A Replicating RNA Molecule Suitable for a Detailed Analysis of Extracellular Evolution and Replication

(phage Q8 RNA/antiparallel complementary strands/Darwinian selection)

D. L. KACIAN, D. R. MILLS, F. R. KRAMER, AND S. SPIEGELMAN

Institute of Cancer Research and Department of Human Genetics and Development, College of Physicians and Surgeons, Columbia University, New York City, N.Y. 10032

Contributed by S. Spiegelman, August 14, 1972

ABSTRACT The aim of the present study is to make available a replicating molecule of known sequence. Accordingly, we sought a molecule that has the following properties: (a) replicates in vitro in a manner similar to phage Q8 RNA; (b) produces antiparallel complementary strands that can be separated from one another; and (c) is small enough to yield its sequence with reasonable effort.

We report here the isolation of a replicating RNA molecule that contains 218 nucleotides and possesses the other features desired for a definitive analysis of the replicating mechanism. Despite its small size, this molecule can mutate to previously determined phenotypes. It will, therefore, permit the precise identification of the base changes required to mutate from one phenotype to another in the course of extracellular Darwinian selection experiments.

Some years ago we isolated a template-specific, RNA-instructed RNA polymerase (1) from Escherichia coli infected with phage Q8, an RNA bacteriophage (2, 3), and established that the purified enzyme (replicase) could mediate the autocatalytic synthesis (4) of biologically competent and infectious viral RNA (5). When RNA from a mutant phage Q8 was used to start the reaction with wild-type enzyme, mutant RNA was produced (6), proving that the RNA was the instructing agent in the synthesis. These findings provided the first opportunity to perform extracellular Darwinian experiments on replicating nucleic-acid molecules under conditions that simulate certain aspects of precellular evolution, when environmental discrimination operated at the level of the gene rather than the gene product.

It was not a priori obvious that such experiments would generate useful information. Much depended on the number of different ways the environment could distinguish one nucleic acid molecule from another and decide which was superior. Happily, the experiments revealed an unexpected wealth of phenotypic differences that replicating nucleic acid molecules can exhibit and from which the environment can select.

The first mutant (V-1) was isolated under selection pressures designed to encourage rapid completion of replication (7). To satisfy the imposed demands, large stretches of genetic information unnecessary for replication were discarded leading to the emergence of a variant containing only 550 of the original 3600 nucleotides found in phage Q8 RNA. Variant molecules of this class were then transferred under conditions conferring advantages on molecules that could institute rapid synthesis when the reaction was initiated with as little as a single strand. In this manner we isolated V-2, a strain that can be cloned (8). By variation of the specific nature of the selection pressures, other variants were derived from V-2 that could multiply at severely limiting concentrations of riboside triphosphates, or in the presence of such inhibitory agents as tubercidin triphosphate (9), ethidium bromide, and proflavin (10).

It was apparent from even the limited set of successful examples that the number of identifiable mutant molecules possessing prespecified phenotypes was restricted only by the ingenuity exercised in devising the appropriate selective conditions. These experiments resolved an interesting dilemma of precellular evolution. They permit one to see, at least dimly, how selective pressures could have forced replicating nucleic acids to greater length and complexity (11), a necessary prelude to the invention of cells and subcellular components as aids in replication. In addition to their evolutionary interest, abbreviated molecules that have eliminated segments unnecessary for replication provide ideal objects for the detailed chemical analysis of the replicative process.

It became obvious after several years that no matter how amusing, continuing Darwinian-selection experiments would not bring us any closer to the kind of understanding we wanted of either the evolutionary process being observed or of the replication mechanism. A truly profound exploitation of the information inherent in this system required that we start with a molecule of completely known sequence. To attain this goal, we sought a replicating molecule with the following properties: (a) It should replicate in a manner recognizably analogous to that of phage Q8 (12-17), (b) It should be possible to separate one of the two complementary strands for independent sequencing, and (c) It should be small enough to yield its absolute sequence without inordinate time and effort.

The variants we had thus far isolated satisfied the first two criteria, but were all about 550 nucleotides long. Our ultimate aim required that the sequence be known with absolute certainty. Even with the elegant sequencing techniques introduced by Weissmann and his colleagues (18), it seemed desirable to ease our task by searching for smaller replicating molecules. This decision was encouraged by the report (19) of "6S" replicating RNA molecules in cells infected with phage Q8. On examination, the "6S" species turned out to be a mixture of similar, but not identical, molecules. One of these "minivariants" (MV-1) was cloned and examined in detail with respect to its 5'-terminal sequence and replicative features (20). Unfortunately, these molecules possessed structural
characteristics that made it difficult to recognize replicative stages or to separate plus from minus strands.

We continued our search for a suitable candidate and succeeded in isolating the molecule we report on here. It is 218 nucleotides long and satisfies all the other criteria desired as an object of intensive study. With its aid, we can hope to specify in chemical detail its mechanism of replication and to specify the base changes required to mutate from one phenotype to another.

MATERIALS AND METHODS

Isolation of Phage Qβ Replicase. Phage Qβ replicase was purified through the diethylaminoethyl (DEAE)-cellulose column step (21). The enzyme was further purified by agarose gel (20) and phosphocellulose chromatography (22). The fraction corresponding to the "heavy" component described by Elkhom and Spiegelman (23) was used for all the work reported here.

Synthesis of Labeled Ribonucleoside Triphosphates. [α-32P]-Ribonucleoside triphosphates were prepared as detailed by Haruna et al. (1). [γ-32P]-Ribonucleoside triphosphates were prepared by the method of Glynn and Chappell (24). Trinitiated ribonucleoside triphosphates were purchased from Schwarz BioResearch, Inc., Orangeburg, N.Y.

RESULTS

Since the "6S" species were isolated from infected cells and found as contaminants of replicase preparations, other variants were looked for in the same source. In the course of examining products of reactions run without added RNA templates, a replicating molecule was obtained that appeared to possess the features we desired. This variant was labeled midivariant (MDV-1) to indicate that it is smaller than the 350 nucleotide V-class of variants, and is some 90 nucleotides longer than the group of "6S" "minivariants."

Characterization of reaction products

After protein removal, the products of a reaction templated by either phage Qβ RNA or V-2 RNA resolve into three distinct structures when examined by acrylamide gel electrophoresis (16): a fast-moving component consisting of single strands, a double-stranded structure of intermediate mobility, and a slow-moving multistranded complex containing double-stranded regions along with single-stranded tails. Fig. 1 shows that a reaction initiated by MDV-1 RNA yields a simpler pattern when the protein-free products are analyzed on acrylamide gels. Again one observes three distinct structures (labeled I, II, and III) moving with different velocities in the gel.

The properties of peaks I and II are examined in Fig. 2 with the products of a late reaction containing very little of the peak III component. Fig. 2A shows the behavior and mobility of the untreated structures, compared with the "6S" midivariant (MV-1). From their mobilities, it would appear that MV-1 is smaller than either I or II of MDV-1 and that peak II has about twice the molecular weight of peak I. If the mixture of I and II is subjected to heat denaturation, peak II is almost eliminated, whereas peak I is unaffected (Fig. 2B), suggesting that fraction II is a double-stranded structure and that fraction I is single-stranded. These conclusions are further confirmed by the effects of nuclease treatment on a mixture of fractions I and II shown in Fig. 2C. Here we see that peak II is left undisturbed, whereas peak I is destroyed.

Fig. 3 examines the response to RNase of 32P-labeled peak III isolated from a previous gel. 3H-labeled MDV-1 product was included as an internal marker to locate the structures. It will be noted from the repeated electrophoresis of untreated peak III (Fig. 3A) that manipulation of this fragile structure gives rise to components I and II. Exposure to ribonuclease completely converts the multistranded peak III to the double-stranded peak II and eliminates the single-stranded peak I.

The responses of peaks I, II, and III of MDV-1 to denaturation and ribonuclease are in complete accord with responses obtained with the analogous structures observed in reactions directed by phage Qβ RNA (16) and the V-class of abbreviated mutants (15).

Nearest-neighbor analysis of the double-stranded structure (peak II)

If peak II corresponds to the double-stranded structure observed in our previous studies (25), it should be an antiparallel...
As in the case of Base enzymatic digests. The self-annealed, for initiation mixture that there this duplex of the plus and minus strands of complementary strands. After self-annealing, peak II was electrophoretically isolated from each reaction, hydrolyzed, and analyzed as described in the legend to Table 1. The observed equivalent molar frequencies of antiparallel dinucleotide complements (Table 1) are clearly in support of an antiparallel duplex composed of complementary strands. It will be noted that there was excellent agreement between the alkali and enzymatic digests. The total absence of ApU (Table 1) means that this variant RNA is devoid of the AUG codon required for initiation of translation into protein.

**Base composition of the MDV-1(+) strand**

As in the case of phage Qβ and V-2, this always corresponds to the plus strand if the synthesis is extensive. We have accordingly designated as MDV-1(+) the strand that predominates in the peak I component of late reactions.

To determine the base composition of the MDV-1 plus strand and of the double-stranded duplex, an extensive synthesis was performed in a reaction containing all four ribosidetriphosphates labeled at the same specific radioactivity with ³²P in the α position. The purified product was self-annealed and the double strands were separated from the single plus strands electrophoretically. The two preparations were then hydrolyzed and analyzed as described in the legend of Table 1. It is evident from the results recorded in Table 2 that both plus and minus strands possess complementary-type base compositions, with guanosine being equal to cytidine and adenosine being equal to uridine. Further, both strands have an unusually large proportion (70%) of guanosine and cytidine. All of these findings predict the possibility of a highly ordered secondary structure.

**The 5′ terminus and the direction of synthesis of MDV-1**

To determine the 5′ terminus, we proceeded as in our previous studies with phage Qβ RNA (16), V-2 RNA (15), and "8S" RNA (20). Four reactions directed by MDV-1 RNA were run, each of which contained one of the ribonucleoside triphosphates labeled with ³²P in the γ position. The extent of chain synthesis was monitored internally in each case by including [³H]CTP. Of the four [γ-³²P]ribonucleoside triphosphates, only [γ-³²P]guanosine triphosphate labeled multivariant RNA. Fig. 4 shows the electrophoretic pattern of the MDV-1 RNA products labeled internally with [³H]CTP and terminally by [γ-³²P]GTP. That ³²P labeling is only terminal was demonstrated, as we had done previously (16), by showing that digestion with T1 RNAse releases [γ-³²P]guanosine tetraphosphate as the only radioactive residue. The distribution of ³²P and ³H in the three peaks of Fig. 4 shows clearly that both

![Graphical representation of the effect of melting and nuclease treatment on multivariant RNA.](image)

**Fig. 2.** Effect of melting and nuclease treatment on multivariant RNA. (A) 130 ng of [³²P]RNA reaction product were co-electrophoresed with [³H]minivariant RNA (20) marker (=130 nucleotides long). Multistranded structures were not present in sufficient quantity to be detected in this late reaction product. (B) 130 ng of RNA were melted at 100° for 2 min in 30 μl of 3 mM EDTA, then frozen in a dry ice–ethanol bath. The sample was thawed in the presence of the marker and immediately analyzed by electrophoresis. (C) 130 ng of RNA were digested with 20 μg of RNase A for 10 min at 37° in 30 μl of 400 mM NaCl–50 mM Tris·HCl (pH 7.4)–3 mM EDTA. In order to inactivate the nuclease, 5 μl of diethylpyrocarbonate (Naftone, Inc., New York, N.Y.) was added and incubation was continued for 30 min. Undissolved diethylpyrocarbonate was removed by centrifugation. The supernatant was mixed with the marker and analyzed by electrophoresis.

![Graphical representation of the effect of nuclease treatment on multistranded RNA.]
TABLE 1. Nearest-neighbor analysis of double-stranded midivariant RNA

<table>
<thead>
<tr>
<th>Dinucleotide pairs</th>
<th>Molar dinucleotide frequencies</th>
<th>Alkali digestion</th>
<th>Enzymatic digestion</th>
</tr>
</thead>
<tbody>
<tr>
<td>CpA, UpG</td>
<td>0.030, 0.037</td>
<td>0.037, 0.036</td>
<td></td>
</tr>
<tr>
<td>ApC, GpU</td>
<td>0.057, 0.059</td>
<td>0.056, 0.062</td>
<td></td>
</tr>
<tr>
<td>CpU, ApG</td>
<td>0.022, 0.030</td>
<td>0.006, 0.006</td>
<td></td>
</tr>
<tr>
<td>UpC, GpA</td>
<td>0.071, 0.077</td>
<td>0.069, 0.068</td>
<td></td>
</tr>
<tr>
<td>GpG, CpC</td>
<td>0.119, 0.119</td>
<td>0.119, 0.115</td>
<td></td>
</tr>
<tr>
<td>ApA, UpU</td>
<td>0.028, 0.028</td>
<td>0.029, 0.033</td>
<td></td>
</tr>
<tr>
<td>GpC</td>
<td>0.113</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cpg</td>
<td>0.141</td>
<td>0.139</td>
<td></td>
</tr>
<tr>
<td>ApU</td>
<td>0.001</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>UpA</td>
<td>0.010</td>
<td>0.009</td>
<td></td>
</tr>
</tbody>
</table>

Single-stranded and double-stranded midivariant RNA prepared from reaction mixtures containing the appropriate [γ-32P]ribonucleoside triphosphate(s) were separated by gel electrophoresis. The RNA was self- annealed before electrophoresis. After elution from the gels, each RNA preparation was analyzed by two methods: digestion with alkali and digestion with a mixture of ribonucleases.

For alkaline digestion, the RNA was dissolved in 10 μl of 0.2 M KOH and incubated at 37° for 17 hr. The sample was then neutralized with HCl, and the mononucleotides were resolved by electrophoresis in 50 mM sodium citrate (pH 3.5) at 4000 V on Whatman 3 MM paper (27). The radioactive digestion products were detected by autoradiography, cut from the paper, and counted in 2,5-bis-(5-tert-butylbenzoxazolyl)-thiophene (BBOT) in toluene.

For nuclease digestion, the RNA was melted at 100° in 3 mM EDTA (pH 7.0) for 2 min and then dissolved in 5 μl of 20 mM sodium acetate (pH 4.5), containing 200 μg each of RNase A, RNase T1, and RNase T2. RNase A was obtained from Worthington Biochemical Corp., Freehold, N.J. and RNase T1 and T2 were obtained from Calbiochem, San Diego, Calif. The samples were incubated at 37° for 30 min, then spotted directly onto Whatman 3 MM paper and separated by high-voltage electrophoresis.

Molar dinucleotide frequencies were calculated by the procedure of Josse et al. (26).

plus and minus strands of MDV-1 begin with [γ-32P]GTP at their 5' ends. This finding also establishes that the direction of synthesis of both plus and minus strands of MDV-1 is in the 5' to 3' direction, again in agreement with our earlier findings with phage Qβ (16) and V-2 (15).

Size of single-stranded MDV-1

There are two ways (16) of determining the length of single-stranded RNA molecules that can be synthesized in vitro.

![Fig. 4. Polyacrylamide gel electrophoresis of midivariant RNA labeled by [γ-32P]GTP and [3H]CTP. None of the other three [γ-32P]ribonucleoside triphosphates was able to label midivariant RNA. 460 ng of the RNA shown above were melted at 100° for 2 min in 8 μl of 3 mM EDTA, then digested with 1 μg of RNase T1 for 30 min at 37° in 10 μl of 10 mM Tris-HCl (pH 7.4)–3 mM EDTA. The digestion mixture was analyzed by high-voltage electrophoresis at pH 3.5. Guanosine tetraphosphate was identified as the only radioactive residue. Both single-stranded and double-stranded RNA (peaks I and II) are labeled equally well by each isotope, indicating that [γ-32P]GTP is incorporated into the 5' end of both complementary RNA strands. The chain length of single-stranded midivariant RNA was determined by comparison of the ratio of [γ-32P]GTP incorporated into the end of the molecule to [3H]CTP incorporated internally. If, as in the present case, the 5' terminus is known, one can compare the number of moles of [γ-32P]GTP and the number of moles of [3H]CTP incorporated internally in the peak I region of Fig. 4. Taking into account the molar frequency of cytosine residues (Table 2), one can readily calculate the number of nucleotides; this turns out to be 230. The same calculation performed on the peak II material yields a similar value, proving that both plus and minus strands have pppGp at their 5' terminus.

The other procedure is to compare the electrophoretic mobility of midivariant RNA in acrylamide gels with internal RNA markers of known size, as is done in Fig. 5. The mobility observed for MDV-1 RNA corresponds to a chain length of...

![Fig. 5. Determination of the chain length of midivariant RNA by its relative electrophoretic mobility. Midivariant RNA was coelectrophoresed with RNA markers of known chain length on a 4.8% polyacrylamide gel. The RNAs were isotopically labeled. The log of the chain length was plotted against the distance each RNA migrated into the gel. The distance midivariant RNA migrated corresponded to a chain length of 213 nucleotides.]

TABLE 2. Nucleotide composition of midivariant RNA

<table>
<thead>
<tr>
<th>Nucleotide</th>
<th>Single-stranded RNA</th>
<th>Double-stranded RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Alkali</td>
<td>RNase</td>
</tr>
<tr>
<td>Guanosine</td>
<td>0.345</td>
<td>0.346</td>
</tr>
<tr>
<td>Cytidine</td>
<td>0.345</td>
<td>0.336</td>
</tr>
<tr>
<td>Adenosine</td>
<td>0.160</td>
<td>0.156</td>
</tr>
<tr>
<td>Uridine</td>
<td>0.150</td>
<td>0.161</td>
</tr>
</tbody>
</table>

Analysis was performed as described in the legend to Table 1.
213, in reasonable agreement with the number obtained from the ratio of the 5’ terminus to total chain length. It may be noted that by actual count (Mills, D. R., Kramer, F. R., & Spiegelman, S., manuscript in preparation), the absolute number of nucleotides is 218.

DISCUSSION

The length of the MDV-1 RNA molecule we describe here is only 5% of that of phage Qβ RNA, and since the initiating AUG triplet is absent, it is untranslatable into protein. In other respects the two replicating molecules have much in common. Both are accepted by phage Qβ replicase, implying that MDV-1 has retained the structure used by phage Qβ replicase to recognize and replicate phage Qβ-like RNA. Both RNAs direct the synthesis of complements and both give rise to similar multistranded, double-stranded, and single-stranded structures in the protein-free product.

Our goal was to isolate a molecule that replicates in a manner similar to phage Qβ RNA and that was small enough to yield the sequences of its complementary strands. This goal has been achieved. Using existing techniques with suitable modifications, we have in independent experiments established the sequences of both the plus and minus strands of MDV-1 RNA. We prefer to delay the discussion of the relation between MDV-1 and phage Qβ RNAs until the sequence data appear in detail (Mills, Kramer, and Spiegelman, in preparation). At present, we may note that the first eleven nucleotides from the 5’ end are identical and the first thirty are nearly so. We have shown in preliminary experiments (Kramer, F. R., Mills, D. R., Nishihara, T., Cole, P. E., & Spiegelman, S., manuscript in preparation) that, small as it is, MDV-1 RNA yields mutants with a prespecified phenotype under appropriate selective conditions.

The results described provide an experimental system that should illuminate in chemical detail the nature of the interaction between the enzyme and a replicating template, the mechanism of replication, and finally what base changes occur in mutating from one phenotype to another.

We thank Mr. Ben Tabenkin and his staff at Hoffman-LaRoche, Nutley, N.J., for their skilled assistance in the preparation of phage Qβ-infected E. coli. This work was supported by the National Science Foundation Research Grant GB-17251X1, National Cancer Institute Research Grant CA-02332, and American Cancer Society Postdoctoral Fellowship PF-538.