The Birth of Recombinant RNA Technology

A newly developed technique promises to allow the production of large quantities of any chosen RNA by hooking it to a replication vehicle

Published in the current issue of the *Journal of Molecular Biology* is a technique that promises to revolutionize the laboratory manipulation of RNA molecules, a development that is reminiscent in a small way at least of the advent of recombinant DNA technology, or genetic engineering.* Indeed, the new technique involves the construction of recombinant RNA; that is, the stitching together of RNA chains from different sources.

The power of the technique, which has been nurtured by Fred Russell Kramer, Donald Mills, and Eleanor Miele of the Institute of Cancer Research, Columbia University, is the production of large quantities of RNA that might otherwise be difficult to obtain in workable amounts. The trick is to confer on any RNA molecule the ability to be replicated easily and in great volume in test tube reactions. If the system fulfills its promise—it is still at a preliminary stage—it will be possible not only to manipulate elusive RNA's, such as those of the tiny infectious particles, the viroids, but also to manufacture commercially significant batches of messenger RNA from which valuable proteins might be generated.

Recombinant RNA technology, in other words, might prove to be yet another highly effective tool in the hands of molecular biologists, and a further extension of molecular biology in the marketplace. The Office of Science and Technology Development at Columbia is preparing patent applications, which will be filed early in 1984.

Test tube synthesis of RNA had been possible for quite some time, so long as the gene coding for the molecule was available: the system involves transcription of RNA copies from a DNA template in the presence of the right enzymes and cofactors. The problem with transcription, however, is that the quantity of product RNA possible is somewhat limited. The transcribing enzyme—an RNA polymerase—merely churns out RNA copies from the given quantity of template, so that the buildup of product is linear. By contrast, the recombinant RNA technology being developed at Columbia should give an exponential accumulation of product, so that 1 milligram of product RNA is generated from an input of 100 nanograms of template in 1 hour. This is unheard of productivity in any transcription system.

The secret engine behind the recombinant RNA technology is no secret at all. It is an enzyme known since 1965 that in nature reproduces the RNA genome of a small bacterial virus known as Qβ: the enzyme is called Qβ replicase. When this enzyme makes the complementary copy of its RNA template, the product strand immediately becomes a potential template itself. Each round of replication therefore doubles the number of available templates, thus yielding the exponential growth of product. Because replication doubles the number of templates at each step, the process is usually characterized as autocatalytic.

> If the replication system works, Qβ replicase will be elevated . . . to being a molecular workhorse in a major industry.

This startling performance of Qβ replicase hasn't escaped the notice of molecular biologists since its discovery two decades ago, but it has not been capitalized on. The difficulty is that the enzyme will replicate only Qβ RNA, or certain related RNA's. The reason for this extreme specificity is not difficult to find. A bacterial cell infected by Qβ is a seething universe of all kinds of RNA molecules and, without specificity to its own RNA, the replicase would wastefully spend most of its time reproducing foreign molecules.

What is useful to the virus has been an insurmountable challenge to biologists—until now. Batteries of biochemical manipulations were applied to the replicase in more or less futile attempts to subvert the enzyme's tight specificity. Various foreign templates—including ribosomal RNA, viral RNA, and even messenger RNA from higher organisms—were copied on occasion, but never in any quantity and never in autocatalytic fashion.

The Columbia team succeeded where others had failed, simply by linking the foreign template to Qβ-like RNA sequences so that the enzyme would accept the package for replication. The recombinant RNA sobriquet refers to the combining of chosen foreign RNA with the Qβ-like RNA. Actually, the success was not simple at all: it derived from Kramer and Mills's long experience with the enzyme and its normal targets and the skilled determination of Miele in developing the systems for stitching the various RNA fragments together.

Kramer has worked with Qβ replicase and its family of templates for 14 years, and Mills a little longer. Both had the late Sol Spiegelman as a guiding light. And both became intimate with idiosyncracies of this curious enzyme system.

One curiosity is the composition of the enzyme itself, which is built from four separate protein subunits. When Qβ RNA infects a cell it codes for just three proteins: a coat protein, a maturation protein, and a protein that joins with three host proteins to form the replicase complex. This combination of viral and host proteins might, it seems, be the source of trouble and intrigue that has dogged researchers for many years.

The enzyme is relatively easily isolated for laboratory use, and is exceedingly stable once prepared. Researchers soon realized, however, that if they incubated this purified enzyme under conditions suitable for replication, but in the absence of added RNA template, after a lag period there would grow up a population of small RNA chains replicated by the enzyme. The size and sequence of the chains produced depended on the exact reaction conditions, so that now there is a series of more or less characterized Qβ variants: they are known as midivariant, minivariant, microvariant, nanovariant, plus one recently discovered species with as yet no name.

These variants are small molecules, ranging from 77 nucleotides in the smallest to 22 in the largest. Compared with Qβ RNA itself, which measures 4220 nucelotides in length, they are *very* small. Where did they come from? Rüdiger Luce, in Manfred Eigen's laboratory in Göttingen, Germany, has hypothesized that they arise de novo, assembled consistently as a result of highly

J. Mol. Biol. **171**, 281 (1983).

specific pressures of natural selection in the test tube.

A more popular view, and one apparently borne out by recent work of David Hill and Thomas Blumenthal at Indiana University, is that the enzyme is usually contaminated with vanishingly small amounts of specific templates for the variants. The variant templates very probably originate in the host, which might use them in self defense against Qβ infection, a kind of molecular decoy.

Whatever the source of the variants, they have provided researchers with a tool for studying the way Qβ replicase makes copies of its template. For replication to proceed normally, it turns out, the template must possess both a specific recognition sequence, or binding site, and an initiation sequence, which is a cytidine-rich region at the 3′ end of the molecule. Hence the reason why foreign RNA's are ignored by the enzyme.

Both Kramer and Mills have been interested in the detailed interactions between protein and RNA in the replication reaction. One striking feature of the template is the extensive amount of secondary structure—hairpins, stems, and loops—in all the variants and the "parent" Qβ. Moreover, this highly configured form then folds further to produce, in the midivariant at least, a more or less spherical structure.

Replication of this type of structure must be a complicated operation, which Qβ replicase clearly performs smoothly. One of the important points about the replication is that as the product chain is formed it does not base-pair with the template, with which, of course, it has a complementary sequence. There may be something in the enzyme activity that prevents this. Or, more likely, secondary structure in the extant and growing chains impedes it.

Mills and Kramer have found that replication proceeds in rapid bursts, the enzyme moving at high speed over certain stretches of the strand and pausing at others. One might have guessed that the replicase would have halted when it stumbled against the beginning of a hairpin stem. In fact, the stutters occur just after the replication of a sequence that can assemble into a hairpin stem in the product strand. It seems likely that an ordered sequence of folding and refolding in energetically different states is an important part of the replication process for Qβ RNA and its relatives.

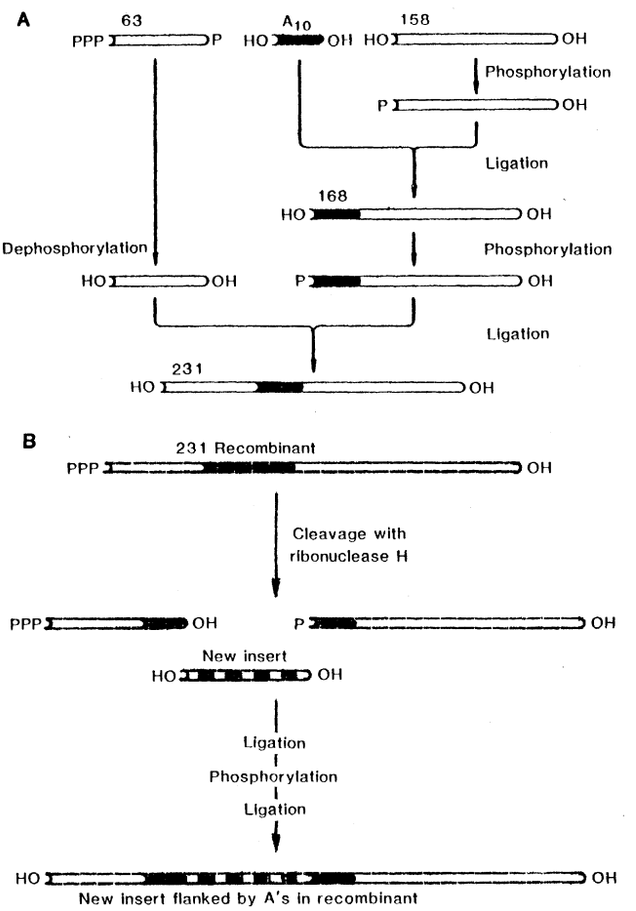
Kramer's interest in the replication of small RNA molecules got him invited to a European Molecular Biology Organization conference on viroids, held in Munich in 1978. Viroids, which measure between 270 and 380 nucleotides long, are the smallest known infectious agents by an order of magnitude. As they are economically of great importance and scientifically mystifying, there is a great deal of interest in studying them closely. But it is difficult to handle the small amounts of material that are usually available in the laboratory. Kramer recalls that at the meeting he was told that, interesting though it was, the Qβ replicase system wasn't of any use to the

The First Recombinant RNA Molecule

Construction of the first recombinant RNA molecule involved the use of bacteriophage T4 RNA ligase, an enzyme that does not like highly structured molecules and therefore promised—and delivered—much trouble in dealing with Qβ variant RNA. Site-directed cleavage of the midivariant RNA was achieved by masking most of the molecule—apart from nucleotides 61 to 64—by a modified complementary DNA: ribonuclease T1 then cut between positions 63 and 64, yielding two fragments with end groups as shown. The 5′ end of the 158 fragment had to be phosphorylated before reaction with the decaadenylic acid insert: the reaction was driven by the presence of a very large excess of decaadenylic acid (see A).

Before ligation of the new 168 fragment to the shorter arm, the 3′ end of the 63 fragment was dephosphorylated and the 5′ end of the 168 fragment phosphorylated. The final ligation was again driven by a large excess of the smaller fragment: the yield of this step was 0.2 percent. Once the recombinant molecule was isolated, however, the power of the technique was demonstrated by rapid amplification by Qβ replicase, which gives a 100,000-fold increase in a 10-minute reaction (see B).

Construction of this small recombinant RNA molecule not only demonstrated that the technique worked, but it also provided a more convenient vehicle for the assembly of further recombinants. Treatment of the molecule with ribonuclease H in the presence of decathymidylic acid, which hybridizes with the run of adenylic acids in the template, causes cleavage in the insert and yields two fragments in which the phosphate and hydroxyl residues are in the correct orientation for more convenient ligation of another foreign RNA.—**R.L.**



viroid researchers because of its tight specificity, so why didn't he try to make it of more general application.

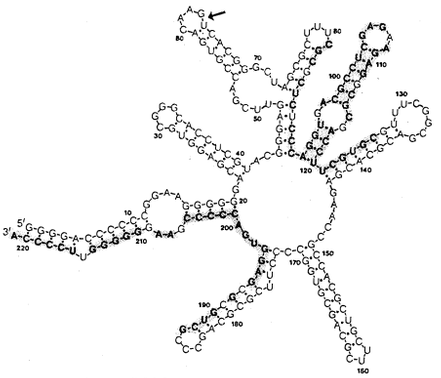
Such a thing had been in the back of Kramer's mind for quite some time, so he determined to try it. Offered to Miele as a thesis project as an alternative to something else that would be sure to work and yield a good degree with sanity intact, she accepted the challenge.

As the midivariant was the most thoroughly characterized member of this little RNA zoo, it was chosen as the vehicle for the new recombinant RNA technology. The task was to find a site on the molecule that would avoid the binding and initiation sequences and would, when cut, cause minimal disruption to the secondary structure. A foreign RNA would then have to be enzymatically stitched into the "gap." The test would then be, would the replicase accept the recombinant as a normal template? The fact that the earlier biochemical attempts to get replicase to copy foreign RNA without benefit of the correct binding and initiation sites had succeeded, albeit at minimal levels, encouraged the belief that foreign RNA sequences per se were not absolutely lethal to the enzyme. But the apparent complexity of secondary structure folding and refolding during replication of normal $Q\beta$-like templates was cause for crossed fingers.

Previous experience in the laboratory with exposure of midivariant to an RNA-cutting enzyme—ribonuclease T1—indicated a loop on one of the hairpins as a likely insertion spot for the foreign RNA. The cut was made here by combining the midivariant with a DNA copy of itself that had been altered at the chosen location: the duplex was then completely base-paired, apart from the cleavage site. Application of the T1 enzyme thus yielded two fragments of the RNA molecule, one measuring 63 nucleotides, the other 158.

As typically happens in these circumstances, the chemical groups at the ends of the fragments—phosphate and hydroxyl—were the wrong way round, and Miele had to reverse them before the foreign RNA—a stretch of ten adenylic acid residues—could be inserted (see box). Once inserted, the resultant recombinant RNA molecule was offered as a template to the replicase, which copied it at a rate no different from the intact midivariant. Recombinant RNA technology was born, if still an infant.

This, then, was the primary route to recombinant RNA molecules, and is the technique outlined in the *Journal of Molecular Biology* paper. The Columbia team has now added a second, and potentially more powerful, route, which makes use of recombinant DNA technology.



Midivariant RNA

Shaded areas through positions 81–126 and 187–221 are nearly identical with positions 84–129 and 4186–4220 of Qβ *RNA: the 5' sequences are the binding site; the 3', the initiation site. Arrow shows cleavage site.*

Mills's interest in the mechanics of replication encouraged him to modify the midivariant template structure, by adding or deleting hairpins for instance. Changes in the process would give insights into the structural requirements of the system. It is much easier to manipulate DNA than RNA molecules, because of the availability of the scores of highly specific restriction enzymes that are the basis of recombinant DNA technology. The obvious thing to do, therefore, was make a DNA copy of the midivariant RNA, and then perform the molecular manipulations with the great precision and facility allowed by restriction enzymes. Placed in a transcription system (that is, in a plasmid with a promoter upstream), the modified DNA can be copied back into RNA. This approach has indeed proved to be extremely powerful in probing the intricate structural interactions during replication.

It was but a small conceptual step—but a giant practical leap—from there to the second route to recombinant RNA molecules. Once a DNA copy of midivariant is emplaced in the plasmid, it is relatively easy to cleave the variant sequence and insert a selected DNA. Transcription of this construct yields a ready-made recombinant RNA, which potentially can then be amplified to large quantities by the $Q\beta$ replicase system. This would make available for study any RNA for which a gene can be isolated, even if the RNA itself has proved too difficult to prepare.

Whether recombinant RNA technology will become potent and widely applicable or simply remain an arcane tool of the few depends on how faithfully and efficiently $Q\beta$ replicase copies the recombinant RNA molecules. As the enzyme's normal substrate is a very large molecule (4220 nucleotides), size is unlikely to be a potential problem. Structural features of the artificial template are more problematical, however.

The enzyme might actually require the highly ordered secondary structure seen in $Q\beta$ RNA and its variants, perhaps for keeping template and product strands apart; in which case relatively unstructured messenger RNA's might not be accepted for replication. Moreover, highly ordered artificial templates might result in steric hindrance, thus impeding access to the binding and initiation sites by the enzyme. Kramer is optimistic in the face of this uncertainty and says the only thing to do is try.

In addition to the initial successful replication of the decaadenylic acid insert reported in the *Journal of Molecular Biology*, the Columbia team now has strong indications that the system works with a 210 nucleotide sequence from adenovirus and the 5*S* ribosomal RNA (120 nucleotides) from *Escherichia coli*. As this fragment of adenovirus RNA is relatively unstructured, it appears that extensive secondary structure is not an absolute requirement for enzyme activity. A transfer RNA from *Sulfolobus solfataricus* has so far proved recalcitrant, though is by no means a confirmed failure yet. A possible difficulty here is that, as one of the replicase subunits is a transfer RNA–binding host protein, there might be specific interaction between this subunit and the transfer RNA in the artificial template, which blocks normal replication. The first messenger RNA to be put to the test almost certainly will be globin messenger, partly because it is readily available.

In their current paper Kramer and his colleagues list several possible benefits of being able to manufacture large quantities of selected RNA at will. These include the production of highly labeled hybridization probes, the construction of templates for studying RNA processing mechanisms, and the manipulation and study of genomic RNA's, such as those of viroids and certain viruses. It is of course the prospect of large-scale synthesis of messenger RNA's that has been exciting a good deal of commercial interest in the work, because of the possibility of laboratory-controlled production of valuable proteins.

If the replication system works, $Q\beta$ replicase will be elevated from the relative obscurity of academic interest in just a couple of university laboratories to being a molecular workhorse in a major industry.—**Roger Lewin**