterogeneity of the substrate, as determined by 5′-end-labelling with polynucleotide kinase.

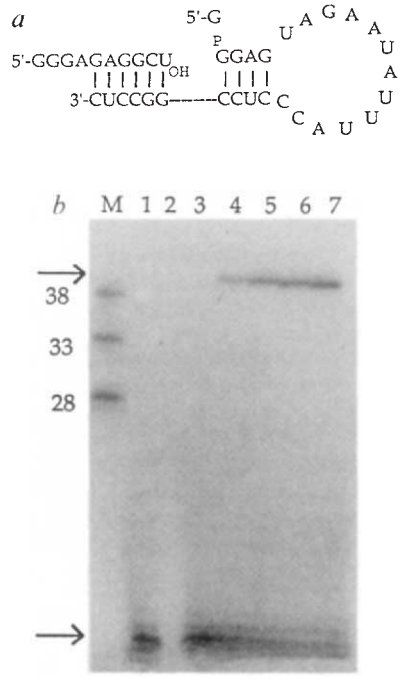

FIG. 3 *a*, Substrate ribo-oligonucleotides for P1 regeneration. RNAs corresponding to the products of a P1 cleavage reaction were prepared and purified as described above, so that the enzymatic reversal of the cleavage reaction could be followed. The 5′ primer oligonucleotide was internally labelled by including 0.5 μCi per μl of [α-^{32}P]GTP in the transcription reaction *b*, Time course of P1 regeneration. Reactions contained 10 mM NH_4Cl, 20 mM $MgCl_2$, 30 mM Tris HCl pH 7.4, 1 mM aurin trichloroacetic acid, 0.2 μM enzyme and 50 μM of each RNA substrate oligonucleotide in a total volume of 5 μl. Reactions were incubated at 58 °C, and stopped by the addition of formamide/dye solution as above. Aliquots were electrophoresed on a 20% acrylamide, 7 M urea gel; the gel was dried and autoradiographed. Lane M, DNA size markers; lane 1, no template oligonucleotide, 60 min reaction; lane 2, no primer oligonucleotide, 60 min reaction; lanes 3–7, complete reactions incubated for 0, 15, 30, 45 and 60 min, respectively. Bottom arrow, labelled primer; top arrow, ligated product.

Ligation of aligned multiple oligonucleotides

To see if the ligation reaction explored above could be used to assemble a longer complementary-strand RNA, three sets of templates and complementary oligonucleotides were made (Fig. 6*a*). In the first set, four identical oligonucleotides align on one long template, with U·G base pairs at each ligation junction. With equimolar template and ribozyme in the reaction, products corresponding to the ligation of two, three or all four oligonucleotides are visible within 5 min of incubation (Fig. 6*b*). The ratio of full-length to shorter products increases over time, with concomitant release of free GMP. With smaller amounts of enzyme, the reaction proceeds more slowly.

We considered the possibility that the longer products in the above reaction were generated by a ligation of two oligonucleotides, dissociation from the template and subsequent reannealing to the template with the original ligator in the first position on the template. Longer oligonucleotides could then form by repeated cycles of a reaction very similar to two-oligonucleotide ligation. This model was eliminated by using a template on which four different oligonucleotides could anneal (Fig. 6*a*), so that each ligation reaction would have to take place at a different position on the template. The course of the reaction was similar to the first multiple ligation (Fig. 6*c*). The sequence of the full-length product was confirmed by dideoxy sequencing with reverse transcriptase.

In the two experiments described above, the product strand is not fully complementary to the template strand, because of the U·G base pairs at the ligation junctions. We therefore synthesized a template on which five small RNAs could anneal, such that each of the four ligation junctions consisted of a

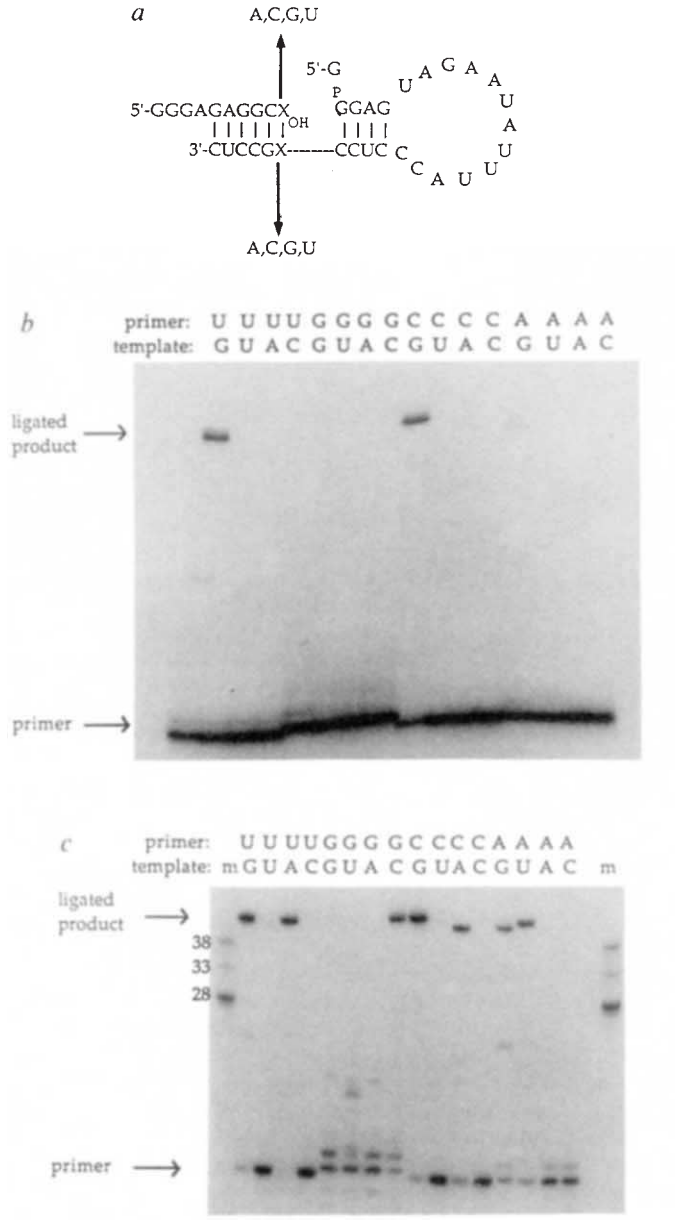

FIG. 4 Effect of the base pair at the ligation junction on P1 regeneration. *a*, We synthesized four primers and four templates, differing at the indicated position. These sets of oligonucleotides can be combined to yield P1 regeneration substrates with any base–base combination at the ligation junction. *b*, P1 regeneration reactions. Reactions were as described above, except that 20 μM oligonucleotides were used, and incubations were at 58 °C for 60 min. Each reaction contained the indicated base–base combination at the ligation junction. The reaction works well only with the U·G and C·G combinations. *c*, Effect of spermidine on P1 regeneration. Reactions were as above, except for the addition of 5 mM spermidine, which allows U·G, C·A, A·G and all Watson–Crick combinations to ligate efficiently. Ligated products with different sequences have slightly different mobilities.

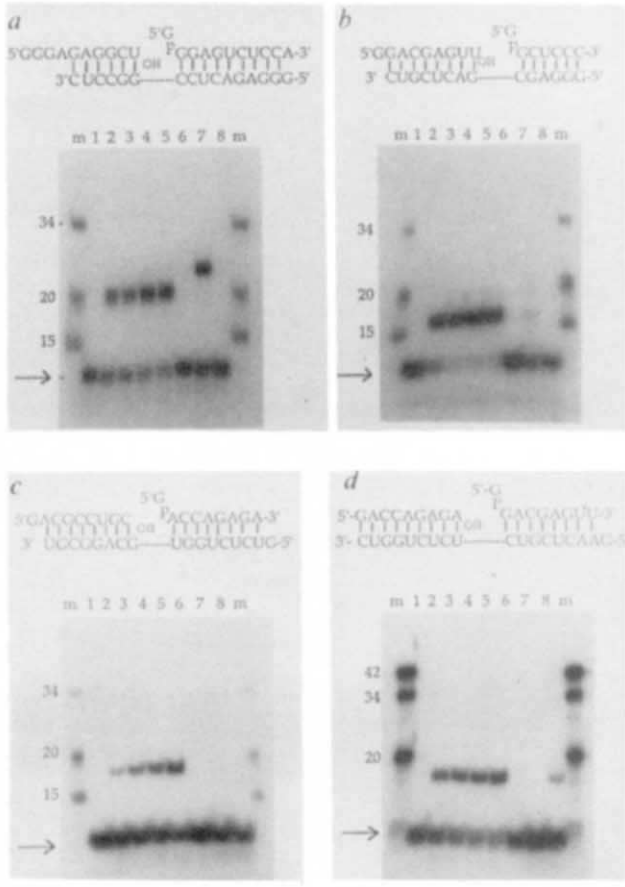

FIG. 5 Template-directed oligonucleotide ligations. Four examples of the ligation of two oligonucleotides aligned by a template oligonucleotide are shown. RNAs were prepared and purified as described above, except that the primer RNAs were labelled with [γ-^{32}P]ATP and polynucleotide kinase. Reaction conditions were as in Fig. 4, with 0.2 μM enzyme, 20 μM oligonucleotides and 5 mM spermidine. Reactions were stopped by the addition of 3 vol 9.3 M urea, 33 mM EDTA. To denature product template complexes completely, reaction mixes were separated by electrophoresis on a 20% polyacrylamide, 90% formamide gel. Lanes M, RNA olgionucleoide size markers; lanes 1–5, time course, 0, 15, 30, 45 and 60 min, respectively; lanes 6, no template oligonucleotide, 60-min reaction; lanes 7, no ligator oligonucleotide, 60 min reaction; lanes 8, no enzyme (except *d*; lane 7, no enzyme; lane 8, no ligator), 60 min reaction. In each case, the arrow represents the position of the primer. *a*, *b*, Substrate complexes with a U·G base pair at the ligation junction; *c*, *d*, complexes with Watson–Crick base pairs at the junction.

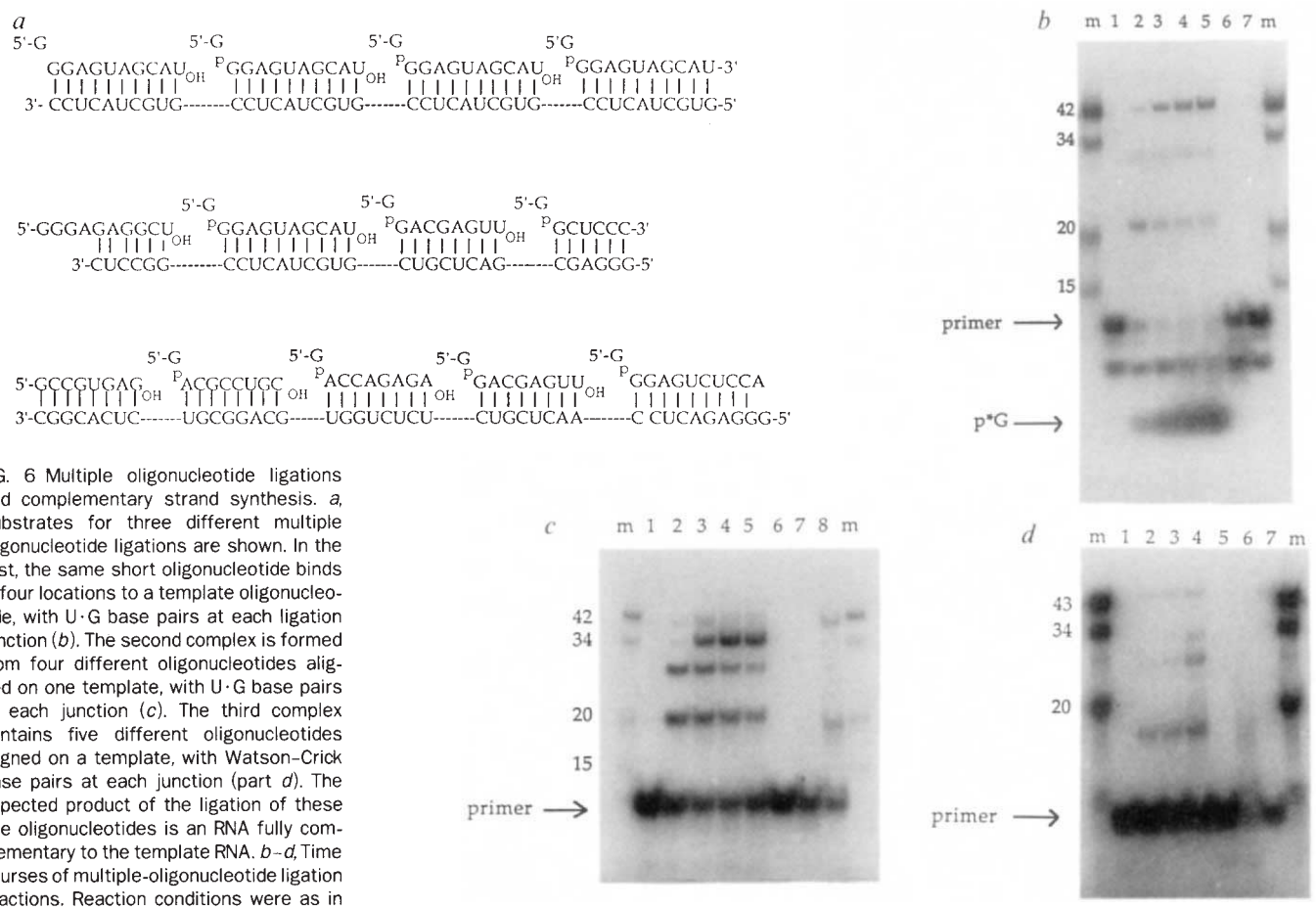

FIG. 6 Multiple oligonucleotide ligations and complementary strand synthesis. *a*, Substrates for three different multiple oligonucleotide ligations are shown. In the first, the same short oligonucleotide binds in four locations to a template oligonucleotide, with U·G base pairs at each ligation junction (*b*). The second complex is formed from four different oligonucleotides aligned on one template, with U·G base pairs at each junction (*c*). The third complex contains five different oligonucleotides aligned on a template, with Watson–Crick base pairs at each junction (part *d*). The expected product of the ligation of these five oligonucleotides is an RNA fully complementary to the template RNA. *b*–*d*, Time courses of multiple-oligonucleotide ligation reactions. Reaction conditions were as in Fig. 5, except that both enzyme and substrate were at 5 μM. Aliquots of each reaction were electrophoresed on 20% polyacrylamide gels prepared in 90% formamide to promote complete denaturation of the samples. Lanes marked 'm' contain end-labelled RNA size markers. *b*, Ligation of four oligonucleotides shown in Fig. 6*a*, top. Lanes 1–5, complete reactions incubated for 0, 5, 30, 60 and 90 min, respectively; lane 6, no template, 90 min; lane 7 no enzyme, 90 min. *c*, Ligation of four oligonucleotides shown in Fig. 6*a*, middle. Lanes 1–7, as in part *b*; lane 8, no second oligonucleotide. *d*, Ligation of five oligonucleotides shown in Fig. 6*a*, bottom. Lanes 1–4, 0, 30, 60 and 120 min, respectively; lane 5, no enzyme, 120 min; lane 6, no template, 120 min; lane 7, no second oligonucleotide, 120 min.

different Watson–Crick base pair. We observed ~5% ligation to full-length product, with larger amounts of shorter products also accumulating (Fig. 6*d*). The reaction was completely dependent on template and enzyme.

Discussion

The ligation reaction that we have characterized differs from all previously described reactions catalysed by group I introns in that base pairing of substrate RNAs to the intron itself is not required. For example, in the polymerase experiments reported by Cech *et al.*[15], the primer to be extended is bound to the internal guide sequence, which is itself covalently joined to the catalytic domain of the intron. In the experiments described here, the template that aligns the oligonucleotides to be ligated is a separate RNA molecule that does not interact with the enzyme by base pairing. One consequence of this difference is much weaker interaction of our substrates with the enzyme. We think that this weakened interaction is the simplest explanation for the rapid turnover we see in the ligation reaction; after ligation, the double-stranded product and the single guanosine residue are rapidly released from the enzyme, allowing another cycle of catalysis to begin.

In principle, a ligation reaction similar to that described here could have been used by a crude primordial replicase, and could be used now in the design of an RNA replicase. The rapid turnover that we see is important in this respect, because many catalytic cycles will be required for the replication of a long template. Our experiments show that a wide range of template sequences can be used, and the addition of spermidine to the reaction allows the formation of a fully complementary product. But several refinements will be necessary to achieve autocatalytic replication in the laboratory. The efficiency of full-length strand synthesis must be increased, the possible deleterious effects of template secondary structure must be explored and overcome, and a way to separate product and template strands must be found. These problems can now be approached experimentally; their solution should allow the design of a self-replicating RNA molecule. □

Received 14 April; accepted 8 May 1989.

1. Cech, T. R., Zaug, A. J. & Grobowski, P. J. *Cell* **27,** 487–496 (1981).
2. Kruger, K. *et al. Cell* **31,** 147–157 (1982).
3. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. *Cell* **35,** 849–857 (1983).
4. Van der Veen, R. *et al. Cell* **44,** 225–234 (1986).
5. Garriga, G. & Lambowitz, A. M. *Cell* **39,** 631–641 (1984).
6. Cech, T. R. *Science* **236,** 1532–1539 (1987).
7. Sharp, P. A. *Cell* **42,** 397–400 (1985).
8. Darnell, J. E. & Doolittle, W. F. *Proc. natn. Acad. Sci. U.S.A.* **83,** 1271–1275 (1986).
9. Gilbert, W. *Nature* **319,** 618 (1986).
10. Cech, T. R. *Proc. natn. Acad. Sci. U.S.A.* **83,** 4360–4363 (1986).
11. Joyce, G. F. *Nature* **338,** 217–224 (1989).
12. Zaug, A. J. & Cech, T. R. *Science* **231,** 470–475 (1986).
13. Zaug, A. J., Been, M. D. & Cech, T. R. *Nature* **324,** 429–433 (1986).
14. Zaug, A. J. & Cech, T. R. *Biochemistry* **25,** 4478–4482 (1986).
15. Been, M. D. & Cech, T. R. *Science* **239,** 1412–1416 (1988).
16. Waring, R. B., Towner, P., Minter, S. J. & Davies, R. W. *Nature* **321,** 133–139 (1986).
17. Been, M. D. *et al. Cold Springer Harb. Symp. quant. Biol.* **LII,** 147–157 (1987).
18. Doudna, J. A., Cormack, B. P. & Szostak, J. W. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
19. Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. *Nucleic Acids Res.* **15,** 8783 (1987).