FEBS

RNA Viruses:
Replication and Structure

Ribosomes:
Structure, Function and Biogenesis

Volume 27

Organized by:
H. BLOEMENDAL
E. M. J. JASPARS
A. VAN KAMMEN
R. J. PLANTA

FEBPBY

FEDERATION OF EUROPEAN BIOCHEMICAL SOCIETIES
EIGHTH MEETING, AMSTERDAM, 1972

RNA VIRUSES:
replication and structure

RIBOSOMES:
structure, function and biogenesis

Volume 27

Organized by

H. BLOEMENDAL, *Nijmegen*
E. M. J. JASPARS, *Leiden*
A. VAN KAMMEN, *Wageningen*
R. J. PLANTA, *Amsterdam*

1972

NORTH-HOLLAND / AMERICAN ELSEVIER

Library of Congress Catalog Card Number: 72 94598
ISBN North-Holland
Series: 0 7204 4300 8
Volume: 0 7204 4327 X
ISBN American Elsevier: 0 444 10422 4

Publishers:

North-Holland Publishing Company - Amsterdam - London

Sole distributors for the USA and Canada:

American Elsevier Publishing Company, Inc.
52 Vanderbilt Avenue
New York, N.Y. 10017

Printed in The Netherlands

RNA VIRUSES:
replication and structure

Part I

Organized by

E. M. J. JASPARS, *Leiden*
A. VAN KAMMEN, *Wageningen*

INTRODUCTION

In this symposium papers have been brought together on topics from bacterial-, animal- and plant virology as it was thought that this might contribute to the exchange of ideas among investigators of very different RNA viruses.

The simplest viruses known are the RNA phages. These phages have been, and are even now, intensively used as tools for the study of problems of messenger RNA translation and the replication of RNA. The attention is focused nowadays on the determination of the nucleotide-sequence of the phage RNA. Detailed knowledge about the structure of this RNA will help to understand the specific interaction of viral proteins with the viral RNA and how the translation of the RNA into viral proteins, the replication of the viral RNA and finally the assembly of the virus particles are regulated.

The structure of simple icosahedral RNA viruses is dealt with in another paper. Plant viruses have been used most for such studies because it is rather easy to obtain several plant viruses in large amounts. It has been found that the nature of the interactions stabilizing the structure of the virions may vary widely. In some cases the interactions between the protein subunits of the capsids predominate, whereas in others there appear to be strong interactions between the RNA and the protein subunits. This appears important to know for understanding of how the nucleic acid can be liberated from the virus particles upon infection and for understanding the process of virus assembly.

The replication of viral RNA in eucaryotic cells is intimately connected with cellular structures. The methods available for the location of viral RNA synthesis in the infected cell are discussed from a study on the synthesis of the RNA of turnip yellow mosaic virus. In that case the outer chloroplast membrane appears to be involved in the synthesis of virus RNA.

In the last few years it has been found that nucleic acid synthetizing enzymes are an integrated part of the virions of several kinds of RNA viruses. Especially the research on tumor inducing RNA viruses was greatly stimulated when it was found that such viruses contain a reverse transcriptase, which transcribes RNA in DNA. This enzyme has been used succesfully now as a tool for tracing viruses in human tumors, as is shown clearly in a paper on this subject.

Another peculiarity of the RNA viruses, which has been recognized in the past five years is the occurrence of viruses with a multipartite genome. A considerable number of viruses have genomes consisting of two or more RNA molecules. These molecules can occur either in one single particle (e.g. influenza virus) or in separate particles (e.g. cowpea mosaic virus and alfalfa mosaic virus). The nature of these viruses and the distribution of the virus properties among the different parts of the genome are discussed in the last two papers of this symposium.

LIST OF CONTENTS

RNA Viruses

Ribosomes

STRUCTURE AND FUNCTION OF PHAGE RNA, 1972.

C. Weissmann[o], M.A. Billeter[o], H.M. Goodman[*], J. Hindley[+] and H. Weber[o]

[o]Institut für Molekularbiologie, Universität Zürich
[*]University of California, Department of Biochemistry, San Francisco
[+]University of Bristol, Department of Biochemistry, Bristol

The RNA-containing coliphages may be classified into three, possibly four serological groups (Watanabe et al., 1967; Sakurai et al., 1968). The first of these comprises phages such as f2, R17 and MS2, which are quite similar among themselves and differ considerably from Qβ, a member of the third group. These four phages have been studied intensively in the last few years (cf. reviews by Zinder, 1965; Weissmann and Ochoa, 1967; Lodish, 1968a; Stavis and August, 1970). Their RNAs consist of about 3500-4500 nucleotides (Sinha et al., 1965; Boedtker, 1971) and comprise three cistrons (Gussin, 1966; Horiuchi et al., 1966; Horiuchi and Matsuhashi, 1970). The capsid consists of about 180 coat protein subunits (Vasquez et al., 1966) and one molecule of a further virus-specific protein, the A or maturation protein (Nathans et al., 1966; Steitz, 1968; Garwes et al., 1969). Qβ contains a further protein, designated A_1, which arises by read-through from the coat cistron into the subsequent nucleotide sequence (Horiuchi et al., 1971; Weiner and Weber, 1971; Moore et al., 1971) (cf. Fig. 1). The third viral cistron codes for the β subunit of the viral replicase (Kamen, 1970; Kondo et al., 1970). The cistron order is A-coat-replicase for both phage groups (Jeppesen et al., 1970a and 1970b; Konings et al., 1970; Hindley et al., 1970; Staples et al., 1971). The complete aminoacid sequence of the coat proteins of several phages has been established (Wittmann-Liebold, 1966; Weber and Konigsberg, 1967; Lin et al., 1967; Konigsberg et al., 1970); only the first few aminoacids of the A proteins and the replicase subunits have been determined (Lodish, 1968b; Lodish, 1969; Weiner and Weber, 1971).

The nucleotide sequences of the phage RNAs are being studied in several laboratories. The strategy currently applied may be broken down into three phases: (1) Reduction of the viral RNA into subsets of about 50-200 nucleotides in length; (2) Determination of the nucleotide sequence of the subsets; (3) Alignment of the subsets in their natural order. Only the first of these steps shall be discussed below (cf. Gilham (1970) for steps (2) and (3).

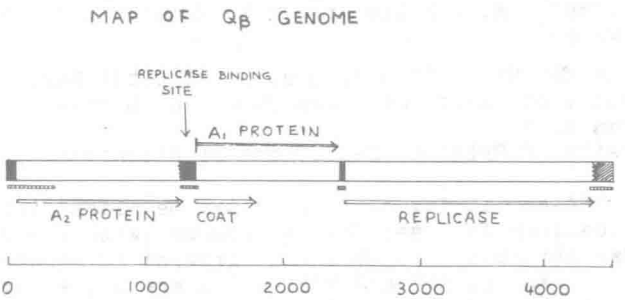

Fig. 1. A map of the Qβ genome.
The map is based on data from Hindley et al. (1970), Staples et al. (1971) and Moore et al. (1971). The extracistronic regions are shaded. The areas of known nucleotide sequence are indicated by the narrow bars under the map.

Reduction of the viral RNA to subsets.

Two principal approaches have been used to obtain RNA segments susceptible to nucleotide sequence analysis, the degradative and the synthetic. The degradative methods are based on the use of nucleases under conditions where only a limited number of the potentially cleavable phosphodiester bonds are split. Relative resistance to nucleolytic cleavage is inherent to the secondary and possibly tertiary structure of the RNA, since linkages between nucleotides located in double-helical regions are less susceptible to enzymatic attack than those present in single-stranded regions (Penswick and Holley, 1965; Gould, 1967; Min Jou et al., 1968; De Wachter et al., 1971); it appears however, that also the primary structure of the RNA confers relative resistance on some internucleotide bonds which in principle are susceptible to the nuclease in question (Pinder and Gratzer, 1970). The isolation of pure RNA fractions from the mixture of degradation products is one of the principal technical bottlenecks. Two main separation techniques have been developed for this purpose, namely polyacrylamide gel electrophoresis (Gould, 1967; Min Jou et al., 1968; Adams et al., 1969; De Wachter and Fiers, 1971) and displacement chromatography (homochromatography) (Brownlee and Sanger, 1969; Jeppesen et al., 1970a).

Partial degradation as described above yields a more or less random collection of fragments and does not permit a priori selection of oligonucleotides from any particular region of the RNA. Isolation of specific segments of viral RNA can be attained by binding ribosomes to viral RNA under

conditions of polypeptide chain initiation, treating the complex with RNase and isolating the RNA piece which remains attached to the ribosome. This procedure has allowed the isolation of RNA segments comprising the ribosome binding sites of the f2 coat cistron (Gupta et al., 1970) of all cistrons of R17 (Steitz, 1969) and of Qβ (Hindley and Staples, 1969; Staples et al., 1971; Staples and Hindley, 1971; Steitz, 1972) (Fig. 2). A similar approach, in which Qβ replicase was

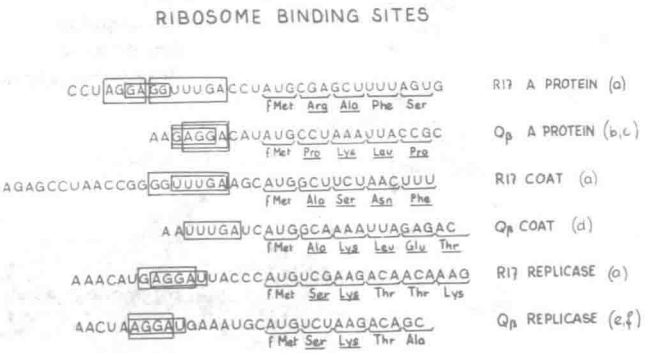

Fig. 2. The ribosome binding sites of RNA phage cistrons. The sequences of the underlined aminoacids have been determined directly (cf. text for references); the others were deduced from the nucleotide sequence. The frames call attention to sequences common to one or more sites (a) Steitz, 1969; (b) Billeter et al., 1969b; (c) Staples et al., 1971; (d) Hindley and Staples, 1969; (e) Staples and Hindley, 1971; (f) Steitz, 1972.

bound to Qβ RNA prior to digestion with T_1 RNase has yielded a piece of RNA containing a binding site for the viral polymerase (Weber et al., 1972); the binding of R17 coat protein to R17 RNA allowed the isolation of a binding site for this protein (Bernardi and Spahr, 1972) (cf. below).

A synthetic approach to the preparation of viral RNA subsets has been reported by Billeter et al. (1969a). Qβ RNA is synthesized in vitro by Qβ replicase, using either a plus or a minus strand as template. Synthesis begins with the 5' terminus and proceeds with an elongation rate of about 6 nucleo-

tides/sec at 20°. By synchronizing RNA synthesis and using α-^{32}P-labelled ribonucleoside triphosphates as substrates during defined periods of synthesis it becomes possible to synthesize either plus or minus strands radioactively labelled, in principle, in any desired segment (Fig. 3). In practice

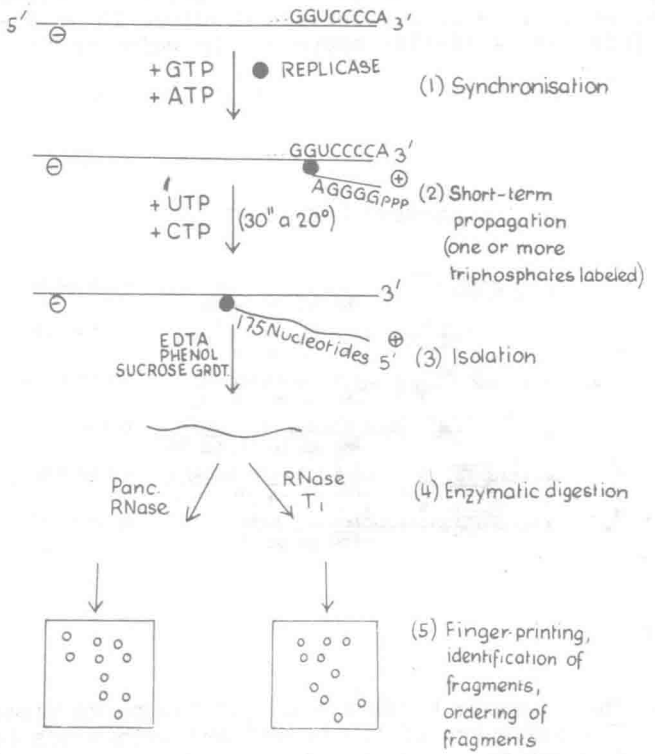

Fig. 3. Synchronized synthesis of 5' terminal segments of Qβ RNA in vitro.
1) Synchronization of Qβ replicase. Qβ minus strands, GTP and ATP are incubated to give initiation and limited chain elongation up to the point where CTP is required for further synthesis. 2) UTP and CTP are added. (Any one or all of the substrates are α-^{32}P-labelled). Incubation is carried out under conditions where the elongation rate is about 6 nucleotides/sec. Synthesis is stopped at any desired point by the addition of EDTA. 3) After deproteinization, the radioactive product is heat-denatured and purified by sucrose gradient centrifugation. 4) Aliquots are digested with pancreatic and T_1 RNase respectively. 5) The oligonucleotides are separated by 2-dimensional electrophoresis and analyzed by Sanger's techniques.

however, synchrony diminishes with elongation so that the method has up to now been utilized only to examine the first few hundred nucleotides at the 5' termini of both plus and minus strands. A new approach has been reported recently allowing resynchronisation of synthesis at an interior position of the RNA (Kolakofsky et al., 1971). As shown in Fig. 4, a ribosome is bound to the coat cistron binding site (the preferred ribosome binding site in native Qβ RNA (Hindley and Staples, 1969; Staples et al., 1971)) and the RNA-ribosome complex is used as a template for Qβ replicase with unlabelled ribonucleoside triphosphates as substrates. Elongation ceases when replicase reaches the ribosome. The ribosome is then removed from the RNA by treatment with EDTA and the replicating complex is separated from substrates and EDTA by chromatography on Sephadex. On addition of radioactive substrates, synchronized synthesis ensues and a labelled minus strand segment is produced, which extends from the region complementary to the coat initiation site into the region complementary to the end of the A protein cistron. To a lesser

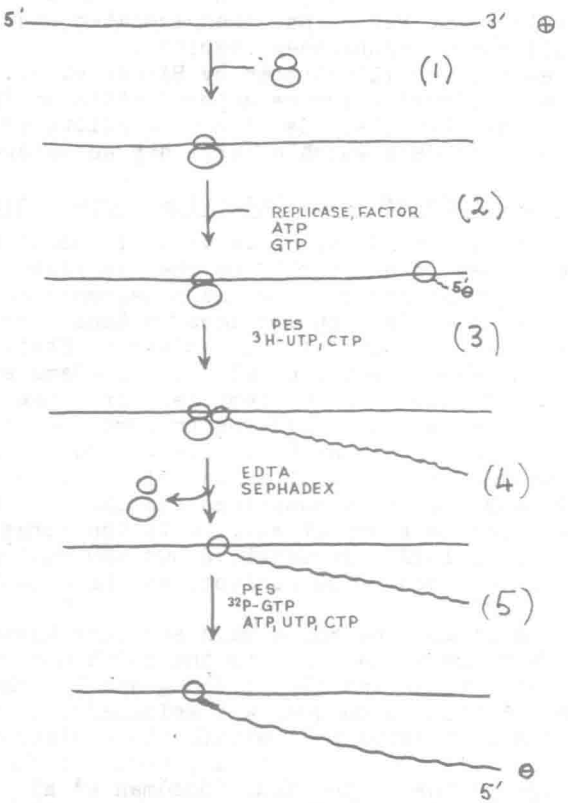

Fig. 4. Resynchronization of RNA synthesis at the beginning of the coat cistron.
(1) Ribosomes are bound to Qβ RNA under conditions of

polypeptide chain initiation. (2) The Qβ RNA-ribosome complex is incubated with Qβ replicase, host factor, ATP and GTP to give an early replicating complex. (3) Polyethylene sulfonate is added to inactivate any free enzyme and thereby prevent initiation during the later phase of the experiment. Elongation is started by adding UTP and CTP (in some experiments ^{3}H-UTP was used to monitor the first phase of synthesis. (4) After 5 min at 37° EDTA is added to remove the ribosome from the complex. EDTA and substrates are removed by Sephadex chromatography. (5) The complex is then incubated with the 4 standard triphosphates (at 0.01 mM), of which one or all are α-labelled with ^{32}P at high specific activity, for 50-100 sec at 20°, to allow synthesis of a labelled segment of about 50-100 nucleotides in length.

extent, ribosome binding occurs also at the replicase cistron, and a shorter minus strand segment, with a labelled region complementary to the beginning of this cistron and the preceding intercistronic region has been isolated and characterized (M.A. Billeter, unpublished results).

As pointed out and illustrated by Bishop et al. (1968) and Billeter et al. (1969b) sequence determination on RNA labelled with a single, α-^{32}P-nucleotide at a time allows exploitation of nearest neighbor data which offers many advantages.

The primary structure of phage RNAs: the current status.

In the case of the RNA of phage group I (considering combined data obtained both from R17 by the Cambridge group and from MS2 by the Ghent group) three main segments have been elucidated. The first (145 nucleotides) extends from the 5' terminus into the beginning of the A cistron (Steitz, 1969; Cory et al., 1970; De Wachter et al., 1971; Adams et al., 1972), the second (about 420 nucleotides) from the intercistronic region after the A cistron through the coat cistron and into the beginning of the replicase cistron (cf. Fig. 5) (Steitz, 1969; Jeppesen et al., 1970a; Nichols, 1970; Min Jou et al., 1972) and the third comprises the last 104 nucleotides of the molecule (Cory et al., 1970; Contreras et al., 1971; Cory et al., 1972). In total, about 680 nucleotides, or 20% of the genome (about 3500 nucleotides) have thus been elucidated.

In the case of Qβ, the three main segments known at present extend from the 5' terminus to the 330th nucleotide (Billeter et al., 1969a and 1969b; Young and Fraenkel-Conrat, 1971; Hindley, Billeter, Goodman and Weissmann, unpublished results), which is located well within the A cistron (see Fig. 6 for the first 200 nucleotides), and from the 160th to last nucleotide to the 3' terminus (Goodman et al., 1970; Goodman, Billeter, Hindley and Weissmann, unpublished results).

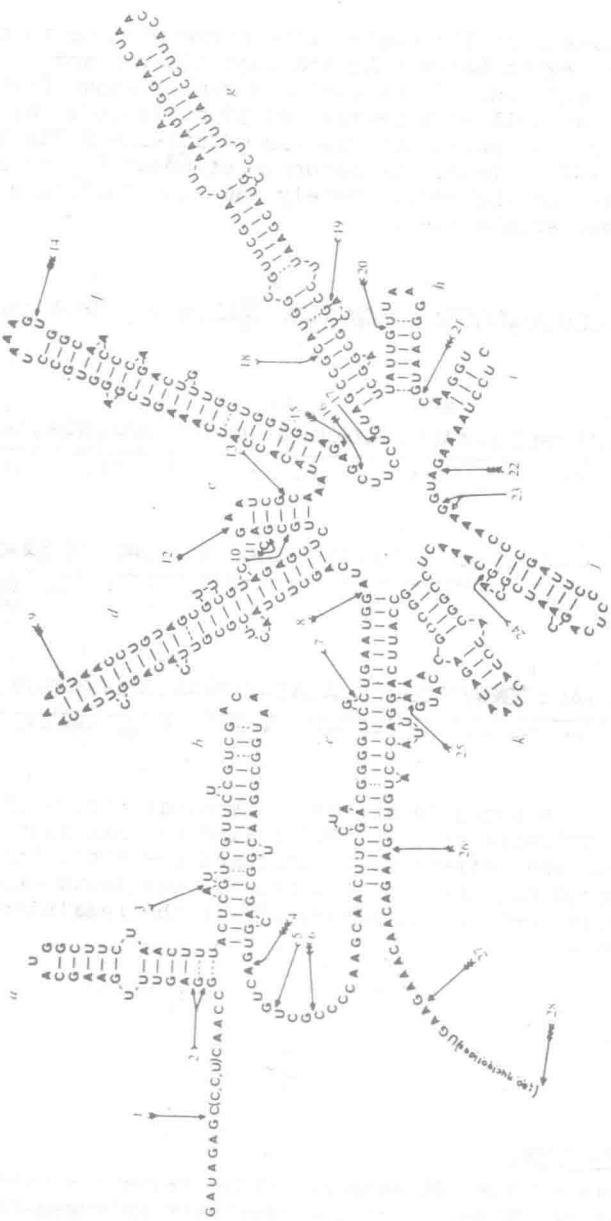

Fig. 5. Nucleotide sequence of the coat cistron of MS2 RNA. (Min Jou et al., 1972).

Moreover, a segment of 117 nucleotides corresponding to the intercistronic region between A2 and coat cistron and extending into the beginning of the coat cistron is known (Weber et al., 1972), as well as a segment of 38 nucleotides around the beginning of the replicase cistron (Staples and Hindley, 1971; Steitz, 1972). Thus, the sequence of about 650 nucleotides (about 15% of the approximately 4500 nucleotides of Qβ RNA) has been established.

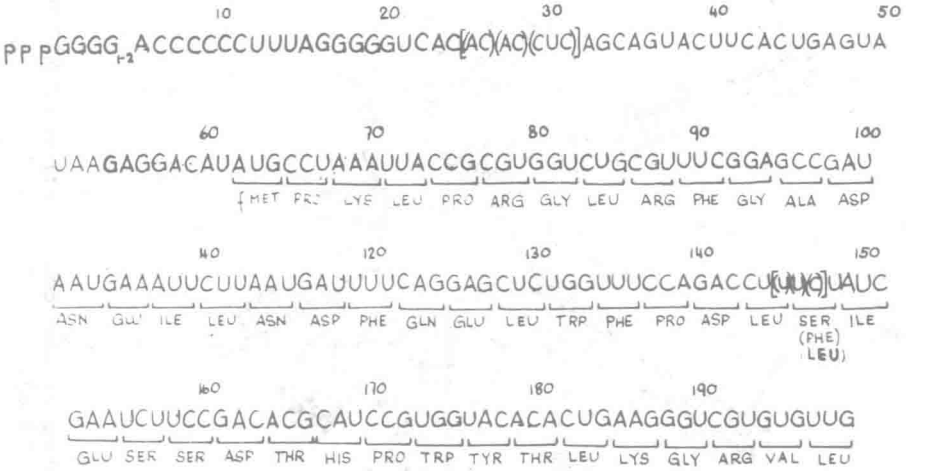

Fig. 6. Nucleotide sequence at the 5 terminal region of Qβ RNA. (Billeter et al., 1969b; Hindley, Billeter, Goodman and Weissmann, unpublished results). The aminoacid sequence Pro-Lys-Leu-Pro was found experimentally (Weiner and Weber, 1971); the remainder was deduced.

Secondary structure.

Large parts of the RNA segments whose sequences have been elucidated can be folded to give extensively hydrogen-bonded hairpins (Adams et al., 1969; Billeter et al., 1969b; Jeppesen et al., 1970b; Cory et al., 1970; De Wachter et al., 1971; Min Jou et al., 1972), using rules such as those ad-

vanced by Tinoco et al. (1971) (however cf. Delisi and Crothers, 1971). In the case of the MS2 coat cistron (Fig. 5) about 70% of the nucleotides are believed to take part in such secondary interactions (Fiers et al., 1971). The proposed structures are supported to some extent by the finding that, under conditions of partial digestion, nucleolytic splits are introduced predominantly (but not exclusively!) into single-stranded regions and into loops and bulges rather than into stem regions of the proposed hairpins.

Organization of the genome.

Since recombination does not occur among RNA phages, the order of their cistrons could not be elucidated by genetic techniques. This information was finally obtained by chemical techniques, in principle by searching for nucleotide sequences corresponding to the beginnings and ends of the phage proteins and locating the absolute or relative positions of these nucleotide sequences on the RNA strand (Jeppesen et al., 1970a; Hindley et al., 1970; Staples et al., 1971). In all cases, the order of cistrons is A-coat-replicase (cf. Fig. 1). Extended non-coding nucleotide sequences are located at the beginning (62 nucleotides in Qβ (Staples et al., 1971), 129 in the MS2-R17 group (Cory et al., 1970; Contreras et al., 1971; Adams et al., 1972)) of the RNAs, as well as between the cistrons (for example, 36 nucleotides between coat and replicase cistron of R17 (Nichols, 1970)).

Deductions regarding the genetic code.

Since the sequences of aminoacids of R17 coat protein and that of the nucleotides in the corresponding cistron are known, it is possible to deduce, at least in part, the genetic code independently of previous data. All assignments made in this fashion agree with the Nirenberg-Ochoa-Khorana code. In addition, it was found that for certain aminoacids not all degenerate codons are used in the translation of the coat. For example, AUU and AUC are each used 5 times for isoleucine, while AUA is not used at all; tyrosine is coded for 4 times by UAC, and never by UAU (Jeppesen et al., 1970b; Fiers et al., 1971).

Recognition of initiation sites of protein synthesis by ribosomes.

The nucleotide sequences around the initiation sites of six different phage cistrons are now known (Hindley and Staples, 1969; Steitz, 1969; Gupta et al., 1970; Staples et al., 1971; Staples and Hindley, 1972; Steitz, 1972). Initiation of phage cistrons is at an AUG triplet in all cases examined, except for the A-cistron of MS2 where GUG occurs (Fiers, personal communication). No extensive sequences are common to all initiation sites, however limited similarities are found in the regions preceding the initiator triplet of some cistrons. Thus, for example, GGUUUGA occurs prior to both the A and the coat cistron of R17, and UUUGA before the

Qβ coat cistron; GAGGA precedes the A cistron of Qβ, GAGG the A cistron of R17, GAGGAU the replicase cistron of R17 and AGGAU the replicase cistron of Qβ. It appears that neither the initiator triplet AUG (Steitz, personal communication) nor the sequence to the right of it (Weber et al., 1972) are required for ribosome binding. If a particular nucleotide sequence rather than some 3-dimensional structure is required for recognition by ribosomes it would seem that more than one such sequence is possible. Another remarkable aspect of ribosomal binding sites is that some are directly available (the coat cistron binding sites in both R17 (Steitz, 1969) and Qβ (Hindley and Staples, 1969)) while others are cryptic, i.e., not available on the mature, native RNA strand (Robertson and Lodish, 1970). For instance, the A cistron binding site of Qβ only binds E.coli ribosomes when present on an RNA fragment or on a nascent strand (Staples et al., 1971). It has been proposed that this unavailability of the binding site is due to the secondary and/or tertiary structure of the RNA (Gussin, 1966; Gussin et al., 1967; Lodish and Robertson, 1969; Lodish, 1971), and in at least one case structural evidence is forthcoming to support this view. Lodish and Robertson (1969) showed that ribosomes do not initiate polypeptide synthesis at the replicase cistron of native viral RNA unless the coat cistron is being translated. They proposed that physical interaction between the replicase initiation site and a region within the coat cistron prevented ribosome attachment to the initiation site. Translation of the coat cistron was thought to be instrumental in reversing this interaction, thereby allowing translation of the replicase cistron to be initiated. Fiers and his colleagues (1972) have pointed out that a complementary relationship in fact exists between an RNA segment of the coat cistron (nucleotides 72 to 96) and the initiation site of the replicase cistron (cf. Fig. 5), lending strong support to the hypothesis described above.

Interaction of Qβ replicase with Qβ RNA.

The viral replicases show a very high template specificity for the homologous, intact viral RNA (Haruna and Spiegelman, 1965) as well as for the complementary minus strands (Feix et al., 1968) to the exclusion of all other, unrelated viral RNAs and most other RNAs examined. It has been suggested that the extracistronic regions at the termini of the viral RNAs are responsible for the recognitive interaction with the replicase (Cory et al., 1970; De Wachter et al., 1971), and evidence (Rensing and August, 1969) has been offered that integrity of the 3' terminus up to the penultimate nucleotide (the terminal A is certainly expendable (Weber and Weissmann, 1970)) is required for template function. In order to explore the structural features required for template function, Spiegelman and his colleagues have taken advantage of the observation that Qβ replicase also replicates (a) species of "6 S" RNA detected in Qβ infected E.coli (Banerjee et al.,