Chemistry and Structure of Nucleic Acids of Bacteriophages

Many forms of nucleic acids of bacteriophages show the ways that information is stored and reproduced.

J. A. Cohen

The chromosomes of microorganisms, particularly those of viruses, are structurally the simplest and consist merely of nucleic acid. Among the viruses the bacteriophages (phages) are the most appropriate for study in view of the ease with which they can be cultured, purified, and tested. Each phage particle contains one molecule of nucleic acid, which may be either deoxyribonucleic acid (DNA) or ribonucleic acid (RNA). These can be isolated from the virus particles and provide extremely homogeneous nucleic acid preparations which are very appropriate for investigation of physical and chemical properties. In the intact virus particle which attacks the bacterial host, the nucleic acid will normally be protected by a protein coat; often the coat is essential to secure the penetration of the nucleic acid, by providing an injection mechanism. After penetration virus multiplication starts. First the nucleic acid molecules are multiplied; they are eventually packed in the protein coats that are also newly synthesized. This "burst" of newly made mature bacteriophage particles then leaves the host, usually after lysis (*1*). However, in a number of cases, it is possible for isolated, naked nucleic acid molecules to induce virus production; to accomplish this they have to be brought within the complex environment of a living cell. There they not only multiply, but also give rise to full-fledged virus particles complete with protein coat. The nucleic acid molecule of a bacteriophage particle will therefore contain information necessary for its own replication (template action), for the production of the proteins of the coat, and for the assembly of nucleic acid and protein into an active virus particle. Usually the synthesis of the enzymes responsible for the replication of nucleic acid, that is, the polymerases and the early enzymes, which switch the host metabolic machinery into action on behalf of this nucleic acid replication, precedes the formation of enzymes required for the synthesis of the coat proteins and the assemblage of the matured nucleic acids within the organized phage particle. The information contained within the nucleic acid molecule of the phage is brought to expression inside the host. The bacteriophages supply the most simple models of the many ways genetic information is stored, reproduced and brought to expression.

The nucleic acid molecule remains unchanged during the extracellular stage of the phage. However, after entering the host bacterium it may be modified before or after acting as a template for reduplication; some of the resulting structures have been described as replicative forms (RF).

Glucose and Methyl Residues on the Nucleic Acids of Bacteriophages

The nucleic acids of bacteriophages occur either as double-stranded DNA arranged in the Watson and Crick double spiral, as single-stranded DNA or as RNA. The molecular size may vary from as few as 1 million to as many as 130 million dalton units. The bacteriophages with double-strand DNA of the Watson and Crick double spiral configuration have accordingly equal amounts of A (adenine) and T (thymine) and of G (guanine) and C (cytosine) repectively (complementarity of base composition) linked by hydrogen bridges. A number of them may contain unusual pyrimidines; 5-methylcytosine or 5-hydroxymethylcytosine may replace cytosine, whereas in other cases thymine residues may be replaced by 5-hydroxymethyluracil or even by uracil. Complementarity of base composition does not necessarily exist in single-strand phages; it does not occur in the single-stranded DNA of the *Escherichia coli* bacteriophage ΦX174 or in RNA bacteriophages which are also single stranded. This does not imply that these single-stranded nucleic acids may not have secondary structures on account of stretches of hydrogen-bonded material with double-strand character. The double-stranded DNA's of the bacteriophages T2, T4, and T6 have been most extensively investigated with regard to the chemical composition of their bases. They contain, instead of cytosine, 5-hydroxymethylcytosine, the hydroxyl group of which may be either free or glucosylated. The various phages of this class (T2, T4, T6) differ in the amount and in the type of sugars bound to hydroxymethylcytosine (*2*). In T2, 25 percent of the hydroxymethylcytosine residues are free, 70 percent are bound to α-glucose, and 5 percent are bound to gentiobiose. In T4, all hydroxymethylcytosine residues are bound to glucose, 70 percent in α- and 30 percent in β-configuration, whereas, in T6, again 25 percent of these residues are free, 3 percent are bound to α-glucose and 72 percent bear gentiobiose molecules.

Kornberg and his associates (*3*) have indicated that, upon infection with T-even phage, specific glucosylating enzymes are induced which catalyze the transfer of glucose residues from uridinediphosphoglucose (UDPG) to hydroxymethylcytosine groups present in the already-formed polynucleotide chain. The patterns described above are dictated in part by the specificity of the various hydroxymethylcytosine-glucosyltransferases (*4, 5*). In the case of T4-DNA, differences in the affinity of the hydroxymethylcytosine for the α- and β-glucosyltransferases are determined by the adjacent nucleotide sequences (*5, 6*). The secondary structure of the intracellular phage DNA at the site of glucosylation (*3*) or the local concentration of magnesium-ions, or both, may be responsible for the constant ratio of α- and β-glucose, which

Dr. Cohen is director of the Medical Biological Laboratory of the National Defense Research Organization TNO and professor of physiological chemistry, Applied Enzymology and Radiobiology of the University of Leiden, The Netherlands.

does not seem to be influenced by the composition of the growth medium (*6*). The information for the glucosyltransferases resides on the phage DNA, as judged by the fact that mutants defective in the α-transferases have been isolated from T2 and T6 bacteriophage (*7*).

Just as glucosylation seems to be a reaction which takes place after the whole unglucosylated DNA has been formed, methylation—at least in a number of cases—represents the finish of the phage DNA molecule. N^6-Methyladenine (*8*) has been found in T2-DNA; it is formed by the action of an induced methylase (*8, 9*). Likewise T4 and T7 induce a DNA methylase; neither T6 nor T5 have this property (*8*). In the case of T3 the absence of methyl groups is due to the presence of an enzyme which breaks down *S*-adenosylmethionine (*10*), which is the normal methyl donor.

From recent data it seems likely that glucosylation and methylation are connected with host-cell modification and restriction. These terms indicate the ability of the bacterial host to affect in some instances the phenotype of the infecting bacteriophage. Luria and Human observed (*10a*) in 1952 that, when bacteriophages T2 or T6 were propagated on a certain *Escherichia coli* strain B/4, the progeny was hardly infectious for the normal host for these phages, *E. coli* B. This phenomenon was called restriction of the phage by the host. If *Shigella dysenteriae* was infected with the same progeny particles, the normal phage resulted. This showed that the property which had been lost and regained by passage through two consecutive hosts (B/4 and *Sh*), affected only the phenotype. Apparently the bacteriophages can be modified within certain hosts. Since it is only the DNA of the phage which penetrates into the host cell, modification and restriction were believed to occur in the DNA.

A connection between host-induced modification and restriction on the one hand and glucosylated DNA on the other hand was found almost simultaneously in three different laboratories (*11*). It turned out that *Escherichia coli* B/4 is unable to synthesize uridinediphosphoglucose and therefore cannot satisfy the requirement of the glucosyltransferases which are induced upon infection with phage. Nonglucosylated T-even bacteriophages can also be produced by the growth on phage mutants which lack the ability to form α-glucosyltransferase (*7, 12*). These unusual phages are characterized by their failure to grow in any but a few *E. coli* B cells but they multiply in some strains of *Shigella dysenteriae*. It could be demonstrated that the inability of both types of nonglucosylated phages to grow on the restrictive host is accompanied by rapid degradation of the nonglucosylated phage DNA occurring in the restrictive host (*12, 13*). These observations suggest that the main biological function of the glucose residues in the DNA of T phages may be to protect the molecule against premature degradation.

In other cases of host-cell modification and restriction, the chemical basis of this phenomenon has been assigned to the state of methylation rather than that of glucosylation of the phage DNA. The best-studied example is a system (*14*) in which lambda phages are modified and restricted by certain hosts. Extensive genetic study during several years has produced evidence pertaining to the chemical basis of this type of modification (*15, 16*). It was demonstrated that lambda phages (grown on certain hosts) probably could not multiply in other hosts because of a lack of methyl groups on their DNA (*16*). Bacteriophage lambda lacking certain methyl groups was formed when methionine deficient (met$^-$) (*17*) strains of a certain mutant of strain K12 were deprived of methionine during part of the latent period. Now, in contrast to the normal situation, the efficiency of plating of the progeny phage proved to be lower on the host, K12, than on strain C (a nonrestrictive host for various modified forms of lambda). These results suggest that host specificity is conferred by methylation of specific sites on the lambda-DNA. As in the case of nonglucosylated T-DNA, the nonmethylated lambda-DNA may be subject to fragmentation in the restrictive host (*18*). However, Gold *et al.* (*19*) could not demonstrate a difference between phage lambda B and lambda B devoid of methyl groups on various strains of *Escherichia coli*, each of which had its own restriction and modification pattern. Therefore at this stage we cannot draw a final conclusion concerning the role of methylated bases in host-cell modification.

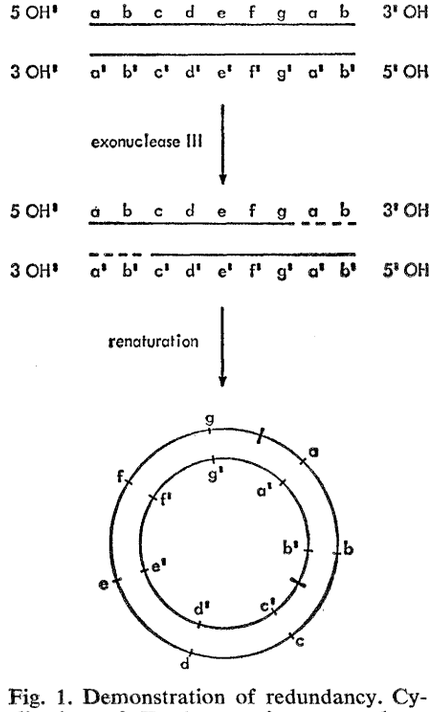

Fig. 1. Demonstration of redundancy. Cyclization of T phages after exonuclease treatment and renaturation.

Chemistry and Structure of the Nucleic Acids of Free Phages

Free infective bacteriophage particles contain either a molecule of DNA or of RNA. So far DNA-RNA hybrids have not been encountered in free phage. RNA bacteriophages (such as MS2, f2, R17, M12) contain a linear single-stranded RNA molecule of slightly more than 1 million daltons. DNA may also occur in single-stranded form in bacteriophages. The first DNA bacteriophage containing single-stranded DNA discovered, was *Escherichia coli* phage ΦX174 (*20*). This single-stranded DNA of noncomplementary base composition is present in a circular form (*21*). The long, rod-shaped phages M13 and fd, specific for male *E. coli*, also contain single-stranded DNA. Like the DNA of ΦX174, the molecular weight is relatively low (between 1.5 and 2×10^6 daltons). Phage fd has also a circular DNA (*22*).

The best-studied bacteriophages are those which contain a double-stranded DNA molecule of larger size than those previously mentioned (up to a molecular weight of approximately 130×10^6 for T2- and T4-DNA). The DNA of bacteriophage lambda has a molecular weight of 33×10^6. Lambda has a particular significance in relation to the phenomenon of lysogeny. Lysogenic bacteria are those which have taken up in their genome at least part of a phage genome and are lysed after induction by ultraviolet light or other agents. This phage DNA multiplies concomitantly with that of the host bacteria. The

phenomenon is of obvious interest because of the analogy with virus tumors such as polyoma, adenovirus, and SV40 virus tumors. These mammalian tumors consist of cells which contain, in addition to DNA of the cell type, probably also DNA characteristic of the virus. This may be inferred from the presence of RNA in these cells, which hybridizes with virus DNA (*23*). In another phenomenon, called transduction, a modified phage is involved, insofar as it has a DNA content different from that of the original bacteriophage. Lambda is a phage whose locus on the chromosome of *E. coli* K12 is closely linked to genes controlling several enzymes of galactose metabolism (*24*). Ultraviolet irradiation of bacteria that are lysogenic for lambda initiates synthesis and release of phage particles. Most of the new phage particles are ordinary lambda. However, approximately 1 out of 10^6 of the new particles is lambda dg (degenerate). Lambda dg transduces galactose genes from the bacterium in which it was produced to the bacterium it infects. Part of the genes of these lambda dg particles are thus replaced by incorporation of galactose genes from the bacterium in which they were produced (*25*).

Apart from accessory information originating from genes of the bacterial host, as in transducing phages, phage DNA may also carry redundant information, that is, information relating to bacteriophage genes present in more than the required quantity. This form of enlarged DNA occurs normally in wild-type T phages. All the available physical and chemical evidence leads to the conclusion that in T phages the double-stranded DNA is present as a linear rod. From the work of Thomas (*26, 27*), it now appears that in all these linear DNA molecules a certain sequence of nucleotides in the beginning of the molecule repeats itself at the end. The terminal repetitive ends represent approximately 1 to 3 percent of the total DNA molecule. This terminal repetitive structure was demonstrated as follows. The linear DNA molecules of T2, T3, or T7 were first treated during a short period with the enzyme exonuclease III. This enzyme acts specifically on double-stranded DNA and successively removes nucleotides of each chain, starting at the 3′-hydroxyl group terminals. The complementary character of the remaining single-strand ends of the molecule became evident from the fact that, under renaturation conditions, intramolecular association of both ends occurred, resulting in stable circular molecules (Fig. 1) which could be observed by electron microscopy. The opposite polarity of the 3′,5′-diphosphate bonds between the deoxyriboses in both rings should be considered. Cyclization (Fig. 1) applies to the DNA of T3 and T7 phage, where the nucleotide sequence of all molecules in a population is the same, and to the DNA molecules of T2 and T4, where the nucleotide sequence is different in various molecules of the same phage population; in the case of T4, DNA is present in a circularly permuted form. In this situation, the various molecules may start with any of the letters *a* to *g*. The difference between T2 and T4, which is circularly permuted on the one hand and T3 and T7 which are not circularly permuted on the other hand, was demonstrated by Thomas from the renaturation behavior after denaturation (complete separation of the two strands) of the native DNA molecules (*26, 27*). It was demonstrated that both T2 and T4 can be cyclized by annealing, after strand separation by alkali or heat. Apparently the annealing procedure yields circles in the case of T2 and T4, but not in the case of T3, T7, and T5. As indicated in Fig. 2, cyclization of the denatured single-strand molecules requires overlapping of complementary strands, which can only be expected in the case of molecules that are permuted such as T2- or T4-DNA but not with T3-, T5-, and T7-DNA. In the case of T5 another reason for noncyclization is the existence of breaks at preferential points (*28, 29*). Thomas (*26*) has suggested that the T5-DNA molecule contains four gaps, one in the chain *a* and three in the chain *b*; these gaps appear to be true interruptions. With the exception of bacteriophages T5 and PB, the phages thus far studied by Thomas, including T2, T4, P1, P22, and lambda, contain polynucleotide chains that are continuous over most of the length of the molecule.

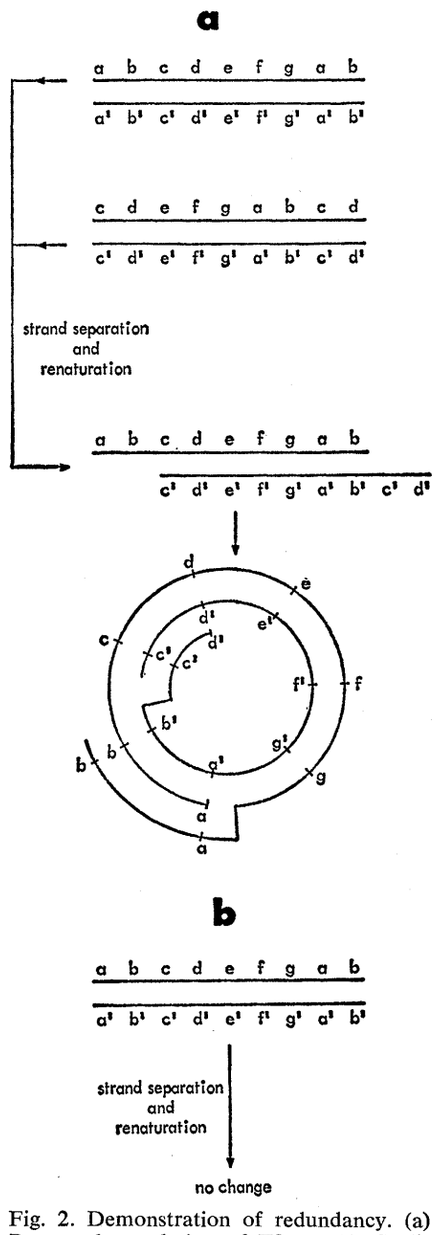

Fig. 2. Demonstration of redundancy. (a) Permuted population of T2 or T4. Cyclization occurs by renaturation after strand separation. (b) Unique population of T3 or T7. Cyclization not possible by renaturation after strand separation.

Investigations on the structure of the DNA of bacteriophage lambda are very interesting. The DNA that can be obtained from this bacteriophage is a linear molecule of double-strand structure, and it has a molecular weight of 33×10^6 (*30*). This DNA, when brought in solution, has an unusual tendency to form aggregates (*29*). From sedimentation analysis it appeared that aggregate formation by lambda-DNA can be formulated as follows:

Open polymers ⇌
linear monomers ⇌ folded monomers

Since the folded molecules appear in one stable configuration only and since folding and polymer formation seem to be mutually exclusive, Hershey *et al.* (*29*) postulated that every linear DNA molecule contains two sites which are mutually cohesive and which are occupying remote sites (for example, at the end of the molecules). These cohesive sites are assumed to occupy only a small fraction of the total length of the molecule. This fine analysis stimulated more study of the molecular

structure of the folded monomer. By the use of electron microscopy it could be demonstrated that the folded monomer is a circular molecule of the same length as that of the linear molecule; and from this it follows that folding should be ascribed to cohesion between head and tail of the linear monomer (*30, 31*). This model was confirmed by the finding that linear molecules of lambda-DNA which had been halved by shear could be made to aggregate to linear molecules of the original length but not to oligomers (*32*). Strack and Kaiser (*33*) showed that the cohesive ends of lambda-DNA consist of short terminal complementary single-strand parts of DNA, as if the linear molecules were derived from a circular one by the induction of single-strand breaks in complementary strands at sites which are not exactly opposite (*33*) (Fig. 3). The complementary parts which are absent in the linear monomer could be completed by use of DNA polymerase. This treatment abolished the cohesive character of the ends. Cohesiveness could be restored by partial degradation of the synthetic product by means of the enzyme exonuclease III. From the known specificity of the enzymes and the number of nucleotides which had to be removed for complete restoration of cohesiveness, Strack and Kaiser concluded that each cohesive end was made up of about ten nucleotides and terminated in a 5′-hydroxyl (or phosphate) group. This number of ten agrees with thermodynamic data on the cyclization reaction of the linear monomer (*34*). Like the DNA of the T-odd phages most of the lambda-DNA molecules have unique sequences (*35*).

Bacteriophage Nucleic Acids inside the Infected Bacterial Host

Bacteriophage nucleic acids inside the extracellular phage particles are covered by a protective protein coat and are not subject to any metabolic changes. However, when the bacteriophages inject their nucleic acids into the host cell, these give rise to a number of changes in the bacterial host, to replication of the phage nucleic acid molecules, and to transcription leading to expression in messenger RNA; all combined result in the production of infectious progeny-phage particles and often in the death of the host. During this process the injected phage nucleic acids may undergo alteration in structure, which is probably necessary to make them suitable as templates for replication and transcription. Thereafter the replicated nucleic acid molecules are subject to a maturation process that eventually leads to the form of nucleic acid which, packed in the protein coat, gives rise to the infective free bacteriophage particle that finally leaves the bacterial host cell.

A specific problem arises with single-stranded forms of nucleic acid, RNA or DNA. From what is known about the replication of genetic material and the transcription into RNA, it appears that for both processes the complementary strand is necessary as a template. In replication the strands to be formed should be identical and not complementary with regard to the infecting DNA or RNA strand, and therefore the infecting strands cannot serve in the function of template. The same applies, at least in the cases thus far specifically studied, with regard to the transcription of single-strand DNA into messenger RNA; this also must be formed on a template complementary to the parent DNA single strand. It now seems that at least in a number of cases the problem is solved in the sense that this single-strand nucleic acid, after entering the host cells, gives rise to a double-stranded structure, the newly formed strand of which may fulfill the template requirement for both replication and transcription. In the case of RNA phages, the data are not as straightforward as in the case of DNA phages. Evidence has been provided that two enzymes are involved in the replication of phage RNA (*36*). One converts single-strand into double-strand RNA, while the other uses double-strand to synthesize single-strand viral RNA. During synthesis the new viral strand displaces its predecessor, while the complementary strand serves as a permanent template (*37, 38*). This replication mechanism may lead to two different forms of double-stranded RNA in infected bacteria, which have different sedimentation characteristics. One of these has been first described for RNA bacteriophage M12 by Hofschneider's group (*38*) and conforms probably to the structure drawn in Fig. 4a; the second represented in Fig. 4b has been described by Franklin (*37, 39*). These structures have been confirmed by electron microscopic data (*40*) on the bacteriophage R17. In bacteria infected with this RNA phage of molecular weight 1.1×10^6, several unique types of RNA were found. These were the single-stranded viral RNA (*41*), double-stranded RNA termed replicative form, and double-stranded RNA with branched single-stranded components, termed replicative intermediate (*37*). So far, single- and double-stranded structures of RNA phage found in free bacteriophage or in infected cells appear to be linear. Spiegelman's group showed that the RNA phages MS2 and Qβ each induce an RNA replicase after infection of *Escherichia coli* (*42, 43*). Both replicases require intact homologous RNA and are inactive with heterologous RNA. From the kinetics of the reaction, it is concluded that the product synthesized by the enzymic reaction can in turn induce the production of complete virus particles in bacterial protoplasts. When purified Qβ replicase is used to synthesize infectious RNA in vitro, there is a latent period which is marked by a loss of the ability of the initiating RNA template to induce plaque-forming units. This eclipse is accompanied by the appearance of a noninfectious complex having the sedimentation properties of the Hofschneider form and containing parental plus (+) strands and newly formed minus (−) strands. The product also agrees with the Hofschneider structure in being noninfectious but giving rise to infectious plus strands on denaturation by heat. Also structures resembling those described by Franklin with branched single-strand components are found in the early states of the synthesis in vitro of RNA (*42, 44*).

When circular single-stranded DNA from bacteriophage ΦX174 is injected into the host cell, a constant formation of infective replicative form (RF) is initiated, whereby the strand complementary to the parental strand in the duplex is synthesized from precursors derived directly from the medium. The duplex containing the parental strand replicates semiconservatively; it seems to be the only template for the production of the more replicative form (*45*). This replicative form is a double-stranded circular DNA having a Watson-Crick structure. Its molecular weight is twice that of ΦX174 single-strand DNA, $2 \times 1.7 = 3.4 \times 10^6$. It occurs in two forms, component I, consisting of a double-stranded circular molecule in which both strands are continuous, with a sedimentation constant of approximately 21*S* and component II, consisting of double-stranded circular molecules,

which contain single-strand breaks (17*S*). The predominant, naturally occurring component is RF-I, whereas RF-II can be formed from RF-I by treatment with pancreatic deoxyribonuclease (*46, 47*). Both components are circular, as indicated by electron microscopy and other studies (*46, 48, 49*). Both the single-stranded DNA of the mature phage and the replicative forms are capable of infecting spheroplasts leading to the induction of ΦX174 phages (*50*).

After ultraviolet radiation of RF its capability to produce phage particles in spheroplasts of certain bacteria was far higher than that of single-strand ΦX174-DNA (*51*). The host bacteria involved contain enzymes that can restore damage by ultraviolet. These enzymes probably excise the ultraviolet-produced pyrimidine dimers (*52*). In single-strand DNA, this excision leads to inactivation, since for single-stranded ΦX174-DNA circularity is a prerequisite for biological activity (*53*). In double-strand DNA circularity is maintained after excision, while the opposite strand can serve as template for a repair polymerase which fills the gap. Therefore under normal conditions the double-strand DNA is far more resistant to ultraviolet than the single-stranded form, and this may also apply to damage other than that caused by ultraviolet. Messenger RNA has a structure similar to that of the parental plus (+) strand and not of the complementary minus (−) strand. In the RF structure the mechanism of strand selection seems to be built in, in view of the fact that in vitro RF-I (*54*), and probably RF-II too (*55*), gives rise only to RNA of the plus (+) type when incorporated with the RNA polymerase system. If broken circles are used as a template, the enzyme produces RNA strands that are complementary to both DNA strands (*54*).

The differences in physical properties between RF-I and RF-II are probably due to the unusual structure of RF-I which has both closed circular duplex DNA and a number of tertiary turns which convey some tension to the structure and which finally result in a twisted form. In RF-II, where at least one of the strands is broken, no such twisting can occur, since the single strand scissions would allow tertiary turns to unwind and relax the circular structure. These twists explain certain observations, such as the fact that RF-I is unusually found accompanied by a slower sedimenting circular DNA (RF-II) into which it can be converted by one or more single-strand scissions (for example by pancreatic deoxyribonuclease). These properties appear in other twisted circular duplex DNA's, such as the DNA's of the mammalian viruses polyoma, SV40, rabbit and human papilloma, and (after injection into the host) phage lambda (*56*). The best evidence comes from electron micrographs. In these, RF-I, consisting of the two intertwined continuous strands, shows a twisted structure whereas RF-II with one or more breaks forms open rings (*57, 58*) (Fig. 5). These micrographs look similar to those published earlier for polyoma DNA (*56*).

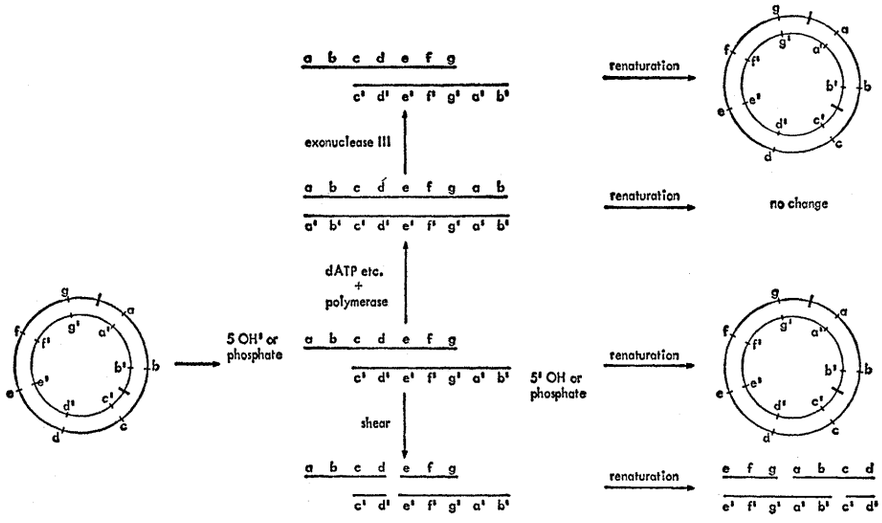

Fig. 3. Maturation of bacteriophage lambda (see text).

Almost all of the DNA injected by bacteriophage lambda into the host cell is very quickly converted into a noninfectious form, apparently owing to cyclization of the linear lambda-DNA. Thus, it could be demonstrated that after infection of lysogenic bacteria with lambda phage, of which the DNA was labeled with tritium, two structures other than the original linear DNA occurred, and these sedimented 1.14 and 1.9 times faster than the linear lambda-DNA (*59*). The first structure represents probably the same circular lambda-DNA that can be formed in vitro from linear DNA (by cohesion); the second component shows the features of a twisted circular structure, due to tertiary turns, as already described for ΦX174-DNA (*60*). During maturation, the twisted form has to be converted into the linear DNA structure, which is eventually packed in the phage coat to give progeny phage. Another circular, possibly intermediary, species with one single-strand break and no twists has been found by Ogawa and Tomizawa (*61*). In this connection an enzyme recently described by Hurwitz *et al.* (*62*) and Weiss *et al.* (*63*) seems most relevant. The enzyme identified in extracts of *Escherichia coli* infected with T4 phage catalyzes the repair of phosphodiester bond interruptions; it repairs single-strand breaks in double-strand DNA by sealing a 3′-OH onto a 5′-P group. The enzyme which requires adenosine

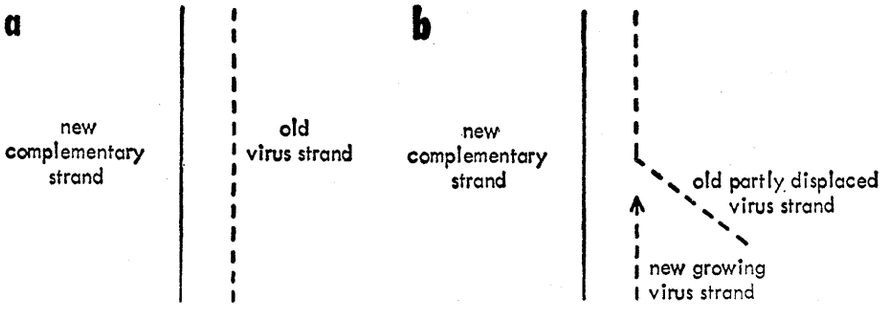

Fig. 4. Probable structures of double-stranded DNA induced by infecting bacteria with RNA bacteriophages or by RNA-replicase in vitro. (a) "Hofschneider structure"; (b) "Franklin structure."

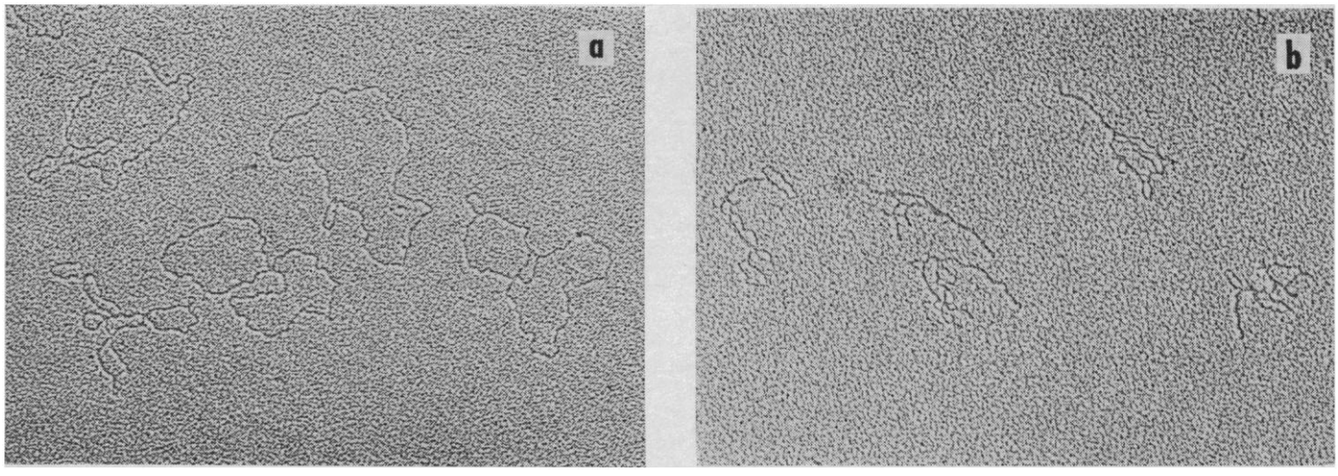

Fig. 5. (a) Open circles of circular RF-II DNA, obtained by treating RF-I with pancreatic deoxyribonuclease (Magnification × 78,000) (*58*). (b) Twisted, double strand, circular RF-I DNA of bacteriophage ΦX174.

triphosphate and magnesium ion for optimum activity is capable of converting the cohesive circular form of lambda-DNA into a circular form with uninterrupted polynucleotide chains.

Neither the fate of the DNA of phages of the T-series after infection of the host, nor the maturation process leading to the various forms of intracellular phage DNA, discussed earlier, are well understood. It was found (*64, 65*) that, at the earlier stages of replication, DNA displayed sedimentation characteristics expected for large macromolecules. The various authors agree that a bizarre structure is involved and that it contains single-stranded regions, as expected from its sensitivity to deoxyribonucleases, from its retention on certain columns, and from electron micrographs. Under the electron microscope, tangled structures often having a total length of two phage DNA units are observed. Frankel has given an explanation for the data available to him (*64*) on the basis of very long molecules resulting from continuous growth. Thomas (*66*) also proposes a giant molecule growing at one end and maturing at the other. However, using density labeling and electron microscopy Kozinski *et al.* (*65*) provided evidence that the maximum mass of the molecules formed on replication was never larger than two phage equivalents. Tomizawa (*67*) observed linear molecules composed of two components derived from the parental DNA joined end to end by hydrogen bonds. At this stage, it cannot be decided whether the long molecules involved are linear or circular.

It would evidently be very attractive, if the events occurring in the infected host cell could be understood as variations on a general theme. Thus it seems tempting to explain the various forms of DNA molecules, encountered in the free bacteriophage particles, as derivatives from a double-stranded circular form present in the infected host (Fig. 6). The circular double-stranded form may appear in the virus as in the case of polyoma, SV40, or papilloma viruses, without further modification. For phages some further maturation must be assumed, consisting in the case of ΦX174 in the semiconservative formation of a single-strand DNA and in the case of lambda, T3, and T7 in the introduction of two single-strand breaks at specific sites, followed by ring opening. In lambda, the resulting linear molecule would then be ready for delivery into the phage coat. For the T-uneven phages like T3 and T7 this only happens after the single-strand terminals have been completed with a complementary chain. This series of events would account for the nonpermuted molecules with repetitive terminals. When maturation would be initiated by the introduction of two single-strand breaks at random sites of the molecule, a linear particle, circularly permuted and with repetitive ends, would be the result after completion of the single-strand terminals (T2 and T4). The concept of a circular double-stranded intermediate in phage maturation is attractive, not only because it has been proven to occur in a number of instances and can be argued to occur in the way just described for T phages, but it also provides a good basis for genetic data. Genetic data, including those on T phages, with regard to the sequences of genes, map distances, and other factors are all in accordance with continuous repetition of gene sequences. Of course this continuous repetition required to explain genetic results can be easily accounted for by the occurrence of a circular intermediate. However, the same genetic results would be observed when the repetition was provided by linear structures carrying more than one phage DNA monomer in repetitive sequence. Obviously the concentrated intracellular forms, found experimentally after T4 infection, and recently also after lambda infection (*68*), support the latter possibility although the existence of circular intracellular structures for T4 phages cannot be excluded. The corresponding data on the RNA phage genome are too scarce to allow speculation about the existence of a circular intermediate.

An interesting complication is met in the case of bacteriophage T5-DNA, which contains interruptions at specific locations in both of its polynucleotide chains. Thomas described some experiments on the continuity of the T5-DNA molecules collected at various times after infection (*26*). The sedimentation rate of intracellular T5-DNA appears normal in neutral sucrose if the cells are broken open between 0 and 25 minutes after infection. However, after 4 minutes the single strands obtained by denaturation seem to be of equal length and move fast, an indication that they do not contain breaks. This suggests that all the gaps are sealed very soon after infection. In this case maturation would include the reintroduction of specifically located gaps.

Biology and Structure of the Bacteriophage Nucleic Acids

The biological function of the bacteriophage nucleic acid is the storage of information required to induce its own replication and expression within the host in order to effect the production and release of a burst of viable bacteriophage particles. This information is encompassed in the base sequences, that is, in the primary structure of the nucleic acid molecule. At first it seems surprising that the appropriate base sequences do not appear in the chromosome of each phage in a uniform, particularly suitable way, for example like double-strand DNA, hydrogen-bonded in Watson and Crick double spirals, a structure which occurs in practically all known organisms. However, the multiformity of nucleic acids of viruses, including bacteriophages, seems understandable at second thought. Each of the very many species of phage has to cope with many widely different factors in order to survive and to multiply. Although some of the various features described earlier may not represent any specific biological function, they may be remnants of evolutionary earlier functional components or just wasteful excesses of some metabolic activity. Earlier, the possible biological significance of host cell modification and restriction due to certain methyl groups was discussed. The methyl groups discussed already may endow the methylated structure with individual properties which, at least in some instances, serve to protect the genetic material against damage either outside or inside the host. The change of host range specificity may for various reasons be useful for the propagation of the phage. The same reasoning applies to the incorporation of glucose groups in the DNA. Here it could be shown specifically that protection against degradation will sometimes ensue.

A single-stranded genetic structure may sometimes be profitable, because it can carry maximum information within minimum mass. Moreover, it is much more flexible than double-stranded nucleic acid and can therefore be accommodated in a small and simply built protein coat. In this way small single-strand DNA and RNA phage particles with a reasonable amount of information are formed, and they can pass into host bacteria in a simple way (all phages with double-strand DNA are much larger and require an elaborate protein coat, providing a kind of

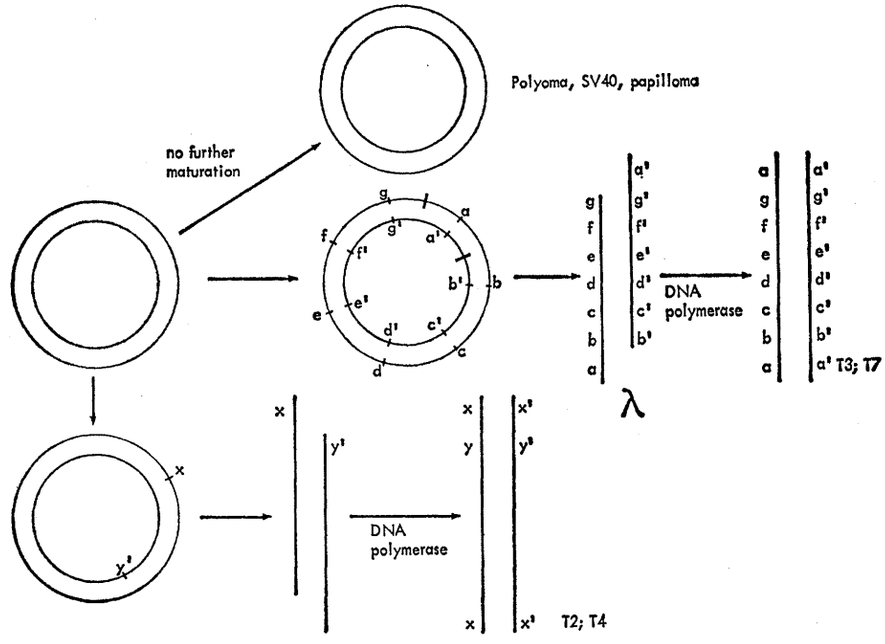

Fig. 6. Maturation of various intracellular phages.

injection mechanism to deliver their nucleic acid to the host cell). It is also understandable that the primary structure of the single-stranded material may sometimes be RNA and at other times DNA; the conditions to which the genetic material of a particular phage during its cycle will be exposed may sometimes be more appropriately met by DNA and at other times by RNA (for example, outside the host an alkaline *p*H may be more readily endured by a DNA than by an RNA molecule, whereas inside the host the relative content of active ribonucleases and deoxyribonucleases will influence the issue, which of the two nucleic acids stands the best chances to survive).

Also the variety of host cells with regard to metabolic equipment for replication and transcription is so large, that it seems likely that either of the two molecular species will be more useful in a particular instance. It should be remembered that an RNA phage introduces into the host a ready-made messenger which is replicated by a new, phage-induced, RNA-directed RNA polymerase. Single-stranded DNA must first be converted into double-stranded DNA by host DNA polymerase before transcription can take place; double-strand DNA replication requires a DNA polymerase, which may be host polymerase or, alternatively, be induced by the phage. The RNA-directed RNA polymerases, induced by RNA phages, which replicate the RNA, appear from Spiegelman's work (*44*) to be extremely specific enzymes, which will not be hampered in their action by competition by the excess of cellular RNA present. This template specificity is probably at least partly due to the ability of the RNA to assume a highly specific secondary structure possibly by cohesion of complementary sequences. The RNA inside bacteriophages seems always to be linear and the same may apply to the double-stranded replicative form inside the infected host.

The known single-stranded DNA phages contain a circular DNA component. A single break in the circular DNA is fatal to the biological activity of the phage. The circular structure of the single-strand DNA makes it less flexible than the linear single-stranded RNA, and this may be a disadvantage. On the other hand, circularity provides stability against the action of exonucleases, and the circular form is probably the most appropriate precursor for the also-circular RF, which is to be formed from it.

In contrast to the multiformity of the nucleic acids of the free phages, the intracellularly occurring forms, insofar as they consist of DNA and are involved in replication and transcription, might often be circular as well as double-stranded. This may comply with a general tendency for replicating and transcribing chromosomes to occur as circular and double-stranded structures. Circularity provides resistance against certain nucleases. A dou-

ble-strand structure would also add to the stability of the DNA molecule; double-stranded structures are not only mechanically more robust but they are also more resistant to many nucleases; moreover they contain the relevant information twice, once in each strand. These features make double-stranded DNA more resistant to ultraviolet and possibly also to other damages, as discussed before; double-stranded DNA but not single-stranded DNA may, in principle, be repaired with excising and repairing enzymes. Although animal viruses with double-stranded circular DNA are known (polyoma, SV40, and papilloma) no bacteriophages of this type have been found. For bacteriophages, the gain consisting in higher stability and a structure bearing great resemblance to the one to be evolved inside the cell is apparently not a sufficient compensation for the penalty to be paid in the form of loss of flexibility and the necessity of a more elaborate protein coat. When phage has to convey a large amount of information, as in the case of the T-even phages, it builds up a large DNA molecule, stabilized by double-strand structure, but not by circularity. This structure is injected through the narrow tube of the injection apparatus of the phage protein envelope. Probably circular double-strand DNA could not be so delivered, although it might otherwise have been the material of choice on behalf of its resistance against certain nucleases and its resemblance to known circular intracellular forms. The robust circular double-stranded structure which in at least a number, if not in all cases, has to be established inside the host is not present during the intraphage stage; however, sticky (single-strand) terminals (bacteriophage lambda) or redundant terminals, which may, after infection, be converted into "sticky" ones (T phages), could be considered as preparatory stages for conversion into a circular form. Another function of the redundant terminals may be to convey stability to the gene in view of the fact that some terminal damage of the DNA molecule may be incurred without permanent loss of function. Thus far, no bacteriophages containing double-stranded RNA have been found, although animal viruses of this type are known (reovirus, wound tumor virus).

The double-stranded DNA component in T-viruses usually has few if any single-strand breaks in either polynucleotide strand. A different situation is met in the case of T5-DNA, which has specific gaps in each of the two chains. The biological function of these gaps may be related to the observation that absorption of T5-DNA from the phage into the cell occurs in two sequential and separable steps. At first only 8 to 10 percent of the phage DNA, which probably arises from the phage nucleic acids by breakage at a preferential point, enters the cell; and in the second step the rest of the DNA follows (*69*). The first arriving DNA portion is responsible for the degradation of host-cell DNA within the infected cell and probably induces the synthesis of one or several proteins which are required for the transfer of the rest of the T5-DNA. It seems likely that the preferential breaking point is located at the site of one or more of the gaps in the DNA. About the biological function of the tertiary twists observed in the circular, continuous, double-stranded DNA of certain animal viruses and of bacteriophages, such as the intracellular lambda-DNA and the DNA of the replicating form of ΦX174 no data are available. They may provide landmarks or punctuations to be used by transcribing or replicating enzymes for initiation or termination of action, for strand selection or otherwise. It may also be that the near-linear shapes they give to the circular DNA molecule may in some way contribute to successful infection.

Summary

The nucleic acids of bacteriophages are characterized by a surprising multiformity. RNA and DNA may occur, the latter in single- or double-stranded form, circular or linear, with or without breaks or single-strand ends. Terminal redundancy may exist and the populations of linear phages may be uniform or randomly permuted. A double-stranded circular DNA does not occur in extracellular bacteriophage, but is often if not always formed after infection of the bacterial host. Phage DNA may be glucosylated or methylated to a certain extent, and the glucose and methyl residues may influence the stability of the DNA inside the host.

The biological meaning of the multiformity in structure of the bacteriophage chromosome may be sought in the diversity of the conditions that the various bacteriophages, and in general all viruses, have to comply with. In order to thrive they have to call upon many if not all available means for storage, reproduction, and expression of information.

References and Notes

1. The lysis of bacteria is their disruption, which is visible by clearing of suspensions or plates of medium, on which they are grown.
2. I. R. Lehman and E. A. Pratt, *J. Biol. Chem.* **235**, 3254 (1960).
3. A. Kornberg, *Enzymatic Synthesis of DNA* (Wiley, New York, 1962).
4. M. R. Lunt, J. C. Siebke, K. Burton, *Biochem. J.* **92**, 27 (1964).
5. A. de Waard, *Biochim. Biophys. Acta* **92**, 286 (1964).
6. ———, Th. E. C. M. Ubbink, W. Beukman, unpublished observations.
7. H. R. Revel, S. Hattman, S. E. Luria, *Biochem. Biophys. Res. Commun.* **18**, 545 (1965).
8. R. Hausmann and M. Gold, *J. Biol. Chem.* **241**, 1985 (1966).
9. D. Fujimoto, P. R. Srinivasan, E. Borek, *Biochemistry* **4**, 2849 (1965).
10. M. Gefter, R. Hausmann, M. Gold, J. Hurwitz, *J. Biol. Chem.* **241**, 1995 (1966).

10a. S. E. Luria and M. L. Human, *J. Bacteriol.* **64**, 557 (1952).

11. A. Shedlovsky and S. Brenner, *Proc. Nat. Acad. Sci.* **50**, 300 (1963); S. Hattman and T. Fukasawa, *ibid.*, p. 292; N. Symonds, K. A. Stacey, S. W. Glover, J. Schell, S. Silver, *Biochem. Biophys. Res. Commun.* **12**, 220 (1963).
12. S. Hattman, H. R. Revel, S. E. Luria, *Virology* **30**, 427 (1966).
13. T. Fukasawa and S. Saito, *J. Mol. Biol.* **8**, 175 (1964); S. Hattman, *Virology* **24**, 333 (1964); T. Fukasawa, *J. Mol. Biol.* **9**, 525 (1964).
14. G. Bertani and J. Weigle, *J. Bacteriol.* **65**, 113 (1963).
15. W. P. M. Hoekstra, thesis, Utrecht, 1966.
16. W. Arber, *J. Mol. Biol.* **11**, 247 (1965).
17. The met⁻ strains lack the capability of synthesizing methionine.
18. D. Dussoix and W. Arber, *J. Mol. Biol.* **5**, 37 (1962).
19. M. Gold, M. Gefter, R. Hausmann, J. Hurwitz, *J. Gen. Physiol.* **49**, 5 (1966).
20. I. Tessmann, *Virology* **7**, 263 (1959); R. L. Sinsheimer, *J. Mol. Biol.* **1**, 43 (1959).
21. W. Fiers and R. L. Sinsheimer, *J. Mol. Biol.* **5**, 424 (1962); D. Freifelder, A. K. Kleinschmidt, R. L. Sinsheimer, *Science* **146**, 254 (1964).
22. D. A. Marvin and H. Hoffmann, *Z. Naturforsch.* **18b**, 884 (1963).
23. T. L. Benjamin, *J. Mol. Biol.* **16**, 359 (1966); K. Fujinaga and M. Green, *Proc. Nat. Acad. Sci. U.S.* **55**, 1567 (1966); K. Fujinaga and M. Green, *Fed. Proc.* **26**, 450 (1967).
24. J. Lederberg and E. Lederberg, *Genetics* **38**, 51 (1953).
25. M. L. Morse, E. M. Lederberg, J. Lederberg, *ibid.* **41**, 142 (1956); J. Weigle, *Virology* **4**, 14 (1957); W. Arber, E. Kellenberger, J. Weigle, *Schweiz. Z. Allegem. Pathol. Bacteriol.* **20**, 659 (1957).
26. C. A. Thomas, Jr., *J. Gen. Physiol.* **49**, 143 (1966).
27. D. A. Ritchie, C. A. Thomas, Jr., L. A. MacHattie, P. C. Wensink, *J. Mol. Biol.* **23**, 365 (1967); L. A. MacHattie, D. A. Ritchie, C. A. Thomas, Jr., C. C. Richardson, *ibid.*, p. 355.
28. J. Abelson and C. A. Thomas, Jr., *J. Mol. Biol.* **18**, 262 (1966).
29. A. D. Hershey, F. Burgi, L. Ingraham, *Proc. Nat. Acad. Sci. U.S.* **48**, 748 (1963).
30. L. A. MacHattie and C. A. Thomas, Jr., *Science* **144**, 1142 (1964).
31. H. Ris and B. Chandler, *Cold Spring Harbor Symp. Quant. Biol.* **28**, 1 (1963).
32. A. D. Hershey and F. Burgi, *Proc. Nat. Acad. Sci. U.S.* **53**, 325 (1965).
33. H. B. Strack and A. D. Kaiser, *J. Mol. Biol.* **12**, 36 (1965).
34. J. C. Wang and N. Davidson, *ibid.* **15**, 111 (1966).
35. D. S. Hogness and J. R. Simmons, *ibid.* **9**,

411 (1964); A. D. Kaiser, *ibid.* **4**, 275 (1962).
36. H. F. Lodish and N. D. Zinder, *Science* **152**, 372 (1966).
37. R. L. Erikson and R. M. Franklin, *Bacteriol. Rev.* **30**, 267 (1966).
38. B. Francke and P. H. Hofschneider, *Proc. Nat. Acad. Sci. U.S.* **56**, 1883 (1966).
39. R. M. Franklin, *ibid.* **55**, 1504 (1966).
40. N. Granboulan and R. M. Franklin, *J. Mol. Biol.* **22**, 173 (1966).
41. N. K. Sinha, R. K. Fujimura, P. Kaesberg, *ibid.* **11**, 84 (1965).
42. N. Pace and S. Spiegelman, *Proc. Nat. Acad. Sci. U.S.*, in press.
43. I. Haruna, K. Nozu, Y. Ohtaka, S. Spiegelman, *Proc. Nat. Acad. Sci. U.S.* **50**, 905 (1963).
44. S. Spiegelman, 51st Annual Meeting Fed. of Amer. Soc. Exp. Biol., Chicago, April 1967, Symp. Nucleic acid metabolism and biological function.
45. D. T. Denhardt and R. L. Sinsheimer, *J. Mol. Biol.* **12**, 647, 663 (1965).
46. P. H. Pouwels, H. S. Jansz, J. van Rotterdam, J. A. Cohen, *Biochim. Biophys. Acta* **119**, 289 (1966).
47. H. S. Jansz and P. H. Pouwels, *Biochem. Biophys. Res. Commun.* **18**, 589 (1965).
48. A. K. Kleinschmidt, A. Burton, R. L. Sinsheimer, *Science* **142**, 96 (1963).
49. B. Chandler, M. Hayashi, M. N. Hayashi, S. Spiegelman, *ibid.* **143**, 47 (1964); A. Burton and R. L. Sinsheimer, *ibid.* **142**, 962 (1963).
50. G. D. Guthrie and R. L. Sinsheimer, *J. Mol. Biol.* **2**, 297 (1960); R. L. Sinsheimer, B. Starman, C. Nagler, S. Guthrie, *ibid.* **4**, 142 (1962).
51. H. S. Jansz, P. H. Pouwels, J. van Rotterdam, *Biochim. Biophys. Acta* **76**, 655 (1963).
52. R. P. Boyce and P. Howard-Flanders, *Proc. Nat. Acad. Sci. U.S.* **51**, 293 (1964); R. B. Setlow and W. L. Carrier, *ibid.* **51**, 226 (1964).
53. W. Fiers and R. L. Sinsheimer, *J. Mol. Biol.* **5**, 424 (1962); R. L. Sinsheimer, *ibid.* **1**, 37 (1959).
54. M. Hayashi, M. N. Hayashi, S. Spiegelman, *Proc. Nat. Acad. Sci. U.S.* **51**, 351 (1964).
55. J. A. Cohen, unpublished results.
56. J. Vinograd and L. Lebowitz, *J. Gen. Physiol.* **49**, 103 (1966).
57. T. F. Roth and M. Hayashi, *Science* **154**, 658 (1966).
58. Electron micrographs by Dr. E. F. J. van Bruggen with technical assistance of Carol M. Runner, Laboratory of Structural Chemistry, University of Groningen.
59. V. Bode and A. D. Kaiser, *J. Mol. Biol.* **14**, 399 (1965).
60. E. T. Young and R. L. Sinsheimer, *ibid.* **10**, 562 (1965).
61. H. Ogawa and J.-I. Tomizawa, *ibid.* **23**, 265 (1967).
62. A. Becker, M. Gefter, J. Hurwitz, *Fed. Proc.* **26**, 395 (1967); Abstract of the 51st Annual Meeting, Chicago, Illinois, April 16–21, 1967; J. Hurwitz, M. L. Gefter, A. J. Becker, Ann. Res. Conf., Gatlinburg, Tennessee, April 1967.
63. B. Weiss, T. R. Live, C. C. Richardson, *Fed. Proc.* **26**, 395 (1967).
64. F. R. Frankel, *J. Mol. Biol.* **18**, 109 (1966); *ibid.*, p. 127; A. W. Kozinski and P. B. Kozinski, *Proc. Nat. Acad. Sci. U.S.* **54**, 684 (1965).
65. A. W. Kozinski, P. B. Kozinski, R. James, *Virology*, in press.
66. C. A. Thomas, Jr., *Progr. Nucleic Acid Res.* **5**, 315 (1966).
67. J.-I. Tomizawa, Ann. Res. Conf., Gatlinburg, Tennessee, April 1967.
68. L. A. Salzman and A. Weissbach, *J. Mol. Biol.* **28**, 53 (1967).
69. D. J. McCorquodale and Y. T. Lanni, *J. Mol. Biol.* **10**, 10 (1964); Y. T. Lanni, D. J. McCorquodale, C. M. Wilson, *ibid.*, p. 19.
70. I thank Dr. F. Berends, Prof. H. S. Jansz, Dr. R. A. Oosterbaan, Prof. A. Rörsch, and Dr. A. de Waard for stimulating discussions and criticisms; and Miss Frieda Leenart for preparation of the manuscript.

NEWS AND COMMENT

Water and Air Pollution: Two Reports on Cleanup Efforts

1. The Battle of Lake Erie: Eutrophication and Political Fragmentation

Cleveland. As any nose can detect on a warm day, Lake Erie, the oldest and shallowest of the Great Lakes, is an ailing body of water. Its affliction, known as eutrophication, is a natural disease of aging lakes. But, in the case of Lake Erie, the inevitable is being accelerated by a daily overdose of some 150,000 pounds of phosphates, which, ironically, are a by-product of man's penchant for cleanliness. Since most of the phosphates have been traced to expended detergents, it might seem that the problem could be alleviated in one way or another. But, as has been demonstrated on innumerable occasions, the technology of pollution control cannot be considered apart from the economics and politics of pollution.

Eutrophication is characterized by changing biota, and particularly by dense algal blooms that deplete oxygen from the bottom layers of lake water when the algae decay, foul beaches when they wash ashore, and sometimes cause odor and discoloration of drinking water (*Science*, 13 Oct., p. 278).

Statistics compiled by the Federal Water Pollution Control Administration's (FWPCA) Great Lakes Program in Cleveland indicate that 80 percent of the phosphates entering Lake Erie are first treated by municipal sewer plants. Of those phosphates, 66 percent originate in detergents. The FWPCA takes the position that the eutrophication problem in Erie can be alleviated by removing phosphates from detergents and from treated sewage effluents.

Because no one favors pollution, the soap and detergent industry, which grosses $2 billion annually, has been left in the position of having to favor eutrophication control measures while quietly pointing out that it is not solely to blame. The industry is especially sensitive since just 2 years ago it completed a $150-million changeover to the manufacture of biodegradable detergents. It did this after having been singled out as responsible for turning numerous rivers, lakes, and water supplies into giant bubble baths.

Charles G. Bueltman, technical director of the Soap and Detergent Association, said during an interview with *Science* that the industry is now being cited as the major cause of eutrophication because the government knows the industry will cooperate to find a solution. The industry has already proved its public-spiritedness by voluntarily seeking and finding a substitute for the suds situation, Bueltman says. Critics of the detergent industry assert that the industry came up with a remedy only because legislation was being threatened which would have forced it to do so anyway. Bueltman denies this, saying the industry started a program in 1951 to find a substitute for the surface active agent that was causing the problem. In 1963 it announced that a substitute substance had been found and that an industry-wide changeover would be made. According to Bueltman, most of the legislation threats were not made until after the switch had been announced.

The industry asserts that most of the new burst of finger-pointing in its direction did not begin until 1965, the year of the first Lake Erie Enforcement Conference. Bueltman says that