Portraits of Viruses: RNA Phage

N.D.

Norton D.

Zinder

Rockefeller University, New York, N.Y.

Address inquiries to: Dr. Norton D. Zinder, Rockefeller University, 66th Street and York Avenue, New York, NY 10021 (USA)

Contents

Introduction

The Particle

Life Cycle

The Genome

Outside the Cell

In vitro Protein Synthesis

Amber Mutant Suppression and Protein Chain

Termination RNA Replication and RNA Replicase The Sequence

Again the Genes, Especially the One Found Last The Origin of RNA Phages

Introduction

The RNA phages were discovered accidentally. A group of phages had been isolated on the basis of their ability to grow on male *Escherichia coli* [Loeb, 1960]. One finding more than any other led to an early chemical analysis of these phages: the titer of the raw lysates was about 10^{12} /ml. Purification was rather simple but also surprising. 20% saturated ammonium sulfate solutions precipitated the in-

¹ This is the fourth in a historical series on the development of virology.

fectivity quantitatively without loss of titer. However, centrifuging at the usual 60,000 g left the phage in the supernatant. At each step a diphenylamine test for DNA and an orcinol test for RNA were done. 100,000 g for 2 h pelleted the infectivity. A technician was most concerned that the pellets had little DNA but lots of RNA 'contamination'. It was most tantalizing. However, we had lost about 10-fold in infectivity, some to mechanical loss and some to inactivation of phage. Dependent upon the amount of nucleic acid per particle, the amount of DNA still present could have explained the residual infectivity. We were sure we had a small phage, as there were small pellets and large titers. The year was 1960 and the only small phage known was ψ X174, described only 1 year before [Sinsheimer, 1959]. Φ X174 made large plaques, had a small burst size and not only was it not male-specific but it did not grow at all on *E. coli* K12. Thus, our phage was not ϕ X174. We subjected the phage to CsCl density gradient centrifugation, setting the average density at the 1.5 of the typical phage. A phage band of about 3 mm width was just below the top of the gradient (phage's $\sigma = 1.4$).

With purified phage at hand, chemical analysis soon revealed that the phage, f2, had to contain an oxypentose sugar and, by pre-

258

Zinder

cedent, probably RNA. Further analysis showed the sugar to be ribose. We had isolated the first RNA phage [Loeb and Zinder, 1961].

The Particle

The molecular weight of the f2 phage particle is 4.0×10^6 , not much larger than a 70S ribosome. It is 25 nm in diameter and icosahedral in shape. The RNA is single-stranded (1.3×10^6) and, except for a small deficiency in adenine, has almost equal amounts of the four bases. A double-stranded RNA virus was found several years later, as if it were necessary to fill in the table of nucleic acid types. The RNA sits in the center of the phage particle, surrounded by 180 subunits of a single species of coat protein molecule (13,700 daltons) and one molecule of maturation protein (40,000 daltons) [Zinder, 1965; Gomatos and Tamm, 1963].

Attachment of the phage to the bacterium is accomplished by a virion protein called the maturation protein. Crawford and Gesteland [1964] showed that phage attachment occurred to special bacterial pili conferred by the F agent when in *E. coli*. Prior to the f2 experiments there was no hint of the existence of sex pili. It is the absence of pili on female bacteria that accounts for the male specificity of f2 and not some other intrinsic property; phage RNA can infect females as well as males [Zinder, 1965].

The phage attaches to the sides of these long pili and, when negatively stained, they can be visualized in the electron microscope (fig.1). The maturation protein is cleaved and enters the host cell with the RNA [Kozack and Nathans, 1971; Krahn et al, 1972]. During the brief period between attachment and penetra-

tion, the RNA is sensitive to RNase and infection can be aborted. But precisely how the RNA penetrates is unknown, perhaps via the central core of the pilus.

RNA phages that infect *E. coli* have been isolated from sewage in countries all over the world: Japan, Australia, Germany, England, Canada and all over the USA. The phages are in four different serological groups. However, the majority of studies have been accomplished with f2, MS2, and R17 of group I and QB of group IV. QB is some 15% larger than f2, both in particle and RNA size. RNA phages have also been isolated that can infect *Pseudomonas* and several other bacterial species. All RNA phages are quite similar: they differ only by a few percentage points in size and they all require a pilus for attachment [Zinder, 1975].

Life Cycle

The nature of the problems facing any virus are delineated by its life cycle. Following infection, f2 is eclipsed as RNA begins to be synthesized. 15 min after infection, the first new virions appear, and small amounts of phage begin to be released at 22 min. The cells do not lyse completely, but most of the phage is in the medium by 70 min. Of the 20,000 progeny per cell, between 10 and 30% are infective.

At a minimum, the following physiological events must occur to produce phage progeny. Since *E. coli* has no RNA-dependent RNA polymerase, the phage must specify one. Enough enzyme must be made to synthesize 20,000 genomes. Moreover, since RNA is the template for protein synthesis, sufficient RNA message must be available. The newly synthesized RNA is not only the message but also

Portraits of Viruses: RNA Phage

Fig. 1. Electron micrograph of particles of RNA phage f2 attached to an F-pilus. The thin filaments in the field are the filamentous phage f1 (800 nm) which attach to the end of a pilus. Negatively stained with phosphotungstic acid. $\times 125,000$.

the progeny genome. Each cell must produce 20,000 molecules of maturation protein and more than 5×10^6 molecules of coat protein. Just prior to lysis, hexagonal and square nets of phage crystals appear. Phage are then released through holes in the cell membrane and wall. The phage is responsible for inducing some kind of lytic mechanism [Zinder, 1968].

The Genome

One million daltons of RNA can code for polypeptides with a total of 1,200 amino acids.

This is scarcely larger than the number of amino acids in the enzyme β -galactosidase, whose function is the cleavage of a simple glycosidic linkage. Few functions can be encoded in so small a genome.

Experiments with amber and temperature-sensitive mutants revealed three complementation groups. Each could be correlated with a particular phage function: one gene specified a major coat protein, one gene a minor coat protein, and another an RNA replicase. One particular mutant failed to complement both coat and replicase mutants. It proved to be a mutation in the proximal portion of the coat gene and had a polar effect on the replicase

260

Zinder

5' H a H b H c H 3'
 o co in ϕ o ϕ --oj m ϕ
 γ o o ro r-CM ipo ϕ n to
 •— 10 to tor- r- σ i ro in
 ^- ■ ”“ — ‘ ro ro

Fig. 2. Genetic map of group I RNA phages. Genes (rectangles) are drawn more or less in proportion to their size. A = Maturation gene; B = coat gene; C = replicase gene; D = lysis gene. The numbers on the line below sum the nucleotides from the 5' end of the RNA molecule to the

significant sites on the genetic map above. (After *Fiers et al.* [1976], *Atkins et al.* [1979], and *Beremand and Blumenthal* [1979].)

gene (see below) [Horiuchi et al, 1966; Gussin, 1966; Horiuchi and Matsuhashi, 1970].

A fourth complementation group was defined by a mutant that grew only in UGA suppressing strains. It failed to lyse the nonpermissive hosts while filling them with thousands of normal phage particles [Beremand and Blumenthal, 1979; Atkins et al., 1979; Model et al, 1979]. Many years passed before this mutation was discovered to be in a gene whose reading frame was different from other genes and overlapped two of the known genes.

The RNA phages do not undergo molecular recombination. Therefore, the order of the genes and the position of mutations could not be readily obtained. Ultimately, it was by structural rather than genetic means that the RNA was mapped (fig. 2).

Outside the Cell

When possible, it is highly informative to study a biological phenomenon outside an organism, a tissue or a cell. Controlled analyses are possible and components more accessible.

There are, of course, disadvantages in that integrative activities may disappear, topographical restraints are more relaxed and concentrations of components distorted.

In vitro Protein Synthesis

The RNA phages arrived on the scene at a rather propitious time. The idea that there were RNA messages transcribed from the genome and translated on any cellular ribosome – the messenger RNA hypothesis – had superseded the notion that different proteins were specified by different ribosomes [Monod and Jacob, 1961]. In addition, extracts of *E. coli* were becoming rather efficient at incorporating amino acids into proteins [Lamborg and Zamecnik, 1960]. The use of phage RNA as a message not only helped to reveal important aspects of phage biology but also many aspects of protein biosynthesis.

In 1961, f2 RNA [Nathans et al, 1962], TMV RNA [Tsugita et al, 1962], and poly-uridylic (poly-U) acid [Nirenberg and Matthaei, 1961] were added to extracts of bacteria capable of synthesizing protein. Addition of f2 RNA led to an understanding of the mechanisms for protein synthesis. The addition of poly-U led to the discovery of the triplets of the genetic code, a profound result. The addition of TMV RNA showed that not every RNA was functional in *E. coli* S30 (supernatant 30,000 g extracts), as no recognizable products were made.

The f2 studies were done in my laboratory and in that of *L. D. Watson*. No single discovery was made in one laboratory alone. Although the procedures were often different, the results were always the same. This competition was hard on the nerves, particularly those of the students.

However, since

Portraits of Viruses: RNA Phage

261

most results had been replicated in two laboratories, they were quickly and generally accepted. 70% of the protein synthesized on f2 RNA was its coat protein. This was unexpected since the 400 nucleotides needed to code for it amounted to only 11 % of the RNA's coding capacity.

Other larger proteins were seen by labeling with histidine, an amino acid not present in the coat [Nathans et al, 1962]. The problem with replicase synthesis (63,000 daltons) is that it terminated prematurely, giving fragments of many different sizes and hence hard to identify.

Ultimately, replicase or its fragments amounted to 30% of the protein synthesized [Lodish, 1968]. No maturation protein could be found. Evidently, the RNA in the virion coded for most, if not all, of the phage protein products.

Amber Mutant Suppression and Protein Chain Termination

Two amber mutations in the coat protein gene were most revealing. Sites No. 6 and No. 50 of the coat protein had glutamine residues [Notani *et al.*, 1965; Weber *et al.*, 1966]. An amber mutation (CAG to UAG) at site No. 6 was polar *in vivo*, while the one at site No. 50 was not [Zinder *et al.*, 1966]. Both mutants grew well on supA strains, and the amino acid serine was found at the respective sites. *In vitro*, RNA from the site No. 6 mutant directed the synthesis of an amino terminal hexapeptide in Sup⁻ extract [Engelhardt *et al.*, 1965]. Some full-sized coat was made in a Sup⁺ extract. Obviously, the components needed for suppression could be assayed. This mutant's coat protein not only terminated prematurely but its RNA did not direct the incorporation of any other amino acids; it was also polar *in vitro* [Zinder *et al.*, 1966]. The other mutant also directed the synthesis of an amino terminal fragment, this time of 50 amino acids. In addition, it directed the synthesis of replicase protein; it was not polar. The properties and ease of assay of the small peptide Fmet -ala - ser - asn - phe - thr led to a series of insights about protein biosynthesis, (i) Chain termination involved release of the peptide from tRNA [Engelhardt *et al.*, 1965]. (ii) Suppression of mutants to allow chain propagation was effected through mutant tRNAs [Engelhardt *et al.*, 1965; Capecchi and Gussin, 1965]. Different suppressors were different mutant tRNAs and inserted such amino acids as serine, tyrosine and glutamine at the amber codon. (iii) Suppression not only promoted chain propagation but also relieved polarity [Zinder *et al.*, 1966]. (iv) At the site of a chain-terminating event, ribosomes dissociated from the phage RNA and the tRNA from the peptide [Webster and Zinder, 1969].

Retrospectively, these points are understood by our detailed knowledge of protein synthesis. However, it was the experiments with phage RNA that led to many of these conclusions. In addition, it was shown that amber mutations were chain-terminating, that chain termination occurred at the site of the amber mutation, and that UAG was a chain-terminating codon. The initiation of protein synthesis was more mysterious. On the one hand, addition of any polyribonucleotide to an *in vitro* protein-synthesizing system directed the incorporation of amino acids [Nirenberg and Matthaei, 1961]. However, phage RNA directed the synthesis of products with the amino acids in a precise and proper order [Nathans *et al.*, 1962]. The sole difference in the extracts used with the two different types of message was that one had a high concentration of magnesium

262

Zinder

while the other had a low concentration. High Mg^{++} concentrations were found to cause spurious initiation. A clue to faithful initiation appeared when the prematurely terminated coat protein hexapeptide was found to have a blocked amino terminus. Removal of the blocking group by mild acid hydrolysis indicated that it was a formyl group. We remembered that, a year before, Marcher and Sanger [1964] had found a formylated methionine tRNA. Before we had finished the experiment designed to test the role of N-formyl methionine, James Watson came flying into the laboratory and told us that Capecchi had found N-formyl methionine on the amino terminus of *in vitro* synthesized phage protein. Within a month the experiments were finished and two papers sent off to the *Proceedings of the National Academy of Sciences of the USA* [Adams and Capecchi, 1966; Webster *et al.*, 1966]. Remarkably, of the myriad of known and unknown factors required for protein synthesis in these cell extracts, only the de-formylase had died, leaving the formyl group on the peptides ready to be discovered. A quick perusal of the

composition of the *E. coli* proteins led to the conclusion that all *E. coli* proteins started N-formyl methionine [Capecci, 1966]. Sometimes the formyl group and sometimes both the formyl group and the methionine were removed in maturing the protein. A special tRNA Fmet is used in the initiation of protein synthesis for all prokaryotes. It was some years before an initiating tRNA for eukaryotes was found. It also carries methionine but is not formylated, which makes it difficult to find [Smith and Marcker, 1970]. This difference between prokaryotes and eukaryotes is strange in that the addition of a formyl group seemed to be a clever ploy by nature to form the first peptide bond in solution. Protein synthesis would therefore have a unique direction. Perhaps Fmet tRNA was important in early evolution when primitive protein synthesis needed direction and is only a relict today. The identification of 'Fmet tRNA' as a universal initiator did not solve the problem of the initiation of protein synthesis. The codon for Fmet is AUG, the codon for methionine is also AUG. Just what was the distinction between initiation and methionine? Obviously, it must be differences between nearest neighbor nucleotides, either in their sequence or in their structure. Some clarification came by analyzing phage RNA ribosome-binding sites.

As RNA phage analysis was proceeding, developments in related areas often made significant experiments possible just at the appropriate time. One example was cited above: the isolation of Fmet tRNA without understanding its function. Two parallel developments led to unravelling the structure of the ribosome binding sites. Lengyel [Gupta et al, 1974] was studying the binding of phage RNA (the only real messenger RNA available for a long time) to *E. coli* ribosomes. Sanger and his colleagues [Barrell, 1971] were developing the procedures for sequencing small amounts of radioactively-labeled RNA. Steitz combined these two techniques and isolated ribosome-bound phage RNA that was RNase-resistant. The bound RNA fragments were sequenced. Steitz [1969] found three ribosome-binding sites corresponding to the three major gene products. The most prominent ribosome-binding site was that of the coat protein gene. Those for the replicase and maturation genes were present in much lesser amounts. They were found only when the RNA was broken by radioactive decay of the ³²P label. The sequences were somewhat disappointing. No common feature that would identify this from some other unknown sequence as a ribosome-

Portraits of Viruses: RNA Phage
263

binding site was apparent [Steitz, 1972]. It is now some 70 sequences later, and the most that can be said is that in addition to AUG (occasionally GUG), a ribosome-binding site has 4-5 purines some 5-10 bases to the left of the AUG [Shine and Dalgarno, 1974]. The purines are often capable of base pairing with the nucleotides at the 3' end of the 30S ribosomal RNA [Steitz and Jakes, 1975]. Currently, this is believed to be the mechanism for specific initiation, although it is probably incomplete.

The examples described above of the application of phage RNA to studies of protein biosynthesis were chosen to represent some of its singular contributions. Almost incidentally, the fidelity of the *in vitro* system was verified and since then other messenger RNAs have been studied reliably. Chain initiation and chain termination were important aspects of the genetic code that were not readily studied by the use of polyribonucleotides. Many principles and components involved in protein biosynthesis were first seen by the use of polyribonucleotide messages. Complementary studies with phage RNA established a firm basis for our understanding of protein biosynthesis [Lucas-Lenard and Lipmann, 1971].

The *in vitro* system also revealed many phage properties. The sizes of the three major phage proteins were defined by their size mobility when electrophoresed in a gel. Since they were the same as the proteins made *in vivo*, the RNA in the virion must be the messenger strand. Moreover, the different proteins which were coded for by a single RNA molecule were made in different amounts, amounts somewhat in proportion to what was found in infected cells [Lodish, 1975]. Further use of the *in vitro* system to probe phage-specific processes will be discussed below.

RNA Replication and RNA Replicase

The first question asked about phage RNA replication was, 'is there a DNA intermediate?'.

Inhibitors of DNA synthesis did not block RNA replication, hence, there was no DNA intermediate for this phage [Cooper and Zinder, 1962]. 10 years later reverse transcription was found for the retroviruses [Temin and Baltimore, 1972].

Infection with f2 leads to the appearance of an RNA-dependent replicase [August et al., 1963; Weissman et al., 1963; Spiegelman and Doi, 1963]. For the type I phages, enzyme instability made purification very difficult. Following inoculation, enzyme activity increased for about half the latent period and then leveled off. Assay of coat and maturation proteins showed that they continued to be synthesized. Surprisingly, mutations in the coat gene led to excess production of enzyme, a loss of control [Lodish and Zinder, 1966]. This phenomenon was useful in obtaining enzyme for purification.

Once again, the *in vitro* protein-synthesizing system clarified a problem, that of replicase biosynthesis. *In vitro* addition of purified coat protein inhibited replicase synthesis [Eggen and Nathans, 1969; Robertson et al., 1968]. The initiation sequence of the replicase gene bound the added coat protein [Bernardi and Spahr, 1972]. The coat blocked ribosome attachment and consequently replicase synthesis, a very simple but efficient control mechanism. *In vivo*, when sufficient enzyme is made and protein synthesis should be turned towards the synthesis of much coat, the coat itself turns off replicase synthesis. A further result of decreased replicase synthesis is an increased production of plus RNA strands for message. If replicase increases in amount, it competes with ribosomes for plus strands and

264

Zinder

converts them into double strands, useless for protein synthesis.

RNA synthesis *in vivo* can be described only in rather gross terms. As for any single-stranded virus, asymmetric synthesis creates special problems. Early in infection the input strand directs the synthesis of some replicase and this in turn synthesizes the minus strand. For about one-fourth of the latent period, minus and plus strands are made in equal amounts. Thereafter, between 5 and 10 plus strands are made per minus strand [Weiss-mann et al., 1968a].

Some insight into the mechanism of RNA synthesis was developed by analyses of RNA synthesis *in vitro*. A group IV phage, Q β , was studied by Spiegelman [Haruna et al., 1963]. Its enzyme was very stable [Spiegelman et al., 1965]. The enzyme itself contained four components: the phage replicase, a protein from the small ribosome subunit S-I, and the factors required for elongation in protein synthesis, EF-Tu and EF-Ts [Kamen, 1970; Blu-menthal et al., 1972; Wahba et al., 1974]. Thus, three of four proteins in the replicase are cell proteins involved in protein biosynthesis. Their precise role in RNA synthesis is not clear. Still, their roles all involve RNA recognition. S-I is important for the proper messenger binding [Noller et al., 1971]. EF-Tu and EF-Ts carry the charged tRNAs from the supernatant to the ribosomes [Lucas-Lenard and

Lipmann, 1971]. It may be significant that the structure of the 3' end of the phage RNA resembles the cloverleaf of tRNAs. Indeed the 3' terminus of some viral RNAs can be charged by specific amino acids; they also have a tRNA-like structure [*Pinck et al, 1970*]. One might assume that their replicases are composed of similar cell components.

In addition to replicase, two factors also isolated from uninfected cells are necessary for full replicase activity [*Frunze de Fernandez et al, 1968; Fedoroff and Zinder, 1973*]. These factors are proteins without known functions in the cell. They are required only to make a minus strand copy of a plus strand [*Feix et al, 1968*], thus providing a means for regulating the ratio of plus to minus strands. The phage-specific intermediate in RNA synthesis is not double-stranded. In all stages of synthesis the nascent RNA strands are associated only at the growing point [*Spiegelman et al, 1968; Weissmann et al, 1968b*]. Since the host *E. coli* contains a potent double-stranded RNase [*Robertson et al, 1968*], double-strandedness might be dangerous. Indeed, when excess replicase is present some double-stranded RNA fragments are found [*Lodish and Zinder, 1966*]. It is believed that *in vivo* both the plus and minus strand templates are themselves single strands with nascent molecules as single-stranded tails. However, double-stranded material is readily found following the usual de-proteinization procedures. Apparently the strands hybridize to each other during extraction [*Eikhorn and Franklin, 1966; Robertson and Zinder, 1969*].

The fact that the replicative form is never double-stranded solves another problem, particularly for a nucleic acid whose major product is plus strands. Consider a viral strand acting as a template for the minus strand. Enzymes and nucleotides flow from its 3' end to its 5' end, giving a double-stranded molecule with the enzyme hanging on by its teeth at a now blunt-ended template. The enzyme might turn around and by a displacement synthesis make a new plus strand. It might once again repeat the process and synthesize a new minus strand. Shuttling back and forth, plus and minus strands would be made in equal numbers. However, in order to synthesize the requisite excess of plus strands the enzyme would

Portraits of Viruses: RNA Phage

265

have to fall off one end of the double strand, return to the other end, and repeat the displacement synthesis. The molecular complications of such models caused most investigators to believe that there must be two different enzymes, one to make plus strands and the other to make minus strands. It is far simpler to synthesize single strands.

The Sequence

Sequence studies of phage RNA were initiated by *Sanger's* group shortly after they developed the technology for sequencing tRNA [*Sanger et al, 1965*]. Pieces that were relatively resistant to RNase were isolated and sequenced. *Steitz* [1969] determined the sequence of the ribosome-binding sites, and *Fiers* [1975] decided to sequence the whole RNA molecule. Gene-by-gene the sequence appeared [*Fiers et al, 1976*].

There are a number of interesting features of the primary sequence. At both ends of the molecule, there are non-coding regions. Moreover, each gene is separated from the next by about 40 nucleotides, the first intergenic spaces to be noted. Two of the three genes start with AUG and one with GUG while they terminate with UAG or UAA. All of the synonymous codons are used and with frequencies that might be expected on a more or less random basis. *E. coli's* vocabulary is equivalent to the dictionary of genetic words.

Free phage RNA has considerable secondary structure. There is hardly a region of the 3,600 bases that cannot be folded on itself to form a not quite perfect set of from 20 to 100 base pairs. Significantly, no perfect 20-base-pair segments are found, perhaps because these would be susceptible to RNase III. The 3' end of the RNA can be folded into a tRNA-like cloverleaf. Appraising the three initiation regions, one finds that that of the maturation gene is surrounded by secondary structure, whereas that of the coat gene is open and accessible and that of the replicase gene is base-paired to the proximal sequence of the coat gene. This last finding explains why the coat gene must be translated prior to effective translation of the replicase gene. Ribosomes do not glide from gene to gene [Webster and Zinder, 1969]. They enter each of the ribosome-binding sites independently. Sites differ both in accessibility and in their affinity for ribosomes. Some of these elements play significant roles in determining the amount of the various proteins that are synthesized.

Again the Genes, Especially the One Found Last

The maturation gene is translated very poorly *in vitro* from phage RNA, possibly because the template is the tail of the replicating RNA molecule. This functions well *in vitro* and gives a ratio of one part maturation protein to six parts of coat [Robertson and Lodish, 1970]. The secondary structure of single-stranded RNA prevents it from directing the synthesis of maturation protein. Obviously, the single strands are more open on the tails of the replicating molecules. As infection proceeds, the ratio of replicating molecules to single strands decreases, and this change of ratio is reflected by a necessary increase in coat protein synthesis. Still, sufficient maturation protein is made [Robertson, 1975].

The coat gene seemed also to be involved with cell lysis. Coat mutants failed to lyse infected cells [Zinder and Lyons, 1968]. Why this was so was not apparent from any attribute of the coat protein. In addition to being in-

266

Zinder

involved with lysis the coat gene holds a 2-fold vise on the expression of the replicase gene. Translation of the early part of the coat gene is necessary for the translation of the replicase gene. Also, coat protein binds to RNA at the origin of replicase synthesis and blocks further translation, the test example to date of a specific translational repressor.

Replicase protein is made in large amounts – about 10% of coat protein – when the controls are released. Normally *in vivo* it is detectable only enzymatically. This protein is also postulated to have a second function. Binding studies [Weber *et al.*, 1972] show that replicase binds to the ribosome-binding site of the coat gene. In addition, *in vitro*, replicase inhibits translation of phage RNA [Kolakofsky and Weissmann, 1971]. Thus, enzyme could prevent a few plus strands from participating in protein synthesis and thereby provide template for minus strand synthesis. For 10 years the lysis function was defined solely by a single UGA mutant [Model *et al.*, 1979]. This mutant failed to lyse infected cells although they contained a more than normal complement of phage. Recently, from *in vivo* experiments Beremand and Blumenthal [1979] noted a phage-specified 75 amino acid pep-tide. In addition, Atkins *et al.* [1979] translated phage RNA in a system composed of rabbit reticulocyte ribosomes and ascites cells supernatant and found the same peptide. In both instances it disappeared with the use of the mutant phage RNA. The gene for this peptide was located in a plus one phase starting at an AUG codon, 47 codons from the distal end of the coat gene continuing through 36 codons of intergenic space and on into 142

codons of the replicase gene. The UGA mutation was the result of a C to U change in the second position of the second codon of the replicase and within its ribosome-binding site [Atkins *et al.*, 1979]. Several puzzling features of the lysis mutant became clear. Its delayed growth was probably due to slow replicase production. However, its overproduction of replicase in time was probably the result of poor coat protein binding, hence poor repression. The position of the lysis gene also explains why mutations affecting coat gene translation lyse cells poorly: the ribosome-binding site of the lysis gene probably never becomes available.

The mechanism by which the lysis protein lyses the cells is obscure. It is not in any way a lysozyme. The problem is not unlike that for the single-stranded DNA phages such as Φ X174. It also has a small lysis gene and it is also located in a second reading frame within another gene [Barrell *et al.*, 1976]. Perhaps the lysis function of these phages was only needed after the size of the genome had been fixed by the particle size. Therefore, they had to enlist a protein produced by occasional misreading from within an existent gene. In time the advantage of being able to leave efficiently infected cells would select for more and more gene-like properties of the primitive lysis gene.

Packed into the small RNA phage genome are four genes, perhaps one more than the minimum needed to make a phage. In addition, there is a set of regulatory functions that give in time minus and plus strands in proper ratio, as well as the needed amounts of the four different proteins. A last example of such regulation would be the small amount of lysis protein made. It must be controlled by a very modest affinity of its ribosome-binding site for ribosomes. Thus, the functional efficiency of these phages is a marvel of natural selection. They have been carefully crafted such that almost each nucleotide performs several functions and in proper proportion.

Portraits of Viruses: RNA Phage

267

The Origin of RNA Phages

Consideration of the origin of phages is at best a difficult exercise. There is a dearth of facts for analysis; the only certain one is that for a phage to survive, a fully functional host cell must exist. Some DNA phages might be considered as vehicles promoting genetic exchange between cells. A segment of bacterial chromosome producing a protein that can en-capsidate DNA joined with the *cis* elements of bacterial replication implies a primitive phage.

Speculation on the origin of the RNA phages has even less substance. Still, several possibilities present themselves. It is generally assumed that the earliest life form was an inefficiently replicating RNA molecule. Consequently, it is not unlikely that the first organized cell had an RNA genome. RNA phage could then be derived from either of two sources. It could be a legacy of the evolving primitive RNA that found survival by becoming a parasite. Alternatively, it could reflect an initial effort at sexuality by the early RNA genomes much like that described for DNA. Thus, RNA phage may be very, very old, a remembrance of the time when genetic material was RNA. Their continued evolution when DNA became the dominant genetic material had as its major stress the maintenance of the RNA replicating system in the face of powerful DNA replicating systems. It is perhaps for this reason that the phage replicases are highly specific and will replicate only homologous molecules [Spiegelman and Doi, 1963].

RNA phages could also have arrived in a DNA world, possibly derived from messenger RNA. Still, it is conceptually difficult to provide *de novo* an efficient mechanism for RNA replication. Perhaps they arose at an early

time when nucleic acid replicases had a low order of specificity and synthesized both RNA and DNA. This imprecise system could evolve to synthesize either RNA or DNA. While it is difficult to put RNA phage into a fully satisfying context, it is clear that it has evolved in exquisite adaptation for itself and to its host. If they are as old as postulated, they have had infinite time to refine ever more sharply their few functions. As we find them today, as long as there are bacteria, there will be RNA phage.

References

- Adams, J. M. and Capecchi, M.R.: N-formyl-meth-ionyl-sRNA as the initiator of protein synthesis. *Proc. natn. Acad. Sci. USA* 55: 141 (1966).
- Atkins, J.; Steitz, J. A.; Anderson, C, and Model, P.: Binding of mammalian ribosomes to MS2 phage RNA reveals an overlapping gene encoding a lysis function. *Cell* 18: 247 (1979).
- August, J.T.; Cooper, S.; Shapiro, L., and Zinder, N.D.: RNA phage induced RNA polymerase. *Cold Spring Harb. Symp. quant. Biol.*, vol. 28, p. 95 (Cold Spring Harbor Laboratory, Cold Spring Harbor 1963).
- Barrell, B.G.: Fractionation and sequence analysis of radioactive nucleotides; in Cantoni and Davis, *Procedures in nucleic acid research*, vol. 2, p. 751 (Harper & Row, New York 1971).
- Barrell, B. G.; Air, G. M., and Hutchison, C. A.: Overlapping genes in bacteriophage ψ X174. *Nature, Lond.* 264: 34(1976).
- Beremand, M. and Blumenthal, T.: Overlapping genes in RNA phage: a new protein implicated in lysis. *Cell* 18: 257(1979).
- Bernardi, A. and Spahr, P.: Nucleotide sequence at the binding site for coat protein on RNA of bacteriophage R17. *Proc. natn. Acad. Sci. USA* 69: 3033 (1972).
- Blumenthal, T.; Landers, T., and Weber, K.: Bacteriophage Q β replicase contains the protein biosynthesis elongation factors EF Tu and EF Ts. *Proc. natn. Acad. Sci. USA* 69: 1313 (1972). 268
- Zinder
- Capecchi, M.R.: Initiation of *E. coli* protein. *Proc. natn. Acad. Sci. USA* 58: 1517 (1966).
- Capecchi, M. and Gussin, G.: The identification of a serine sRNA as a nonsense suppressor. *Science* 149: 111 (1965).
- Cooper, S. and Zinder, N.D.: The growth of an RNA bacteriophage: the role of DNA synthesis. *Virology* 75:405 (1962).
- Crawford, E. M. and Gesteland, R.: The adsorption of the bacteriophage R17. *Virology* 22:165 (1964).
- Eggen, K. and Nathans, D.: Regulation of protein synthesis directed by coliphage MS2 RNA. II. In vitro repression by phage coat protein. *J. molec. Biol.* 39: 293 (1969).
- Eikhorn, R. and Franklin, R.: Formation and properties of a replicative intermediate in the biosynthesis of viral ribonucleic acid. *Bact.Rev.* 30: 267 (1966).
- Engelhardt, D.; Webster, R.E.; Wilhelm, R.C., and Zinder, N.D.: In vitro studies on the mechanism of suppression of a nonsense mutation. *Proc. natn. Acad. Sci. USA* 54: 1791 (1965).
- Fedoroff, N. and Zinder, N.D.: Factor requirement of the bacteriophage f2 replicase. *Nature new Biol.* 241: 105 (1973).
- Feix, G.; Pollet, R., and Weissmann, C.: Replication of viral RNA. XVI. Enzymatic synthesis of infectious viral RNA with noninfectious Q β minus strands as template. *Proc. natn. Acad. Sci. USA* 59: 145 (1968).

Fiers, W.: Nucleotide sequence studies of MS2 RNA; in Zinder, RNA phages, p. 353 (Cold Spring Harbor Laboratory, Cold Spring Harbor 1975).

Fiers, W.; Contreras, R.; Duerinck, F.; Haegeman, G.; Iserentant, D.; Merregaert, J.; Min Jou, W.; Molemans, F.; Raeymaekers, A.; van den Berghe, A.; Volckaert, G., and Ysebaert, M.: Complete nucleotide sequence of bacteriophage MS2 RNA: primary and secondary structure of the replicase gene. *Nature, Lond.* 260: 500 (1976).

Franze de Fernandez, M.T.; Eoyang, L., and August, J.T.: Factor fraction required for the synthesis of bacteriophage Q β RNA. *Nature, Lond.* 219: 588 (1968).

Gomatos, P. and Tamm, I.: The secondary structure of reovirus RNA. *Proc. natn. Acad. Sci. USA* 49: 707 (1963).

Gupta, S.L.; Chen, J.; Schaefer, L.; Lengyel, P., and Weissman, S.M.: Nucleotide sequence of a ribosome attachment site of phage f2. *Biochem. bio-phys. Res. Commun.* 39: 883 (1974).

Gussin, G.: Three complementation groups in bacteriophage R17. *J. molec. Biol.* 21: 435 (1966).

Haruna, I.; Nozu, K.; Ohtaka, Y., and Spiegelman, S.: An RNA 'replicase' induced by and selective for a viral RNA: isolation and properties. *Proc. natn. Acad. Sci. USA* 50: 905 (1963).

Horiuchi, K.; Lodish, H., and Zinder, N.D.: Mutants of the bacteriophage f2. VI. Homology of temperature-sensitive and host-dependent mutants. *Virology* 28: 438 (1966).

Horiuchi, K. and Matsushashi, S.: Three cistrons in bacteriophage Q β . *Virology* 42: 49 (1970).

Kamen, R.: Characterization of the subunits of Q β replicase. *Nature, Lond.* 228: 527 (1970).

Kolakofsky, D. and Weissmann, C.: Q β replicase as repressor of Q β RNA-directed protein synthesis. *Biochim. biophys. Acta* 246: 596 (1971).

Kozack, M. and Nathans, D.: Fate of maturation protein during infection by coliphage MS2. *Nature new Biol.* 234: 209 (1971).

Krahn, P. M.; O'Callaghan, R. J., and Paranchych, W.: Stages in phage R17 infection. VI. Injection of A protein and RNA into the host cell. *Virology* 47: 628 (1972).

Lamborg, M. and Zamecnik, P. C.: Amino acid incorporation by extracts of *E. coli*. *Biochim. biophys. Acta* 42: 206 (1960).

Lodish, H.: Bacteriophage f2 RNA: control of translation and gene order. *Nature, Lond.* 220: 345 (1968).

Lodish, H.: Regulation of in vitro protein synthesis by bacteriophage RNA; in Zinder, RNA phages, p. 301 (Cold Spring Harbor Laboratory, Cold Spring Harbor 1975).

Lodish, H. and Zinder, N.D.: Mutants of the bacteriophage f2. III. Control mechanisms for phage-specific syntheses. *J. molec. Biol.* 19: 333 (1966).

Loeb, T.: Isolation of a bacteriophage specific for F and Hfr mating types of *E. coli* K12. *Science* 131: 932 (1960).

Loeb, T. and Zinder, N.D.: A bacteriophage containing RNA. *Proc. natn. Acad. Sci. USA* 47: 282 (1961).

Lucas-Lenard, J. and Lipmann, F.: Protein biosynthesis. *A. Rev. Biochem.* 40: 409 (1971).

Marcker, K. and Sanger, F.: N-formyl methionyl-sRNA. *J. molec. Biol.* 8: 835 (1964).

Model, P.; Webster, R.E., and Zinder, N.D.: Characterization of op3, a lysis defective mutant of the bacteriophage f2. *Cell* 18: 235 (1979).

Portraits of Viruses: RNA Phage

Monod, J. and Jacob, F.: Teleonomic mechanisms in cellular metabolism, growth and differentiation. Cold Spring Harb. Symp. quant. Biol., vol. 26, p. 389 (Cold Spring Harbor Laboratory, Cold Spring Harbor 1961).

Nathans, D.; Notani, G.; Schwartz, J., and Zinder, N. D.: Biosynthesis of coat protein of coliphage f2 *E. coli* extracts. Proc. natn. Acad. Sci. USA 48: 1424 (1962).

Nirenberg, M. W. and Matthaei, J. H.: The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. Proc. natn. Acad. Sci. USA 47: 1588 (1961).

Noller, H.F.; Chang, C; Thomas, G., and Aldridge, J.: Chemical modification of the transfer RNA and polyuridylic acid binding sites of *E. coli* 30S ribo-somal binding site subunits. J. molec. Biol. 61: 669 (1971).

Notani, G.; Engelhardt, D.; Konigsberg, W., and Zinder, N.D.: Suppression of coat protein mutant of bacteriophage f2. J. molec. Biol. 12: 43 (1965).

Pinck, M.; Yot, P.; Chapeville, F., and Duranton, H.: Enzymatic binding of valine to TYMV RNA. Nature, Lond. 226: 954 (1970).

Robertson, H.D.: Functions of replicating RNA; in Zinder, RNA phages, p. 113 (Cold Spring Harbor Laboratory, Cold Spring Harbor 1975).

Robertson, H.D. and Lodish, H.: Messenger characteristics of nascent bacteriophage RNA. Proc. natn. Acad. Sci. USA 67: 710 (1970).

Robertson, H.D.; Webster, R.E., and Zinder, N.D.: Bacteriophage coat protein as repressor. Nature, Lond. 218: 533 (1968).

Robertson, H.D. and Zinder, N.D.: Purification and properties of nascent f2 phage ribonucleic acid. J. biol. Chem. 244: 5790 (1969).

Sanger, F.; Brownlee, G. G., and Barrell, B. G.: A two dimensional fractionation procedure for radioactive nucleotides. J. molec. Biol. 13: 373 (1965).

Shine, J. and Dalgarno, L.: The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. natn. Acad. Sci. USA 71: 1342 (1974).

Sinsheimer, R.L.: Purification and properties of bacteriophage ϕ X174. J. molec. Biol. 1: 37 (1959).

Smith, A.E. and Marcker, K.: Cytoplasmic methionine transfer RNAs from eukaryotes. Nature, Lond. 226: 607 (1970).

Spiegelman, S. and Doi, R.H.: Replication and translation of RNA genomes. Cold Spring Harb. Symp. quant. Biol., vol. 28, p. 109 (Cold Spring Harbor Laboratory, Cold Spring Harbor 1963).

Spiegelman, S.; Haruna, I.; Holland, B.; Beaudreau, G., and Mills, D.: The synthesis of a self-propagating and infectious nucleic acid with a purified enzyme. Proc. natn. Acad. Sci. USA 54: 919 (1965).

Spiegelman, S.; Pace, N.; Mills, D.R.; Levisohn, R.; Eikhorn, T.; Taylor, M.; Peterson, R.L., and Bishop, D.H.L.: The mechanism of RNA replication. Cold Spring Harb. Symp. quant. Biol., vol. 33, p. 101 (Cold Spring Harbor Laboratory, Cold Spring Harbor 1968).

Steitz, J.: Polypeptide chain initiation: nucleotide sequences of three ribosome binding sites in bacteriophage R17. Nature, Lond. 224: 957 (1969).

Steitz, J.: Oligonucleotide sequence of replicase initiation site in Q β RNA. Nature new Biol. 236: 71 (1972).

Steitz, J. and Jakes, K.: How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. Proc. natn. Acad. Sci. USA 72: 4734 (1975).

Temin, H. and Baltimore, D.: RNA-directed DNA synthesis and RNA tumor viruses; in Smith, Lauf-fer and Bang, Advances in virus research, vol. 17, p. 129 (Academic Press, New York 1972).

Tsugita, A.; Fraenkel-Conrat, H.; Nirenberg, M.W., and Matthaei, J. H.: Demonstration of the messenger role of viral RNA. Proc. natn. Acad. Sci. USA 48: 846 (1962).

Wahba, A. J.; Miller, N.; Niveleau, A.; Landers, T. A.; Carmichael, G.C.; Weber, K.; Hawley, D.A., and Slobin, L.I.: Subunit I of Q β replicase and 30S ribosomal protein SI of *E. coli*. J. biol. Chem. 249: 3314 (1974).

Weber, H.; Billeter, N.A.; Kahane, S.; Weissmann, C.; Hindley, J., and Porter, A.: Molecular basis for repressor activity of Q β replicase. Nature new Biol. 237: 166 (1972).

Weber, K.; Notani, G.; Wickner, N., and Konigsberg, W.: Amino acid sequence of f2 coat protein. J. molec. Biol. 20: 423 (1966).

Webster, R.E.; Engelhardt, D., and Zinder, N.D.: In vitro protein synthesis chain initiation. Proc. natn. Acad. Sci. USA 55: 155 (1966).

Webster, R.E. and Zinder, N.D.: Fate of message-ribosome complex upon translation of termination signals. J. molec. Biol. 42: 425 (1969).

270

Zinder

Weissmann, C.; Colthart, L., and Libonati, M.: Determination of viral plus and minus ribonucleic acid strands by an isotope dilution assay. Biochemistry, N.Y. 7: 865 (1968a).

Weissmann, C.; Feix, G., and Slor, H.: In vitro synthesis of phage RNA: the nature of the intermediates. Cold Spring Harb. Symp. quant. Biol., vol. 33, p. 83 (Cold Spring Harbor Laboratory, Cold Spring Harbor 1968b).

Weissmann, C.; Simon, L.; Borst, P., and Ochoa, S.: Induction of RNA synthetase after infection by the RNA phage MS2. Cold Spring Harb. Symp. quant. Biol., vol. 28, p. 99 (Cold Spring Harbor Laboratory, Cold Spring Harbor 1963).

Zinder, N.D.: RNA phages. A. Rev. Microbiol. 19: 455 (1965).

Zinder, N. D.: Phage RNA as genetic material. Harvey Lect. Ser. No. 62 (Academic Press, New York 1968).

Zinder, N. D. (ed.): RNA phages (Cold Spring Harbor Laboratory, Cold Spring Harbor 1975).

Zinder, N.D.; Engelhardt, D., and Webster, R.: Punctuation in the genetic code. Cold Spring Harb. Symp. quant. Biol., vol. 31, p. 251 (Cold Spring Harbor Laboratory, Cold Spring Harbor 1966).

Zinder, N.D. and Lyons, L.: Cell lysis: another function of the coat protein of bacteriophage f2. Science 159: 84 (1968).