Sequence Analysis of RNA Species Synthesized by Qβ Replicase without Template†

Christof K. Biebricher* and Rüdiger Luce

Max Planck Institute for Biophysical Chemistry, D-3400 Göttingen, Federal Republic of Germany

Received November 5, 1992; Revised Manuscript Received February 16, 1993

ABSTRACT: Qβ replicase amplifies certain short-chained RNA templates autocatalytically with high efficiency. In the absence of extraneously added template, synthesis of new RNA species by Qβ replicase is observed under conditions of high enzyme and substrate concentrations and after long lag times. Even under identical conditions, different RNA species are produced in different experiments. The sequences of several independent template-free products have been determined by cloning their cDNAs into plasmids by a novel cloning procedure. Their nucleotide chain lengths are small, ranging from 25 to about 50 nucleotides. While their primary sequences are unrelated except for the invariant 5′-terminal G and 3′-terminal C clusters, their tentative secondary structures show a common principle: both their plus and minus strands have a stem at the 5′ terminus, while the 3′ terminus is unpaired. Direct accumulation of sufficient quantities of early template-free synthesis products by Qβ replicase is prevented by the inherent irreproducibility of the synthesis process and by the rapid change of the products during amplification by evolution processes, but large amounts of such RNA can be synthesized in vitro by trancription from the cDNA clones. RNA species produced in template-free reactions replicate much more slowly than the optimized RNA species characterized previously. These experimental results illustrate how biological information can be gained in small bits by trial and error.

The mechanism of template-instructed RNA synthesis by Qβ is well understood (Biebricher et al., 1983, 1984, 1985). Surprisingly, RNA is often also synthesized when no template is added to the incorporation mix. The most plausible and in many cases also correct explanation for this synthesis is amplification of residual RNA contaminants from various sources. However, the reaction can also be observed when all efforts have been taken to avoid contamination by RNA, albeit only under special reaction conditions and after long lag times (Sumper & Luce, 1975; Biebricher et al., 1981a,b). It has been excluded by careful kinetic studies that this synthesis is due to residual contamination of the enzyme preparation or the buffers by RNA that is replicated by Qβ replicase; furthermore, it was shown that high concentrations of enzyme and triphosphates and low ionic strength of the incubation mixture are required for this reaction to occur, while synthesis under other conditions is strictly template-dependent. In the first phase of the template-free reaction, nucleoside triphosphates are slowly condensed to form oligonucleotides of a form $pppGpR(pN)_n$ where n is in the range of 5–20, and in smaller concentrations up to $n = 100$ (Biebricher et al. 1986). Once a replicating sequence, no matter how poorly it may replicate, is produced, it is amplified and optimized. As one would expect for noninstructed processes, the products are irreproducible; the properties of different isolates differ from experiment to experiment even when the isolates are synthesized under identical conditions. RNA templates are optimized by recombination of RNA molecules (Munishkin et al., 1988, 1991; Biebricher & Luce, 1992) and mutations; RNA recombination may also play an important role in the still unknown early steps of template-free synthesis. In this paper, we show the sequences of several RNA species generated by Qβ replicase without template. The sequences were obtained from cDNA clones of the RNA products by a new method; the RNA can be reproduced from these cDNA clones by transcription with T7 RNA polymerase in vitro. The chain lengths of early template-free synthesis products are in the range of 30–45 nucleotides; while their primary sequences are not related, their secondary structures show significant similarities.

MATERIALS AND METHODS

Materials. Qβ replicase devoid of replicatable RNA impurities was isolated from an *Escherichia coli* strain carrying the replicase gene on a plasmid (constructed by Dr. M. A. Billeter, Zürich) by a method described previously (Sumper & Luce, 1975; Biebricher et al., 1986). Poly(A) polymerase has been prepared from *E. coli* K12 according to Sippel (1973). T7 RNA polymerase was isolated from an overproducing strain (Davanloo et al., 1984), a kind gift from Dr. F. W. Studier, by the method of Grodberg and Dunn (1988). Modified T7 DNA polymerase devoid of 3′-exonuclease activity (Tabor & Richardson, 1987) was obtained from USB.

Phosphorylation of Primers. The oligonucleotide primers were synthesized by the phosphite method in a Pharmacia Gene Assembler and purified according to the instructions of the manufacturer. Five nanomoles of oligonucleotide and 10 nmol of [γ-^{32}P]ATP were dissolved in 50 μL of kinase buffer (50 mM Tris·HCl (pH 7.9), 10 mM $MgCl_2$, and 10 mM DTT) and incubated with 50 units of T4 polynucleotide kinase for 30 min at 37 °C; after addition of 40 nmol of ATP, incubation was continued for 10 min. One microliter of 0.5 M EDTA was added, and the oligonucleotides were isolated by gel exclusion chromatography. The oligonucleotides used for first-strand cDNA synthesis was pGAATTCTAGG-GATCCATTTTTTTGGG, while for second-strand cDNA synthesis pTAATACGACTCACTATAGGG and pTAT-AGTGAGTCGTATTAAGCT in an equimolar mixture were annealed and ligated. In some experiments, a mixture of annealed and ligated pTAATACGACTCACTATTGGG and pATAGTGAGTCGTATTAAGCT was used in addition.

† This work has been supported by the Max Planck Society and by the European Community (PSS*0396).

0006-2960/93/0432-4848$04.00/0

Production and Isolation of the RNA Templates. Qβ buffer contained 50 mM Tris·HCl buffer (pH 7.5), 10 mM $MgCl_2$, 0.1 mM dithioerythritol, 10% glycerol, and the indicated concentrations of nucleoside triphosphates and salt. Standard conditions were 0.5 mM ATP, CTP, GTP, and UTP and 100 nM Qβ replicase. Template-free synthesis requires very long incubation times (Biebricher et al., 1981b); the experimental conditions are indicated. The RNA was isolated by phenol extraction and alcohol precipitation and purified by gel chromatography.

Limited Polyadenylation of RNA. Thirty to seventy-five picomoles of RNA (1–2.5 μg single- or double-stranded RNA) was dissolved in 50 μL of polyadenylation buffer (50 mM Tris·HCl (pH 7.9), 0.25 M NaCl, and 50 μM [^{3}H]ATP) and incubated with 5 units of poly(A) polymerase for 2 min at 37 °C. The reaction was started with 0.5 μL of 1 M $MgCl_2$, and the incubation continued for 5 min at 37 °C. The reaction was stopped with 2.5 μL of 0.5 M EDTA, and the RNA was isolated by phenol extraction and ethanol precipitation. On average, 10 A-residues were attached to the 3′ termini of single- and double-stranded RNA under these conditions.

Cloning of RNA. Twenty picomoles of polyadenylated double-stranded RNA and 100 pmol of first-strand primer were dissolved in 40 μL of H_2O, heated for 1.5 min to 100 °C to melt the double strands, and quickly cooled in dry ice mixture. Five microliters of 10× retrobuffer (final concentrations: 50 mM Tris·HCl, pH 8.3; 10 mM $MgCl_2$; and 10 mM DTT) and 2 μL of 0.025 M dNTP were added, and the volume was adjusted to 50 μL; the mixture was incubated with 25–50 units of AMV reverse transcriptase for 60 min at 42 °C. The cDNA (15–30% of the RNA input) was collected by phenol extraction and EtOH precipitation. The RNA template was hydrolyzed by heating incubation in 200 μL of 20% piperidine for 3 h at 50 °C. After lyophilization, the residue was dissolved in 20 μL of water and purified by EtOH precipitation and gel chromatography or gel electrophoresis. The cDNA was heated to 100 °C for 1 min and rapidly cooled. Fifty picomoles of phosphorylated second-strand primers was added, and the mixture was adjusted to 50 mM Tris·HCl (pH 7.5), 10 mM $MgCl_2$, 1 mM DTT, and 20 μM each dATP, dCTP, dGTP, and dTTP in a volume of 50 μL. The mixture was incubated with 0.5 μL of modified T7 DNA polymerase for 10 min at 4 °C and then for 10 min at 37 °C. The polymerase was destroyed by heating for 15 min at 65 °C, and the DNA was incubated for 1 h at 37 °C with 5 units each of *Hin*dIII and *Bam*HI. The DNA was collected by phenol extraction and EtOH precipitation and purified by Sepharose 4B chromatography or gel electrophoresis. The cDNA was ligated into 200 ng of *Bam*HI/*Hin*dIII-cut vector pUC18 DNA (Yanish-Perron et al., 1985) by standard methods (Maniatis et al., 1982) and used for transformation of *E. coli* HB101 made competent by the method of Hanahan (1983). Plasmids were purified according to the method of Birnboim (1983), replacing the ribonuclease step by ethidium bromide–CsCl equilibrium centrifugation.

Transcription by T7 RNA Polymerase. The transcription mixture contained 50 mM Tris·HCl buffer (pH 7.5); 10 mM $MgCl_2$; 1 mM dithiothreitol; 1.25 mM each ATP, CTP, GTP, and UTP, 0.5 mg/mL plasmid DNA digested with restriction endonuclease *Bst*XI; and 50 μg/mL T7 RNA polymerase. Incubation was for 30 min at 37 °C. For determination of the template activity of RNA, an aliquot of the transcript was used after appropriate dilution without further purification. Larger amounts of RNA were prepared from cDNA MNV-11 clones by the method of Jahn et al. (1991). While transcription from cloned cDNA MNV-11 plus strand gave a homogeneous mixture with correct ends, transcription from the corresponding minus strand clone produced RNA with 2–10 additional A residues at its 3′ terminus (Milligan & Uhlenbeck, 1989). Addition of 175 mM NaCl to the transcription buffer reduced the yield by one-half but improved the homogeneity of the RNA transcripts considerably. MNV-11 obtained by transcription had physical, chemical, and replication properties indistinguishable from MNV-11 obtained by replication.

Sequence Analysis. Sequences were determined by the dideoxy method (Sanger et al., 1977) using Sequenase (USB) according to the protocol of the supplier. Computer manipulation of the sequences (searching for sequence homologies in the GenBank or the EMBL bank, folding the sequences to secondary structure, mapping, etc.) were done with the GCG package of the University of Wisconsin on the VAX of the Gesellschaft für Wissenschaftliche Datenverarbeitung, Göttingen.

RESULTS

Procedure of Template-Free Synthesis. A prerequisite of observing template-free synthesis is the availability of Qβ replicase preparations which are devoid of RNA templates. Enzymes prepared according to the different methods of Sumper and Luce (1975), Hill and Blumenthal (1983), and Bauer et al. (1989) proved to be devoid of RNA templates, but rigorous precautions must be taken to prevent introduction of RNA contaminants from other sources. Pipets should be reserved for this kind of work and never be used for pipetting RNA; solutions must be made in rooms where RNA is not handled in large amounts; and use of glassware and pH-meter electrodes must be avoided. The absence of impurities from buffers and replicase solutions can be readily measured by incubation of replicase and triphosphate solutions overnight at low enzyme concentrations (≤100 nM) and low triphosphate concentrations (≤150 μM). Under these conditions, no RNA synthesis is observed in the absence of templates (Biebricher et al., 1981b, 1986), while contaminating RNA is readily amplified. All preparations used in this work were checked by this method for the absence of templates. Under these conditions, there is also no synthesis observed after incubation overnight by using tRNA, rRNA, or viral Qβ RNA as templates. On the other hand, RNA species that are replicated even with very low replication efficiencies can be reproducibly amplified without emergence of unrelated RNA sequences. The products of template-free synthesis have characteristic patterns, easily distinguishable from those of template-instructed amplification of contaminants (Biebricher et al., 1981a; Biebricher, 1987). Figure 1 shows an electropherogram of the template-free incorporation experiment used for the sequencing studies. Several bands, some of them corresponding to RNA with chain lengths of less than 50, are seen in one and the same sample, and the band patterns of different samples differ. While efficiently replicating RNA species had chain lengths of 80–220, early products of template-free synthesis had chain lengths below 60. The mobilities of the RNA products decrease during further amplification by serial transfers, because the inefficiently replicating early products undergo an evolution process accompanied by an increase in the chain lengths of the products (Biebricher et al., 1981a; Biebricher, 1983, 1987), probably by recombination processes (Biebricher & Luce, 1992).

Fingerprinting experiments have indicated that different bands in the same sample have related sequences, while

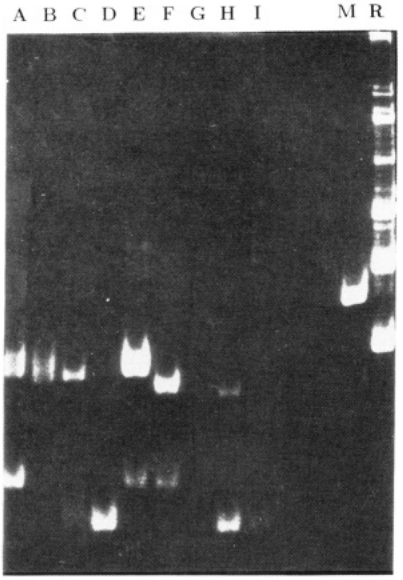

FIGURE 1: Electropherogram of products of template-free synthesis. Samples (100 μL) of template-free standard reaction mixtures containing 0.5 mM triphosphates were incubated with 75 nM Qβ replicase lacking R1 subunit for 21 h at 30 °C. Ten-microliter aliquots were diluted with an equal volume of stop mixture and loaded onto 12.6% polyacrylamide gels (Biebricher et al., 1982). M is MNV-11 double strand (87 nucleotides); R, a mixture of SV-11 and MDV-1 (the bands with the highest mobilities correspond to double strands of 60-, 115-, and 221-nucleotide lengths).

fingerprints of different samples were unrelated (Biebricher et al., 1981a). However, direct sequencing of the RNA was not feasible: most bands seen in electropherograms contained double-stranded RNA, because the short-chained complementary sequences rapidly combine to double strands (Biebricher et al., 1982, 1984). Therefore, separation of the different bands as well as separation of the complementary sequences would be required prior to sequencing. It was impossible to synthesize the required amounts of pure single-stranded RNA because the short-chained early products were replaced by more efficiently replicating species with longer chain lengths during their amplification. A method was thus required to amplify faithfully the sequences and to clone the different components of the sequence mixture (Biebricher, 1987). The method involved transcription into cDNA and amplification of the cDNA in vivo (Taniguchi et al., 1978), because DNA replication is far more accurate than RNA replication and lacks the strong selection bias for certain sequences (Biebricher, 1987).

Cloning Method. A novel cloning method (Figure 2a) applicable to all RNA species replicated by Qβ replicase was developed. It makes use of the fact that they have invariant 5′-terminal GG(G) and 3′-terminal (C)CCA sequences. For investigating the properties of replicating RNA species, transcription of the RNA from the cDNA clones must start with the correct 5′ terminus and end at the required 3′ terminus. A T7 transcription system was highly suitable, because the sequences required to start the transcription are fully compatible with the natural 5′ termini of replicating RNA (Studier & Dunn, 1983).

The attachment of the promoter sequence was achieved by using the double-stranded T7 promoter sequence with a single-stranded 3′-CCC overhang as primer for converting a cDNA sequence obtained by retrotranscription into a double-stranded DNA. An overhang as short as two bases was found to suffice to prime the second-strand cDNA synthesis with a suitable DNA polymerase. However, most DNA polymerases could

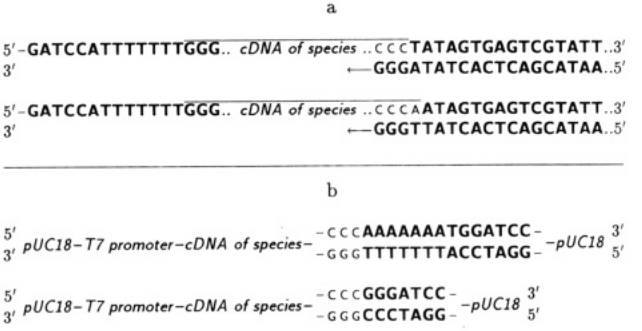

FIGURE 2: (a) Cloning strategy for replicating RNA species. The RNA was polyadenylated, heated to melt double strands, and retrotranscribed with an appropriate primer. The cDNA was isolated, the majority being adenylated at its 3′ terminus. Second-strand synthesis was performed by a DNA polymerase lacking 3′-exonuclease acitivity, using a double-stranded T7 promoter oligonucleotide bearing a GGG overhang (top). The part of the cDNA that has been terminally adenylated could be converted into double strand by an appropriately modified primer (second line). The primers used for first- and second-strand cDNA synthesis carried suitable restriction sites for cloning the cDNA into pUC18 (Yanish-Perron et al., 1985); the first-strand primer also supplied a restriction site allowing cutting at the 3′ terminus of the original sequence (e.g., by *Bst*XI, *Mbo*II, or *Gsu*I). Transcription from plasmids with T7 RNA polymerase in vitro produces RNA of the same polarity as the RNA cloned. With well-characterized RNA species, the properties of the transcription products do not deviate from those of the original RNA. The synthesized primer oligonucleotides are shown in bold type. (b) For transcription the plasmid was cut with *Bst*XI. The resulting transcripts often contained several A residues attached to their 3′ termini. A partial remedy was replacement of the *Bst*XI site by a *Sma*I site done as follows: the plasmid was cut with the restriction endonuclease *Bst*XI; the 3′-terminal A residues were removed by incubating the DNA with T4 DNA polymerase in the presence of dGTP and dCTP; and the 5′ T-overhang was clipped off with mung bean nuclease. After attachment of *Bam*HI linkers (pGGGATCCC), cutting of the plasmid with *Bam*HI restriction endonuclease, and ligation, a *Sma*I site was obtained with excellent efficiency which allowed cutting directly at the 3′ terminus of the cDNA insert.

not be used because the 3′ overhang was removed by their 3′-exonuclease activities, and no second-strand synthesis took place. A commercially available modified DNA polymerase from phage T7 lacking most of the 3′-exonuclease activity (Tabor & Richardson, 1987) worked with excellent yield on a model single-stranded DNA containing a CCC end. However, second-strand cDNA synthesis consistently gave much lower yields when single-stranded cDNA produced by retrotranscription of RNA was used.

The reason for the lower yield of a template produced by reverse transcriptase was found to be addition of an adenylate residue at the 3′ terminus of the cDNA by reverse transcriptase; terminal adenylation of the replica is a feature common to most DNA and RNA polymerases. We found that only 20% of the retrotranscribed cDNA contained the correct C termini, while 80% had an A end. Terminal adenylation prevents annealing of the second-strand primer pair, but in most cases the resulting reduced yield of cDNA clones was still acceptable. However, when the RNA input was exceptionally small, terminally adenylated strands were also cloned with the help of an appropriately modified primer pair (Figure 2a). When this modified primer was used in addition to the standard primer, most T7 promoter sequences were found to deviate from the consensus sequence but to nevertheless trigger transcription with satisfactory efficiency. Transcription from plasmids with this modified promoter sequence was inhibited by the addition of 100 mM salt.

In principle it is possible to use the primers of first- and second-strand cDNA synthesis for further amplifying the

pppGGGAGUUUCACAGGAUGAGACUACCCCUCACCCCA
pppGGGUUUUAAGUGAGAGAAACCCGAAGUUUGCCCCA
pppGGGGGGAAAAAAAGAGGCCCCGAAGGAACCCA
pppGGGAUCUAACAGGUCCUUUUAACCCA
pppGGGGGACUUUCAACGAGGGAGUCUCUCAACCCCA
pppGGGUUUAAAAUGUAAUAGGACCCACAUGAUCCCA †
pppGGGGUUUUCGAUUGGAGACCUUUUAUCAUCCCGUACACCGGCUUCCUCCA †
pppGGGUUAUCUAAGCCUGUUCGCUCCCCCA
pppGGGAAACUUUGGGUUUCUGUUUUACACCUACACGGUGUUUUACAC-
-CUACACGGUGUUCACCCCA †

pppGGGUUCCUUCGGGGCCUCUUUUCCCCCCACGGGGUUCCCA
pppGGGUCCUUCGGGGCCUCUUCCCCCCCA
pppGGUUCCUUCGGGGCCUUUUUCCCUCCACGGGGGGCUUUUCCCA *
pppGGUUCCUUCGGGGCCUCUUUUCCCCCCCGACGGGGGUUCCCCA *
pppGGUUCCUUCGGGGCCUCUUUUUCCCUCCCACGGGGGGCUUUUCCCA
pppGGGUUCCUUCGGGGCCUCUUUUUCCCCA
pppGGGAGACCCCCCGUGGAGGGGAAAAAGAGGACCCGAAGGAACCCCA

FIGURE 3: Sequences of some products of template-free synthesis. Sequencing from plasmids constructed as described in Figure 2 was done by the dideoxy chain termination method (Sanger 1977). Top: Sequences found in independent template-free incorporations. While the sequences look similar at first glance, they could not be aligned to sequence homologies. For sequences marked with † the full complementary sequence has also been found. Note the exact duplication of 22 nucleotides in the fifth sequence, which is missing in another clone from the same incorporation mixture. Bottom: Sequences of different clones from one template-free incorporation mixture. Note the similarity of the sequences. The last sequence shown is complementary to the others. The overlined part is a frequent motif important for RNA structures (see Discussion).

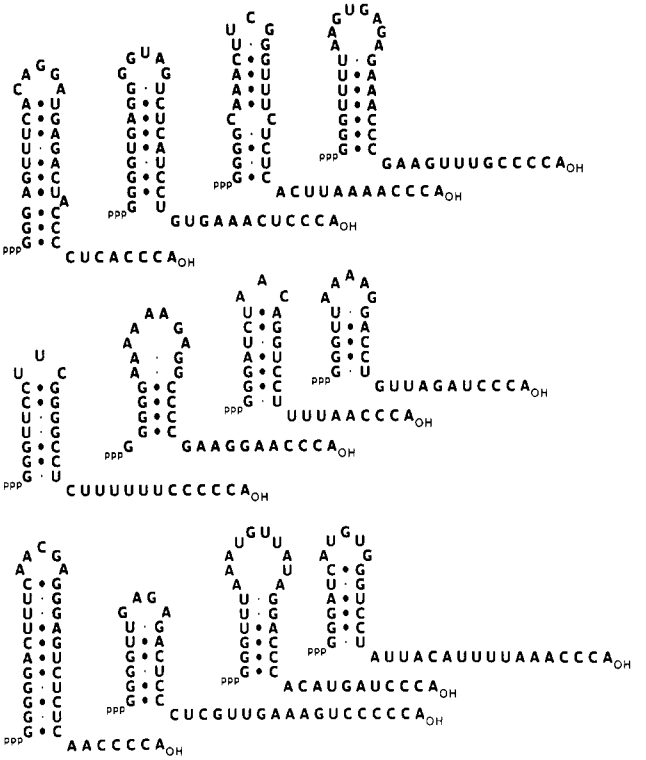

FIGURE 4: Tentative secondary structures of both complementary strands of short-chained RNA species replicated by Qβ replicase. The secondary structures of the first six sequences shown in Figure 3 were calculated by using the algorithm of Zuker and Stiegler (1981). While the 5′ termini are involved in a double-stranded structure, the 3′ ends are unstructured. This feature is also seen in almost all efficiently replicating RNA species.

sequences by the polymerase chain reaction (Mullis et al., 1986); this was not done in this study, since a sufficient number of clones were obtained without amplification.

Sequences of the Products of Template-Free Synthesis. The sequences of some products of template-free incorporation with Qβ replicase are listed in Figure 3. Note the short lengths of the sequences found, some of which were as short as 25 nucleotides. Such short sequences, however, could be artifacts of abortive first-strand cDNA synthesis. Since the cloning method conserves the polarity of the sequences, the fact that both plus and minus strands of the sequences were often found is a clear indication that the RNA strands (a) must have undergone a replication process and (b) are full strands and were not produced by degradation or abortive replication products.

Except for the ends and a relative abundance of nucleotide repetitions, little sequence homology was found among products generated by template-free incorporation in different test tubes, even when the products were prepared under identical conditions. On the other hand, the different RNA sequences derived from one and the same incorporation mixture are strongly related to one another. The mutational events seen include base substitutions, base insertions (frequently at positions of nucleotide repetition), and sequence part repetitions, probably by a copy choice mechanism (Biebricher & Luce, 1992). These results are in full agreement with previous fingerprinting experiments (Biebricher et al., 1981a). Obviously, the replicase must recognize some common features in the sequences, because only a small fraction of RNA sequences are accepted for replication. Since single-stranded RNA molecules, and particularly the replicating ones (Biebricher et al., 1982), have strong secondary and tertiary structures, the calculated secondary structures of the short-chained RNA species found were compared.

Structure of Short-Chained Replicating RNA Species. Tentative secondary structures of the sequenced RNA species were calculated by the algorithm of Zuker and Stiegler (1981). Indeed, a striking similarity among their structures was found (Figure 4). The similarity is emphasized by the fact that the structures of the complementary strands of one and the same species are also similar: the 5′ ends were always found to be in a double-stranded stem, while the 3′ termini were unstructured. This structure has been found also for most RNA species that are efficiently replicated (Schaffner et al., 1977, Biebricher, 1987; Biebricher & Luce, 1992). If only normal Watson–Crick base-pairing would take place, complementary strands would be expected to have mirror structures, i.e., a 5′-stem region would lead to a 3′-stem region in the complementary strand. The homologous structures of the termini of the complementary strands were favored by the inclusion of G:U pairs at strategic positions in the stems, thereby destabilizing antiparallel structures by nonpairing A:C and favoring a more stable stem at the 5′ ends.

Replication of the Short-Chained RNA Species. RNA can be readily prepared by run-off transcription with T7 RNA polymerase from the DNA plasmids cut with restriction endonuclease *Bst*XI. When the cloning and transcription method was tested with the optimized RNA species MNV11 (Biebricher, 1987), it was found that the transcripts had heterogeneous 3′ termini with 0–6 terminal A residues. Addition of salt (175 mM NaCl) reduced the heterogeneity, but in most cases this was not necessary, because the transcripts could be shown to have the same properties in directing replication by Qβ replicase. Improvement in the 3′ heterogeneity was also obtained by replacing the *Bst*XI site with a *Sma*I site by the method shown in Figure 2b. Transcription yields, as well as the heterogeneities of the transcripts and the tendency of transcripts to aggregate, varied widely among different RNA species cloned, in particular among the shorter

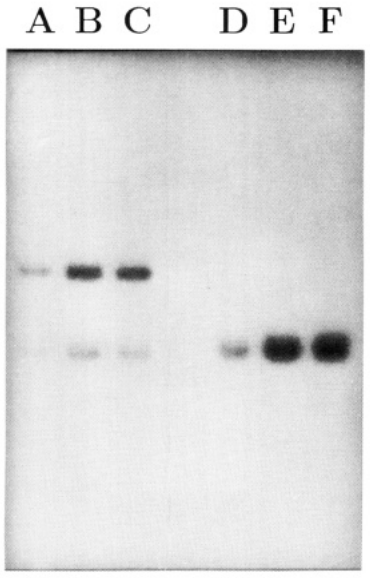

FIGURE 5: Amplification of transcripts of clones (marked in Figure 3 with *) by Qβ replicase. Conditions were as in Figure 1, except that ^{32}P-labeled ATP (200 Ci/mol) and 35 nM transcript were used as template. Shown is an autoradiogram of an electrophoresis gel after amplification of two transcripts with Qβ replicase for 5 min (A,D), 7 min (B, E), and 9 min (C, F) without heating (A,B,C) or after heating for 1 min at 100 °C (D,E,F). The mobilities of the single- and double-strand bands agree with the expected chain length of 45.

transcripts. Some of the sequences found had chain lengths of as little as 25. They show many features shown also by the longer chains and may thus indeed be replication products, but a direct proof would require that RNA strands produced by transcription are replicated by Qβ replicase. The plasmid preparations used for synthesizing the RNA by transcription were not absolutely free from contaminating RNA; however, the small amount of contaminants could be diluted out of the amplification mixture by starting synthesis with about 10^6 strands of transcribed RNA. When the shortest RNA strands (<30 nucleotides) were used as template, no RNA synthesis was detectable. Since these extremely short sequences were partial sequences of longer molecules also found in the same mixture and no sequences exactly complementary to them were found, they are probably artifacts of abortive first-strand synthesis having by chance the CCC ends required for second-strand synthesis in the cloning method. When RNA species with chain lengths ~35 were prepared by transcription of the appropriate plasmids, they were amplified faithfully by Qβ replicase in vitro (Figure 5), although rather slowly. Extensive kinetic data from many sepecies are not yet available, but for the species shown in Figure 5 an overall replication rate in the linear growth phase of about 1.8×10^{-3} s^{-1} was estimated, about one-half that of the optimized RNA species MNV-11. The exponential autocatalytic amplification is even more retarded. It is unlikely that the reduced replication efficiency was due to its synthesis by transcription: MNV-11 produced by transcription had kinetic properties identical to MNV-11 synthesized by Qβ replicase. In the absence of salt, the short RNA species shown in Figure 5 could be amplified 10^{20}-fold by serial transfers without selection of RNA of higher chain length; after the RNA obtained after many rounds of transcript replication was recloned and sequenced, the determined sequences were identical to the template sequences or closely related mutants belonging to the quasispecies distribution. The plus and minus strands of short RNA strands rapidly combine, forming double-stranded RNA (Figure 5). A detailed study of the replication kinetics of early products of template-free RNA synthesis is now readily possible, because these products can be obtained in virtually unlimited quantities by transcription from DNA clones.

DISCUSSION

In addition to replicating its natural templates Qβ RNA itself and its complement, Qβ replicase replicates a number of short-chained RNA species formed in vitro by template-free RNA synthesis. RNA material with similar properties is also found late in the infection process in vivo (Banerjee et al., 1967). Except for one species which probably has been formed by recombination from two pieces of RNA present in infected *E. coli* cells (Munishkin et al., 1988), the genetic origin of the RNA sequences is unknown. Sumper and Luce (1975) presented evidence that highly purified Qβ replicase can produce after long lag times RNA species de novo. In later studies, it was shown (Biebricher et al., 1981a,b) that (1) de novo synthesis can be observed only under special conditions; (2) the lag times are usually orders of magnitude longer than the time required for the amplification of a single strand of the species formed; and (3) even under completely identical conditions, different products are formed. Partial instruction by oligonucleotides or by RNA not replicated by Qβ replicase could not be ruled out. The interpretation of template-free RNA synthesis as de novo synthesis was contested (Hill & Blumenthal, 1983; Chetverin et al., 1991); the controversial interpretation is discussed elsewhere (Biebricher et al., 1993).

In this paper we show that the first products of template-free synthesis are RNA species of sequences with nucleotide chain lengths of 30–50 that are replicated by Qβ replicase quite inefficiently. The primary sequences of the RNA species isolated from different template-free incorporation mixtures, even when incorporation is performed under identical conditions, are unrelated; no significant sequence homologies were found when searching the *E. coli* sequences of the Genbank, the Qβ sequence, or the numerous other RNA species sequences replicated by Qβ replicase with computer programs. In one and the same template-free incorporation mixture, different sequences were also found, but they showed strong sequence homologies and are apparently derived from a common ancestor. The enormous sequence diversity of the early products of template-free synthesis can be explained by the selection pressure to improve the replication efficiency. During extensive replication periods by several serial transfers, the replication efficiencies of the early RNA species are optimized by mutation and recombination (Biebricher, 1983; Biebricher & Luce, 1992) resulting eventually in selection of evolutionarily stable RNA species (Biebricher et al., 1981a) of chain lengths of 70–220. Recombination (Munishkin et al., 1988; Biebricher & Luce, 1992) plays an important role in optimizing the RNA.

The synthesis of inefficient templates of short lengths with differing sequences is what one would expect from a de novo mechanism. Instruction by a heterogeneous population of nonreplicatable RNAs which may be present in undetectable amounts in the enzyme preparation may also contribute, but no evidence could be found in our experiments for its presence or its contribution. It has been shown that an RNA species that is amplified by *E. coli* DNA-dependent RNA polymerase can be selected from a random copolymer (Biebricher & Orgel, 1973). Qβ replicase apparently is able to build its own random copolymer: a slow accumulation of oligonucleotides with the composition pppGpR(pN)$_n$ unable to replicate has been observed (Biebricher et al., 1986). The newly developed cloning procedure presumably gives a roughly representative

sample of the population of the RNA mixture in the test tube, but neglects RNA which does not have the correct termini.

The high diversity and short lengths of the early products of template-free synthesis may give clues to criteria an RNA has to meet in order to be amplified by Qβ replicase. While the sequences of the short RNA species obtained do not unequivocally reveal a building principle, the secondary structures of all short-chained species have been found to be strikingly similar. The 5′ ends form a double-stranded stem, while the 3′ termini are unstructured. The last condition has been proposed by Schaffner et al. (1977). Folding the elongating replica into a double-stranded structure early in the replication process is probably required for efficient strand separation, a process not yet understood in detail. The conditions listed do not suffice for replication; another condition, the presence of both a 3′-terminal and an interior C-cluster (or pyrimidine cluster; Küppers & Sumper, 1975), is also fulfilled. Other structural similarities among RNA species replicated by Qβ replicase found (Voronin, 1992) are not present but may be required for higher replication efficiencies.

It is likely that the tertiary structure of the RNA must complement the replicase structure to produce the replication activity. A further requirement is escaping destruction of the RNA structure by double strand formation (Biebricher et al., 1982, 1984). This is probably the reason for the occurrence of a motif present in many but not all species: a (stable) stem with (C)UUCG(G) (marked in Figure 3) in the loop near one of the ends. This motif is also present in many other ribosomal and viral RNA sequences of different eukaryotic organisms, probably because it contributes to a stable secondary structure (Tuerk et al., 1988). Therefore, the presence of this motif cannot be used as an argument for instructed synthesis: convergent evolution by the selection pressure to protect the short-chained replicating RNA against double strand formation seems more plausible.

ACKNOWLEDGMENT

We thank Dr. J. S. McCaskill for critical reading of the manuscript and Professor M. Eigen for his interest and support.

REFERENCES

Banerjee, A. K., Eoyang, L., Hori, K., & August, J. T. (1967) Replication of RNA viruses. Replication of a Natural 6 S RNA by the Qβ RNA polymerase, *J. Mol. Biol. 45*, 181–193.

Bauer, G. J., McCaskill, J. S., & Otten, H. (1989) Traveling Waves of in Vitro Evolving RNA, *Proc. Natl. Acad. Sci. U.S.A. 86*, 7937–7941.

Biebricher, C. K. (1983) Darwinian Selection of Self-Replicating RNA, in *Evolutionary Biology* (Hecht, M. K., Wallace, B., & Prance, G. T., Eds.) Vol. 16, pp 1–52, Plenum Press, New York.

Biebricher, C. K. (1987) Replication and Evolution of Short-Chained RNA Species Replicated by Qβ Replicase, *Cold Spring Harbor Symp. Quant. Biol. 52*, 299–306.

Biebricher, C. K., & Orgel, L. E. (1973) An RNA That Multiplies Indefinitely with DNA-Dependent RNA Polymerase: Selection from a Random Copolymer, *Proc. Natl. Acad. Sci. U.S.A. 70*, 934–938.

Biebricher, C. K., & Luce, R. (1992) In Vitro Recombination and Terminal Elongation of RNA by Qβ Replicase, *EMBO J. 11*, 5129–5135.

Biebricher, C. K., Eigen, M., & Luce, R. (1981a) Product Analysis of RNA Generated de novo by Qβ Replicase, *J. Mol. Biol. 148*, 369–390.

Biebricher, C. K., Eigen, M., & Luce, R. (1981b) Kinetic Analysis of Template-Instructed and de novo RNA Synthesis by Qβ Replicase, *J. Mol. Biol. 148*, 391–410.

Biebricher, C. K., Diekmann, S., & Luce, R. (1982) Structural Analysis of Self-Replicating RNA Synthesized by Qβ Replicase, *J. Mol. Biol. 154*, 629–648.

Biebricher, C. K., Eigen, M., & Gardiner, W. C. (1983) Kinetics of RNA Replication, *Biochemistry 22*, 2544–2559.

Biebricher, C. K., Eigen, M., & Gardiner, W. C. (1984) Kinetics of RNA Replication: Plus-Minus Assymetry and Double-Strand Formation, *Biochemistry 23*, 3186–3194.

Biebricher, C. K., Eigen, M., & Gardiner, W. C. (1985) Kinetics of RNA Replication: Competition and Selection among Self-Replicating RNA Species, *Biochemistry 24*, 6550–6560.

Biebricher, C. K., Eigen, M., & Luce, R. (1986) Template-Free RNA Synthesis by Qβ Replicase, *Nature (London) 321*, 89–91.

Biebricher, C. K., Eigen, M., & McCaskill, J. S. (1993) Template-Directed and Template-Free RNA Synthesis by Qβ replicase, *J. Mol. Biol.* (in press).

Birnboim, H. C. (1983) A Rapid Alkaline Extraction Method for the Isolation of Plasmid DNA, *Methods Enzymol. 100*, 243–255.

Chetverin, A. B., Chetverina, H. V., & Munishkin, A. V. (1991) On the Nature of Spontaneous RNA Synthesis by Qβ Replicase, *J. Mol. Biol. 222*, 3–9.

Davanloo, P., Rosenberg, A. H., Dunn, J. J., & Studier, F. W. (1984) Cloning and Expression of the Gene for Bacteriophage T7 RNA Polymerase, *Proc. Natl. Acad. Sci. U.S.A. 81*, 2035–2039.

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.

Hanahan, D. (1983) Studies on Transformation of *Escherichia coli* with Plasmids, *J. Mol. Biol. 166*, 557–580.

Hill, D., & Blumenthal, T. (1983) Does Qβ Replicase Synthesize RNA in the Absence of Template?, *Nature (London) 301*, 350–352.

Jahn, M. J., Jahn, D., Kumar, A. M., & Söll, D. (1991) Mono Q Chromatography Permits Recycling of DNA Template and Purification of RNA Transcripts after T7 RNA Polymerase Reaction, *Nucleic Acids Res. 19*, 2786.

Küppers, B.-O., & Sumper, M. (1975) Minimal Requirements for Template Recognition by Qβ Replicase, *Proc. Natl. Acad. Sci. U.S.A. 72*, 2640–2644.

Maniatis, T., Fritsch, E., & Sambrook, J. (1982) *Molecular Cloning: a Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Milligan, J. F., & Uhlenbeck, O. C. (1989) Synthesis of Small RNA Species by T7 RNA Polymerase, *Methods Enzymol. 180*, 51–62.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., & Erlich, H. (1986) Specific Enzymatic Amplification of DNA in Vitro: The Polymerase Chain Reaction; *Cold Spring Harbor Symp. Quant. Biol. 51*, 263–273.

Munishkin, A. V., Voronin, L. A., & Chetverin, A. B. (1988) An in Vivo Recombinant RNA Capable of Autocatalytic Synthesis by Qβ Replicase, *Nature (London) 333*, 473–475.

Munishkin, A. V., Voronin, L. A., Ugarov, V. I., Bondareva, L. A., Chetverina, H. V., & Chetverin, A. B. (1991) Efficient Templates for Qβ Replicase Are Formed by Recombination from Heterologous Sequences, *J. Mol. Biol. 221*, 463–472.

Sanger, F., Nicklen, S., & Coulsen, A. R. (1977) DNA Sequencing with Chain-Determining Inhibitors, *Proc. Nat. Acad. Sci. U.S.A. 74*, 5463–5467.

Schaffner, W., Rüegg, K. J., & Weissmann, C. (1977) Nanovariant RNAs: Nucleotide Sequence and Interaction with Bacteriophage Qβ Replicase, *J. Mol. Biol. 117*, 877–907.

Sippel, A. M. (1973) Purification and Characterization of Adenosine Triphosphate:Ribonucleic Acid Adenyltransferase from *Escherichia coli*, *Eur. J. Biochem. 37*, 31–40.

Studier, F. W., & Dunn, J. J. (1983) Organization and Expression of Bacteriophage T7 DNA, *Cold Spring Harbor Symp. Quant. Biol. 47*, 999–1007.

Sumper, M., & Luce, R. (1975) Evidence for de novo Production of Self-Replicating and Environmentally Adapted RNA Structures by Bacteriophage Qβ Replicase, *Proc. Natl. Acad. Sci. U.S.A. 72*, 162–166.

Tabor, S., & Richardson, C. C. (1987) DNA Sequence Analysis with a Modified Bacteriophage T7 DNA Polymerase, *Proc. Natl. Acad. Sci. U.S.A. 84*, 4767–4771.

Taniguchi, T., Palmieri, M., & Weissmann, C. (1978) Qβ DNA-Containing Hybrid Plasmids Giving Rise to Qβ Phage Formation in the Bacterial Host, *Nature (London) 274*, 223–228.

Tuerk, C., Gauss, P., Thermes, C., Groebe, D. R., Gayle, M., Guild, N., Stormo, G., d'Aubenton-Carafa, Y., Uhlenbeck, O. C., Tinoco, I., Brody, E. N., & Gold, L. (1988) CUUCGG Hairpins: Extraordinary Stable RNA Secondary Structures Associated with Various Biochemical Processes, *Proc. Natl. Acad. Sci. U.S.A. 85*, 1364–1368.

Voronin, L. A. (1992) Structural Consensus of RNA Templates Replicated by Qβ Replicase, *Biochimie 74*, 491–494.

Yanisch-Perron, C., Vieira, J., & Messing, J. (1985) Improved M13 Phage Cloning Vectors and Host Strains: Nucleotide Sequences of the M13mp18 and pUC19 Vectors, *Gene 33*, 103–109.

Zuker, M., & Stiegler, P. (1981) Optimal Computer Folding of Large RNA Sequences Using Thermodynamics and Auxiliary Information, *Nucleic Acids Res. 9*, 133–148.