Bacteriophage T7

Genetic and biochemical analysis of this simple phage gives information about basic genetic processes.

F. William Studier

Bacteriophages are among the simplest biological entities known, and yet they carry out biological processes that are basic to even the most complex organisms. In trying to understand such processes, phage-host systems have one big advantage over more complex systems: they are very easy to manipulate genetically. Mutations can be obtained in virtually any phage function (and in many host functions as well), and biochemical analysis of mutant strains can often reveal the molecular interactions that make up a biological process. Nucleic acid replication, genetic recombination, regulation of gene expression, and assembly of complex structures can all be studied in bacteriophage systems.

Among phages that contain double-stranded DNA, T4 and lambda and their relatives have been favorite objects of study. Phage T7 appears to be even simpler, and is currently receiving considerable attention. A strictly virulent phage, T7 (along with its relative, T3) is the smallest of the seven T phages originally described by Demerec and Fano (*1*) and Delbruck (*2*). It has a polyhedral head, a small, simple tail (*3*), and contains a single piece of double-stranded DNA of molecular weight 25×10^6, about one-fourth the size of the DNA from T even phages (*4*). The DNA from T7 contains the four usual DNA bases (*5*), and its base sequence is not circularly permuted across the population, as it is in the T even phage DNA's (*6*). Mature T7 DNA has a terminal repetition, but, unlike lambda DNA (*7*), the ends are double stranded and do not associate unless first treated with an exonuclease (*6*). After infection, T7 specifies approximately 30 proteins, which account for virtually all of the coding capacity of T7 DNA (*8*). Thus, T7 seems to be a manageable size, and it may well be possible to find a mutation in each of its genes, to determine the function of each of its proteins, and to define the molecular details of the processes directed by T7 after infection.

The author is biophysicist in the biology department, Brookhaven National Laboratory, Upton, New York 11973.

Genetic Analysis of T7

Conventional genetic techniques include the isolation of mutant strains, the analysis of their patterns of recombination and complementation, and the construction of a genetic map. In T7 it is also possible to analyze the end products of gene expression, the protein chains. The combining of physical and genetic techniques has permitted an unusually thorough genetic analysis of T7.

T7 proteins. The proteins of T7 are analyzed by electrophoresis on polyacrylamide gels in the presence of the detergent sodium dodecyl sulfate (SDS), a simple yet powerful technique (*8, 9*). A culture of host cells growing in minimal medium is infected with T7, and ^{14}C-labeled amino acids are added to label the proteins that are synthesized. At the end of the labeling period the infected cells are harvested by centrifugation, suspended in a buffer solution containing SDS, and heated briefly to 100°C. This treatment solubilizes virtually all of the proteins in the cell and dissociates them to individual protein chains, each complexed with large amounts of SDS. Since SDS is negatively charged, the protein-SDS complexes are also negatively charged, and all migrate in the same direction in an electric field. Electrophoresis through a polyacrylamide gel resolves the protein-SDS complexes according to their size (*10*). After electrophoresis, the gels can be dried and autoradiographed to determine the positions of the labeled proteins. Thus, all proteins that incorporate ^{14}C-labeled amino acids during a labeling period are displayed, and their sizes can be estimated from their positions on the gel. Autoradiograms showing the time course of protein synthesis during a normal T7 infection are given in Fig. 1a.

In a normal infection, any host proteins that are being synthesized at the same time as the T7 proteins are also labeled and appear on the autoradiogram. However, if uninfected host cells are irradiated sufficiently with ultraviolet light, they become unable to synthesize their own proteins but retain the ability to synthesize T7 proteins after infection (*8*). In this system, only T7 proteins incorporate added ^{14}C-labeled amino acids, and it is possible to visualize T7 proteins without any interference from host proteins. The time of appearance of T7 proteins seems to be unaffected by the prior irradiation of the host (Fig. 1b).

Approximately 30 T7 proteins can be identified on SDS-polyacrylamide gels. Their molecular weights range between approximately 7,000 and 150,000, and together they account for virtually all of the coding capacity of T7 DNA (*8, 11*).

T7 mutations. Amber mutants and deletion mutants are particularly useful in the genetic analysis of T7. Not only are they easily isolated but both types of mutation alter the size of any protein they affect, thus permitting the protein to be identified electrophoretically.

For most of the genetic analysis of T7, amber mutants have been utilized (*11–14*). Amber mutants make a protein of normal size when grown on a permissive host, but make a shorter fragment, the amber peptide, when grown on a restrictive host (*15*). If the affected protein is essential to the growth of the phage, the mutant will be unable to grow on the restrictive host but will usually be able to grow in the permissive host. Such mutants are referred to as conditional-lethal mutants (*16*). These mutants can be isolated by plating a phage stock that had been treated with a mutagen on the permissive host and then testing individual plaques to find strains that do not grow on the restrictive host. Conditional-lethal amber mutations are easy to map genetically. The mutants are crossed in the permissive host, and the progeny are plated on both the permissive and restrictive hosts. All of the progeny

grow on the permissive host, but only wild-type recombinants can grow on the restrictive host; thus, the ratio of plaques found on the two hosts is readily converted to percent recombination.

Deletion mutants, or strains of T7 in which a portion of the DNA has been deleted, are also easy to isolate. Parkinson and Huskey (*17*) found that phage particles containing a piece of DNA shorter than normal length are more resistant to disruption at high temperatures (perhaps because the smaller amount of DNA exerts less pressure on the head structure). This makes it possible to select deletion mutants at random by heating a solution of T7 particles to the point where most are inactivated, and then growing the survivors. Many of the heat-stable strains isolated from T7 in this way are deletion mutants (*18, 19*). Conventional genetic mapping of deletions is difficult unless the deletion causes a change in plaque morphology or some easily measured character. However, the location and extent of a deletion can be determined in the DNA molecule itself. This physical mapping is done by hybridizing DNA molecules and looking at them in the electron microscope (*20*). When a wild-type strand of DNA is hybridized to a complementary strand from a molecule containing a deletion, the unmatched region of the wild-type strand will loop out, marking the position and length of the deletion.

Deletion mutants are viable only if they delete nonessential regions of the DNA, that is, nonessential genes, nonessential portions of a gene, or nonessential extragenic DNA. (In this article the term gene refers to that portion of a DNA molecule which specifies a single protein chain.) Thus, deletion mutations are complementary to conditional-lethal amber mutations in the sense that deletions are found only in nonessential genes and conditional lethals only in essential genes. The combination should, theoretically, yield a mutation in every T7 gene.

The protein affected by an amber mutation or a deletion should disappear from its normal position in the patterns obtained upon electrophoresis of extracts of infected cells on SDS-polyacrylamide gels. This is because the protein chains are resolved on the basis of size, and both types of mutation affect the size of the protein synthesized. The size of the amber peptide will depend on the location of the amber mutation, and this property has been used to determine the direction of translation (and thus

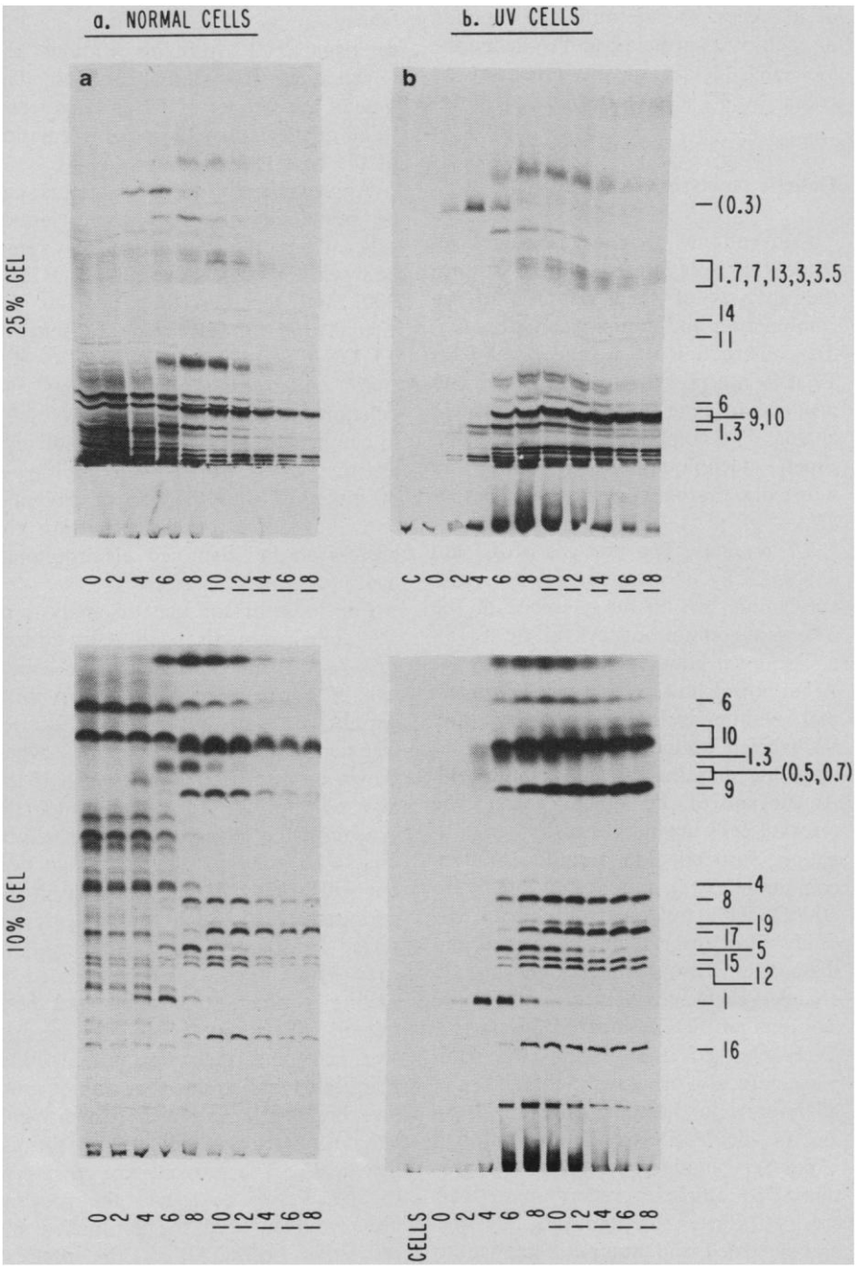

Fig. 1. Time course of protein synthesis in normal (a) and ultraviolet-irradiated (b) *E. coli* B. Cells were grown and infected in M9 minimal medium at 30°C. At 2-minute intervals, samples of infected culture were added to a mixture of ^{14}C-labeled amino acids. Cultures that had been irradiated with ultraviolet light received ten times as much ^{14}C as those that had not been irradiated. After 2 minutes in the presence of ^{14}C, a great excess of unlabeled amino acids was added in order to greatly reduce further incorporation of ^{14}C into proteins. The samples were incubated another 4 minutes to allow any ^{14}C-containing protein chains to be completed, chilled to 0°C, and centrifuged. The cells were suspended in 0.05*M* tris-Cl. *p*H 6.8, containing 1 percent SDS, 1 percent mercaptoethanol, and 10 percent glycerol. After being heated for 2 minutes in a boiling water bath, the samples were subjected to electrophoresis in a discontinuous, SDS-containing buffer system (*47*) on slabs of 10 and 25 percent polyacrylamide gel [a modification of the system of Reid and Bieleski (*52*)]. The origin of electrophoresis is at the bottom, and the farther a protein moves the smaller its molecular weight. The time after infection at which the pulse of ^{14}C began is given beneath each pattern. The number of the gene which specifies each band (where known) is given at the right of the patterns. The numbering of genes 0.3, 0.5, and 0.7 is tentative, since it has not been definitely established whether these protein bands are the product of two genes or three genes nor what the gene order is (see text).

transcription) relative to the genetic map (*8*). The size of any new protein chain that might appear in a deletion mutant will depend on whether the deletion eliminates all or a part of a gene or fuses two genes. Presumably, all of the proteins coded for by T7 are displayed in the protein patterns; therefore, it should be possible to know when a mutation has been found in every T7 gene.

T7 genes. Amber mutants have been isolated in 19 essential genes of T7 (*12–14*). These mutations fall on a linear genetic map, and the 19 genes are numbered in order from left to right. More than 800 conditional-lethal amber mutants have been characterized, but no new gene has been found since mutant number 107 (*11*); either all of the essential genes have been identified, or amber mutations in the remaining essential genes are relatively rare. The proteins specified by 17 of these 19 genes have been identified in the protein patterns on SDS-polyacrylamide gels (see Fig. 2) (*8, 11*).

Amber mutations can also be found in nonessential genes if a test for gene function is available, for example, an assay for enzyme activity. In this case, strains are isolated randomly from a heavily mutagenized stock and are tested for the enzyme activity. Since T7 has only 30 or so genes, if a stock were mutagenized to an average of several mutations per phage particle, only a moderate number of strains should have to be tested before the desired mutant is obtained. A strain containing an amber mutation in the gene for T7 ligase was isolated in this way by Masamune, Frenkel, and Richardson (*21*). It was found among the first 20 strains tested.

T7 ligase mutants grow in normal *Escherichia coli* hosts, but not in a ligase-deficient host (*21, 22*). Thus it is possible to map ligase mutants relative to conditional-lethal amber mutations by using a restrictive host which is also ligase deficient. The ligase gene falls between genes 1 and 2 (*11*). An advantage of the genetic map of T7 is that it makes it possible to recognize the position of a gene by its number. To preserve this feature, new genes are given a decimal number which indicates their relative position. Thus, ligase is designated gene 1.3, since it falls between genes 1 and 2, and, as discussed below, at least one other gene is located between it and gene 2. An amber mutation in lysozyme has also been found by enzymatic test (*11*), and it maps between genes 3 and 4; thus, the lysozyme gene is designated 3.5. Both the ligase and lysozyme proteins have been identi-

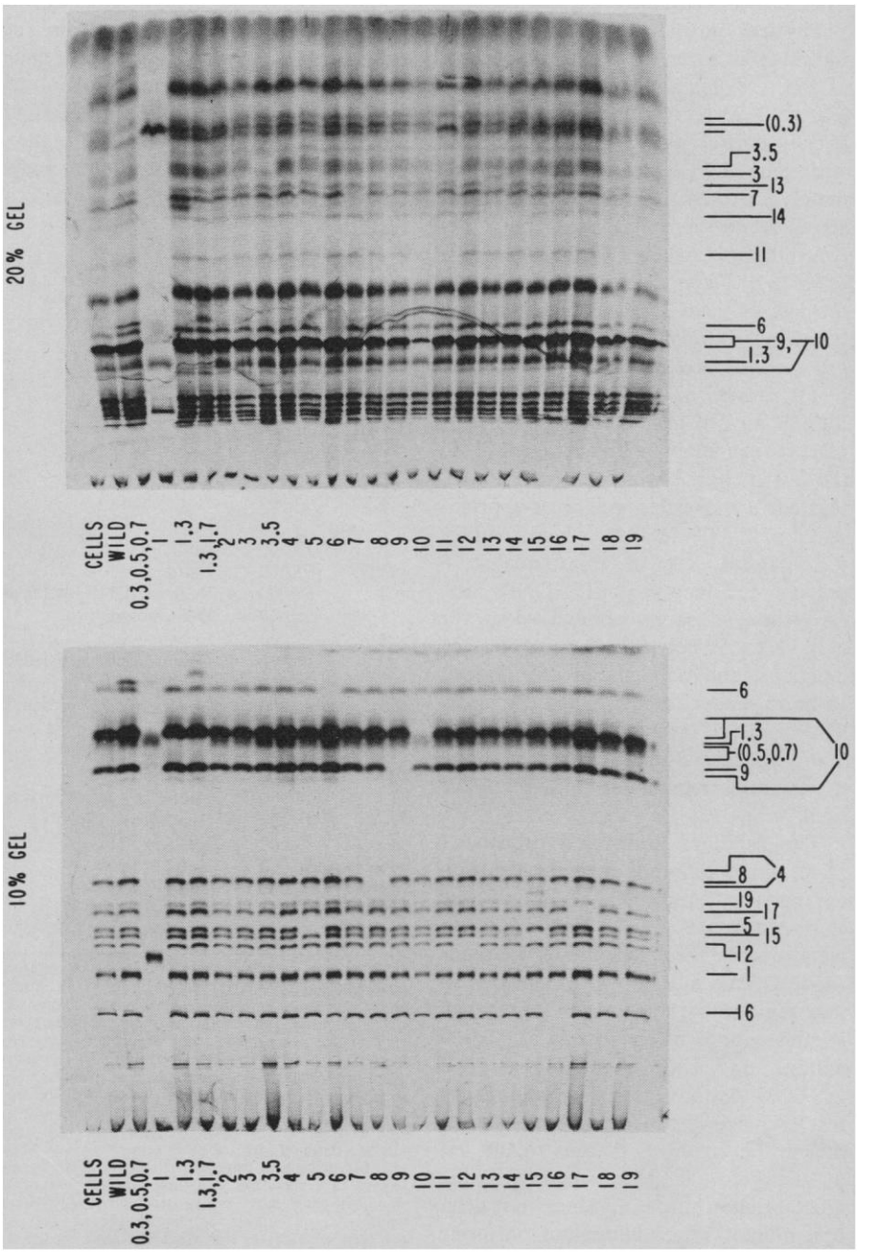

Fig. 2. Protein patterns of different T7 amber and deletion mutants. Patterns were obtained essentially as described in the legend to Fig. 1, with the following exceptions: the upper patterns were obtained on a 20 percent gel instead of a 25 percent gel; twice as much sample was analyzed on the 20 percent gel as on the 10 percent gel, in order to increase the relative intensity of the fainter bands; and ^{14}C-labeled amino acids were present from 2 to 20 minutes after infection so that all T7 proteins would be labeled. The host cells were irradiated with ultraviolet light before infection; therefore, only T7 proteins appear. The patterns are oriented as in Fig. 1. The number or numbers beneath each pattern identifies the gene or genes in which each strain has mutated. Those patterns labeled with a single gene number are amber mutants; those with more than one are deletions which cover the numbered genes. The numbers to the right of the patterns identify the genes which specify individual proteins. (A slight distortion has shifted the outside patterns down and away from the center patterns; the lines on the right point to the bands in the center patterns.) Except for the pattern given by the mutant defective in gene 1, each pattern lacks only the protein band specified by the gene itself. Amber peptides can be seen on the 10 percent gel in the patterns from gene 1 and gene 15, and deletion peptides can be seen in the patterns from genes 0.3, 0.5, 0.7 and from genes 1.3, 1.7. Possible amber peptides are more difficult to distinguish on the 20 percent gel, but may be present in the patterns for genes 1.3, 11, 13, and 16.

fied in the protein patterns on SDS-polyacrylamide gels (*11*).

Deletions have been found in two different regions of the T7 DNA molecule (*19*). Since the T7 ligase is not essential in the normal hosts, it has been possible to isolate deletions which include at least part of the ligase gene. Ligase has been mapped genetically, and thus ligase deletions provide a link between the genetic and physical maps, identifying the left end of the T7 DNA molecule. Genetic mapping shows that some ligase deletions come very close to the right end of gene 1, and these seem to lack only the ligase protein; others come very close to gene 2, and these lack at least one additional protein, which indicates that at least one gene lies between ligase and gene 2. This gene is designated 1.7.

Physical position on the T7 DNA molecule is given in terms of percent of the total length, measured from the left end. Deletions in the region between genes 1 and 2 (including ligase and gene 1.7) extend from approximately 15 to 24 percent (*19*). A second series of deletions has been found between approximately 3 and 8 percent (*19, 23*). These deletions must fall to the left of gene 1, as indicated by the following combination of facts: (i) the protein specified by gene 1 is essential to the growth of T7, which means that little, if any, of gene 1 could be deleted; (ii) genetic mapping places gene 1 to the left of ligase; (iii) the gene 1 protein has a molecular weight of approximately 100,000 (*8, 24*), which requires a coding capacity of approximately 8 percent the length of T7 DNA; and (iv) there is barely enough room for gene 1 to fit between the two sets of deletions, and not enough room for it to be to the left of the 3 percent mark. Thus, gene 1 must be located between approximately 8 and 15 percent, and nonessential regions lie to either side of it.

The deletions to the left of gene 1 affect three different protein bands in the polyacrylamide gel, an intense one at approximately 9,000 daltons and a weak double band near 40,000 daltons (*19*). Of the mutants with deletions in this region that have been tested, all eliminate both of the bands at 40,000 daltons, and some eliminate the band at 9,000 daltons as well. Preliminary results place the gene which specifies the protein at 9,000 daltons to the left of the gene or genes specifying the 40,000-dalton protein. Since no deletion mutant which eliminates only one of the bands at 40,000 daltons has been found, it is not clear whether these bands are the product of one gene or two. Provisionally, the gene for the band at 9,000 daltons is numbered 0.3 and those for the two bands at 40,000 daltons 0.5 and 0.7, but this numbering may have to be revised after more is known about this region.

Protein patterns of mutants in the 25 genes of T7 for which mutants are available are given in Fig. 2. These patterns were obtained in cells that had been irradiated with ultraviolet light before infection, so that only T7 proteins would be displayed, and ^{14}C-labeled amino acids were present throughout the infective cycle so that all T7 proteins would be labeled. All mutants are either ambers or deletions, so the proteins directly affected by the mutation should be missing from the pattern. Two patterns are given for each mutant, one on a 10 percent gel under conditions which resolve proteins of molecular weight 25,000 and higher, and a second on a 20 percent gel under conditions which resolve proteins of molecular weight less than 25,000.

Table 1. Approximate molecular weights (M.W.) and functions of the proteins specified by T7 genes.

Gene	~ M.W.*	Function †
0.3 ‡	8,700	Nonessential
0.5	40,000	Nonessential
0.7	42,000	Nonessential
1	100,000	RNA polymerase
1.3	40,000	Ligase
1.7	17,000	Nonessential
2		Reduced DNA synthesis
3	13,500	Endonuclease
3.5	13,000	Lysozyme
4	67,000	Reduced DNA synthesis
5	81,000	DNA polymerase
6	31,000	Exonuclease
7	14,700	Found in phage particle
8	62,000	Head protein
9	40,000	Head assembly protein
10	38,000	Major head protein
11	21,000	Tail protein
12	86,000	Tail protein
13	14,000	Found in phage particle
14	18,000	Head protein
15	83,000	Head protein
16	150,000	Head protein
17	76,000	Tail protein
18		DNA maturation
19	73,000	DNA maturation

* Molecular weights are estimated from relative positions on SDS-polyacrylamide gels (*11*). Some proteins change relative position with change in electrophoresis buffer or in concentration of acrylamide, but most seem to retain approximately the same relative position. The values given, particularly at low molecular weights, are rather crude approximations, which will be refined as better marker proteins become available and as molecular weights of specific T7 proteins are determined by absolute methods. † See text for discussion and references. ‡ The numbering of genes 0.3, 0.5, and 0.7 is tentative, since it has not been definitely established whether these protein bands are the product of two genes or three, nor what the gene order is.

All patterns except that from the gene 1 mutant are substantially the same as wild type, except that they lack the protein(s) directly affected by the mutation. A mutation in gene 1 affects most of the proteins specified by T7; this behavior will be discussed in the next section. In some patterns, amber peptides or deletion peptides can be found, but most of the amber mutations used are near the left end of the gene and the amber peptides would be very small. Mutations have been identified for most of the bigger proteins of T7, but several bands appear on the 20 percent gel for which no mutation is yet available. It may be possible to obtain amber or deletion mutants affecting these proteins by analyzing protein patterns of strains isolated randomly from a heavily mutagenized stock, in much the same way strains defective in ligase were found. Gene numbers and approximate molecular weights of the proteins affected by T7 mutants currently available are given in the first two columns of Table 1. The proteins that have been identified with specific genes account for approximately 90 percent of the coding capacity of T7 DNA.

Regulation of Gene Expression by T7

In a normal infection, T7 proteins are synthesized according to a characteristic time course (Fig. 1), and at least three classes of proteins can be distinguished (*8, 11*): class I, those synthesized between approximately 4 and 8 minutes after infection; class II, those synthesized between approximately 6 and 15 minutes after infection; and class III, those synthesized from approximately 6 to 8 minutes after infection until lysis. The class I proteins are the same ones that are synthesized at normal rates in gene 1 mutants (Fig. 2). As far as can be determined, these three classes are strictly correlated with position on the genetic map (*11*): class I includes those proteins specified by genes 0.3 to 1.3, the left-most genes of T7; class II includes the next group of proteins, specified by genes 1.7 to 6; and class III includes the right arm of the genetic map, the proteins specified by genes 7 to 19. (Gene 1.3, ligase, is assigned to class I because it is synthesized at normal rates in gene 1 mutants; however, the ligase protein usually continues to be synthesized until class II

protein synthesis ends.) Thus, the T7 genes seem to be expressed according to their map position, left being early and right being late. The orderly appearance of T7 proteins implies some type of control over gene expression, which could be mediated at the level of transcription or translation or both.

All transcription from T7 DNA after infection proceeds from left to right (*8, 25*), the same direction as the early to late polarity in T7. Summers (*26*) has identified 12 or 13 species of T7 messenger RNA in infected cells, which together can account for the total capacity of T7 DNA. If these represent nonoverlapping natural messengers, then a number of them must contain information for more than one protein chain. However, the polar effects on protein synthesis that might be expected from amber mutations in a polygenic messenger RNA (*27*) are not very pronounced, if present at all (Fig. 2). The time course of T7 RNA synthesis has not yet been studied in sufficient detail to determine whether three classes of messenger RNA molecules, corresponding to the three classes of proteins, can be distinguished. However, as discussed in the next section, Siegel and Summers (*28*) and Summers (*29*) have identified the class I messenger RNA's. Summers (*30*) has also found that T7 messenger RNA synthesized between 6 and 8 minutes after infection seems to be stable, unlike the messenger RNA of uninfected cells which has a half-life of 2 to 3 minutes. This suggests that the messenger RNA's for class III proteins, and perhaps other T7 messenger RNA's as well, are not degraded.

Gene 1 action. As was mentioned in the previous section, gene 1 mutants are the only ones isolated so far that seem to affect the regulation of gene expression after T7 infection (*8, 11*). When gene 1 is inactive, all class I proteins are synthesized at normal rates, but synthesis of class II and class III proteins is greatly depressed. By analyzing the messenger RNA produced after infection with normal T7, using gene 1 mutants or the presence of chloramphenicol (which prevents the synthesis of any T7 proteins, including the gene 1 protein), Siegel and Summers (*28*) have shown that gene 1 action in vivo is at the level of transcription. They identified four or five messenger RNA molecules, ranging in size from 200,000 to 1,000,000 daltons, which must be the messenger RNA's for the class I proteins.

Chamberlin *et al.* (*24*) have shown that the gene 1 product is a new RNA polymerase. They have purified it to near homogeneity and shown that it differs from host RNA polymerase in several respects, including size, template preference, conditions for optimal activity, and sensitivity to inhibitors. Thus, it seems likely that the host RNA polymerase transcribes only a limited region of the T7 DNA in vivo (the class I genes), and that the gene 1 RNA polymerase is necessary for efficient transcription of the class II and class III genes. Summers and Siegel (*31*) have shown that partially purified T7 RNA polymerase does transcribe the late regions of T7 DNA in vitro.

Transcription by host RNA polymerase. When gene 1 does not function in vivo, all transcription is presumably by the host RNA polymerase. Apparently the host RNA polymerase can transcribe only the left-most genes of T7 and cannot proceed past a point in the T7 DNA just to the right of gene 1.3 (ligase).

Experiments in vitro by Davis and Hyman (*32*) indicate that T7 DNA has only a single starting point for transcription by host RNA polymerase. They used the electron microscope to look at complexes of T7 DNA and purified *E. coli* RNA polymerase which had been allowed to synthesize RNA for varying lengths of time, and concluded that RNA synthesis starts near one end of the T7 DNA molecule. Examining hybrid molecules made by annealing newly synthesized RNA to T7 DNA, they found a single initiation point approximately 1.3 percent from the left end of the T7 DNA, which would be to the left of the class I genes. Hyman (*33*) has examined hybrids between isolated T7 DNA and class I RNA made in vivo. He found that the hybrid region began approximately 1 percent from the left end, in good agreement with the results in vitro.

A specific signal in the T7 DNA itself stops transcription by host RNA polymerase at a point just to the right of ligase, and deletions which extend through the right end of the ligase gene seem to remove the stop signal (*11, 19*). This can be seen in double mutants containing a mutation in gene 1 plus a ligase deletion; the time course of protein synthesis after infection by two such double mutants is shown in Fig. 3. In the first double mutant (1, LG3), class II and III proteins are made in the small amounts characteristic of any gene 1 mutant, but in the second (1, LG37) they appear in much greater amounts, which suggests that the stop signal has been deleted. The amounts of class II and III proteins are not as great as the amounts found during infection with wild-type T7, but one might expect the T7 RNA polymerase to be more active than host RNA polymerase in transcribing these genes.

The physical position of the stop signal in vivo in the T7 DNA can be determined from the positions of the right-hand ends of ligase deletions which do and do not delete the stop signal. Only a few such deletions have been analyzed so far (*19*) and results indicate that the stop signal is located between approximately 20 and 20.8 percent from the left end of T7 DNA. Hyman's mapping of the class I RNA made in vivo (*33*) places the stop signal approximately 20 percent from the left end, in excellent agreement with the data from the deletion mutants. The four or five class I proteins could account for as much as 18.5 percent of the coding capacity of T7 DNA, which would leave little room for anything else to the left of the stop signal.

Experiments in vitro have given conflicting results on what is required for termination of transcription at the end of the class I region. Using isolated T7 DNA and purified host RNA polymerase, Davis and Hyman (*32*) did not find specific termination unless rho factor (*34*) was added to the reaction. Without rho (a termination factor) transcription proceeded all the way to the right end of the T7 DNA; in the presence of rho, most transcription did not proceed past a point approximately 18 percent from the left end, and the RNA synthesized was not uniform in size. These authors suggested that termination at the right end of the class I region requires rho and that the host RNA polymerase is prevented from moving past this point. They further suggested that there are termination sites within the first 18 percent of the T7 DNA molecule at which rho causes the release of RNA chains but does not prevent the RNA polymerase from proceeding.

On the other hand, the results of Millette *et al.* (*35*) suggest that rho is not needed for termination at the right end of the class I region. Again using purified host RNA polymerase, these workers found that, in the absence of rho, RNA molecules of approximately 2.2×10^6 daltons are produced. Almost all of this RNA has uracil as the 3′ terminal base, which suggests that all

of these molecules may have terminated at a unique point. An RNA molecule this size would represent almost 18 percent of the length of T7 DNA, approximately the length of the class I region.

These two sets of experiments do not conflict on the possibility that rho-mediated termination events may occur in the first 18 percent of the T7 DNA, but they disagree on what is required to prevent transcription past the stop signal at the end of the class I region. It is possible that unrecognized factors play a role. It should also be pointed out that termination at the end of the class I region is not completely effective in vivo (see below).

After infection with mutants which lack gene 1 activity and in which the stop signal is deleted, the order of appearance of the T7 proteins is the same as their order on the genetic map (Fig. 3). Once the host RNA polymerase passes the stop signal, it apparently transcribes the genes in order all the way to the right end, which indicates that the rest of the T7 DNA is free of stop signals for host RNA polymerase. Thus, the time of appearance of a protein in such mutants provides a crude measure of its position on the DNA molecule. The last protein appears approximately 13 minutes after infection; therefore, the minimum rate of transcription of T7 DNA (25×10^6 daltons) by *E. coli* RNA polymerase in vivo at 30°C is approximately 10^6 daltons of RNA per minute, or 50 nucleotides per second. This is almost twice as fast as previous estimates for the rate of transcription in uninfected *E. coli* (*36*).

The stop signal to the right of ligase is apparently not 100 percent effective in vivo. Small amounts of class II and III proteins are made after infection with gene 1 mutants even when the stop signal is present (Fig. 3). This indicates that there is some transcription beyond the stop signal. Such transcription could be caused by a small amount of active T7 RNA polymerase or by some host RNA polymerase molecules which transcribe past the stop signal. The latter possibility is probably correct for two reasons: (i) double mutants in gene 1 produce the same small amounts of class II and III proteins (*11*), and (ii) the time course of appearance of the class II and III proteins in gene 1 mutants is the same whether or not the stop signal is present (only the amounts are different), and it is very different from the time course observed in a wild-type infection (see Figs. 1 and 3).

On the whole, experiments in vivo and in vitro agree quite well, and their combined results give the following picture of transcription after T7 infection. The DNA of T7 is first transcribed by the host RNA polymerase, which starts at a single point close to the left end of T7 DNA and stops at a point approximately 20 percent from the left end. The messenger RNA is found in four or five pieces, which may arise from specific termination events, perhaps mediated by rho or some other termination factor. The class I proteins are made, and they account for virtually all of the coding capacity of this region. One of these proteins is the gene 1 RNA polymerase, which in turn transcribes the rest of the T7 DNA. The stop signal at approximately 20 percent from the left end is not completely effective, a certain fraction of the host RNA polymerase molecules reading through all the way to the right end.

Other regulation. The interpretation just described explains the dramatic effect of gene 1 action, but leaves unanswered a number of questions concerning other aspects of gene expression after T7 infection. Why does the synthesis of host proteins and some class I proteins stop approximately 8 minutes after infection? It is tempting to speculate that inactivation of host RNA polymerase could be responsible for both. [If this were the case, the ligase gene (1.3) should be transcribed by the gene 1 RNA polymerase as well as by the host RNA polymerase, since ligase continues to be synthesized until 15 minutes after infection, shutting off with the class II proteins.] Gene 1 mutants are the only ones found so far that are defective in shutting off the synthesis of host proteins and class I proteins (*11*). The gene 1 RNA polymerase could actively participate in this process, but it seems more likely that a class II or III gene is responsible and the effect of the gene 1 mutation is indirect. However, none of the class II or III genes identified so far seems to be responsible (*11*).

A similar unanswered question is: why does the synthesis of class II proteins stop midway through the infective cycle? In some experiments, including that shown in Fig. 1, the synthesis of class III proteins slows at the same time. No mutant has been found in which this process does not occur, but the shutoff seems to be much less pronounced in cells irradiated with ultraviolet light (Fig. 1). It is not known whether this process is regulated at the level of transcription or translation.

Another element in the control of gene expression concerns the mechanism for determining the relative rates of synthesis of different proteins. Much of the difference in rates among class II and III proteins disappears when transcription is directed by host RNA polymerase (compare Figs. 1 and 3). In fact, when transcription is by the host RNA polymerase there seems to be a gradient in the rate of synthesis of class II and III proteins, the rate generally decreasing from left to right (Fig. 3). This suggests that a major determinant of the rate of protein synthesis may be the number of copies of messenger RNA present, and Summers (*26*) did observe widely varying amounts of different T7 messenger RNA's in a normal infection. Specific messenger RNA's for individual genes have not been identified, so it remains to be determined whether a correlation exists between the number of copies of a messenger RNA and the rate of protein synthesis. If the relative rate of protein synthesis does depend on the relative quantities of different messenger RNA's, then the question becomes: how is the rate of synthesis of different messenger RNA's regulated? This presumably depends on the number and location of specific initiation and termination points for T7 RNA polymerase, whether the enzyme is more active at some sites than others, and whether other factors which interact with the enzyme or the DNA might be involved. Again, this information is not yet available. There may also be factors which regulate rates of protein synthesis at the level of translation, and Morrison and Malamy (*37*) have suggested that the bacterial sex factor interferes with T7 gene expression at the translational level.

Functions of T7 Genes

Genes with related functions tend to cluster along the genetic map in T7 (*13*), as in lambda (*38*) and T4 (*16*). The functions of genes 0.3 to 0.7 are not known, but they are not essential for growth in normal hosts. Gene 1 specifies the T7 RNA polymerase, as was discussed before. Gene 1.3 specifies the T7 ligase, which presumably has a role in DNA metabolism after infection,

although the host ligase can substitute for it. The function of gene 1.7 is unknown, but it is not essential for growth in normal hosts. All of genes 2 to 6 affect the kinetics of DNA synthesis (*12–14*). Genes 7 to 17 specify proteins that are found in mature T7 particles or in partly assembled particles (*8, 11, 13*). Genes 18 and 19 seem to be involved in the maturation of T7 DNA from the replicating intermediate (*14, 39, 40*). Thus, early functions are to the left and late functions to the right: genes located between 1.3 and 6 seem to be mostly if not entirely involved in the early steps of DNA metabolism; genes 7 to 19 seem to participate in phage assembly, including the late stages of DNA metabolism. Known functions of T7 genes are summarized in Table 1.

DNA metabolism. The breakdown of host DNA and the synthesis of T7 DNA are among the early steps of DNA metabolism after infection by T7. The kinetics of DNA synthesis after infection (at 30°C) suggest the following (*12–14*): synthesis of host DNA con-

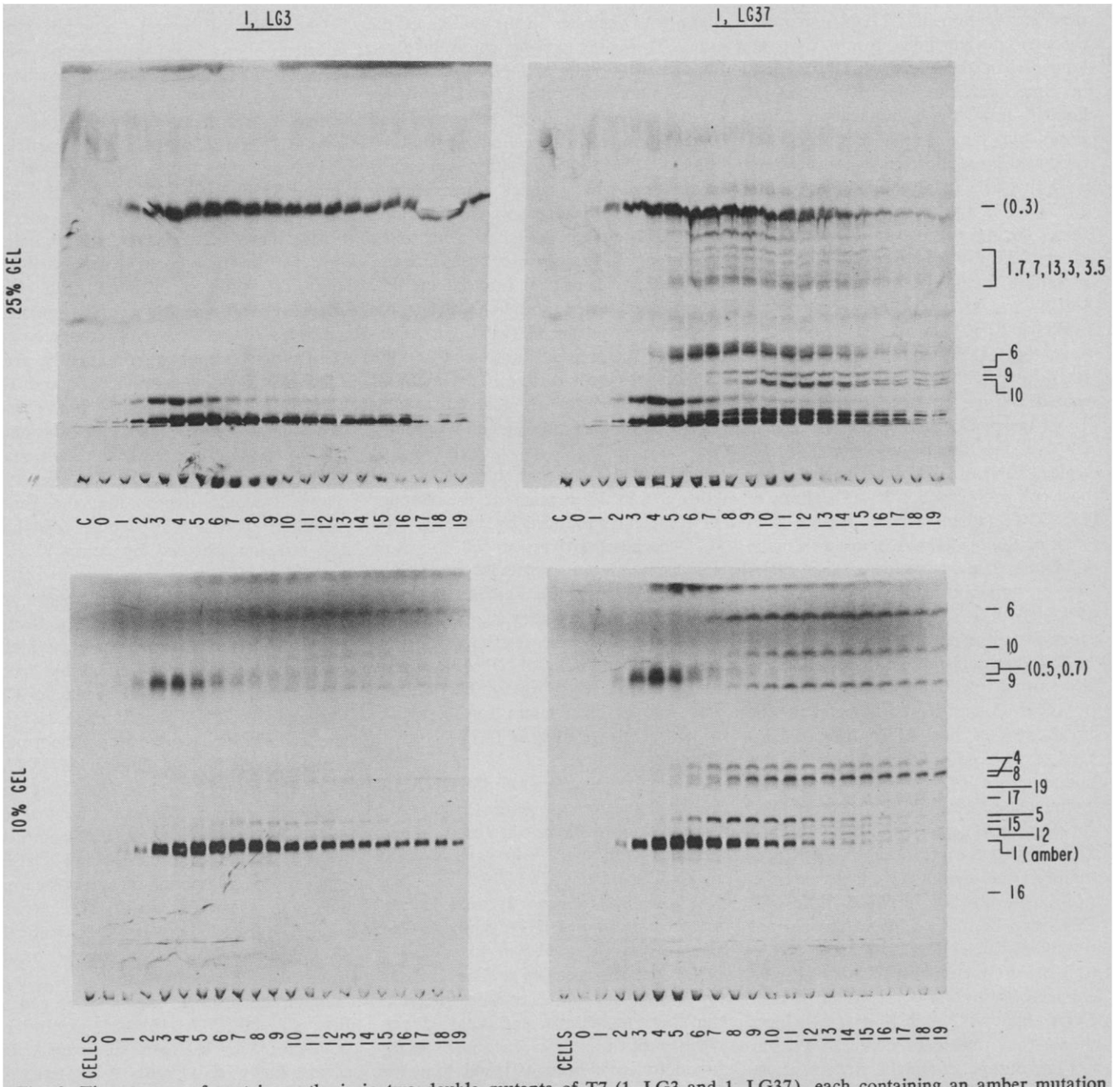

Fig. 3. Time course of protein synthesis in two double mutants of T7 (1, LG3 and 1, LG37), each containing an amber mutation near the right end of gene 1 plus a deletion in the ligase region. Patterns were obtained essentially as described in the legend to Fig. 1, except that the cultures were exposed to ^{14}C-labeled amino acids for 1 minute instead of 2 minutes and the autoradiograms were exposed almost three times as long. The host cells were irradiated with ultraviolet light before infection; therefore, only T7 proteins appear. The patterns are oriented as in Fig. 1, and the numbers beneath the patterns give the time after infection at which the pulse of ^{14}C began. The numbers to the right identify the genes which specify individual proteins; the amber peptide from gene 1 is also identified. The ligase protein (gene 1.3) is missing from the patterns because of the deletions. The right end of the deletion in the double mutant 1, LG3 is approximately 20 percent from the left end of T7 DNA; the right end of the deletion in 1, LG37 is approximately 22 percent from the left end of T7 DNA (*19*). The protein patterns indicate that the LG3 deletion apparently does not delete the stop signal for host RNA polymerase whereas the LG37 deletion apparently does (see text).

tinues at the preinfection rate for approximately 5 minutes after infection and then stops between 5 and 10 minutes after infection; synthesis of T7 DNA begins between 5 and 10 minutes after infection, reaches its maximum rate (perhaps 5 to 10 times the preinfection rate) between 15 and 20 minutes after infection, and shuts off just before or along with lysis, which begins between 25 and 30 minutes after infection. Replication of T7 DNA proceeds through intermediate forms that are longer than the piece of DNA found in phage particles (*39, 40*). It is not known how these replicative forms arise, but they could presumably be produced by replication, recombination, or both (*40*). It is thought that mature T7 DNA is cut from the intermediate forms during phage assembly in a process which also generates the terminal repetition found in mature T7 DNA (*40*).

Phage T7 is very efficient at breaking down host DNA to utilize nucleotides for its own DNA. Eighty-five to 90 percent of the nucleotides found in mature T7 phage particles were present in host DNA at the time of infection (*41*). A typical burst of 250 phage particles per host cell produces 6.3×10^9 daltons of T7 DNA, which would require that most of the host DNA be used to make T7 DNA. The nucleotides of host DNA are presumably incorporated into T7 DNA about as fast as they are released, since almost none of the host DNA can be found as acid-soluble material during the normal infective cycle (*12, 42, 43*).

Genes 3 and 6 participate in the breakdown of host DNA after T7 infection (*42, 43*). Gene 3 specifies an endonuclease (*42–44*) and gene 6 an exonuclease (*45*), and both are required for any significant degradation of host DNA. Synthesis of T7 DNA stops prematurely in mutants of genes 3 or 6 (*13*), presumably because little T7 DNA can be made without the normal supply of nucleotides from host DNA.

A third function of T7 may also be connected with the degradation of host DNA. Sadowski and Kerr (*43*) have shown that release of host DNA from its "membrane complex" occurs at an early stage of degradation. Such release occurs in mutants defective in genes 3 or 6 but not in gene 1 mutants (which are deficient in all class II and III proteins). This suggests that there may be a class II protein which causes release, perhaps to facilitate breakdown by the gene 3 and 6 nucleases. The release function cannot be ascribed to the products of genes 2, 3, 4, 5, or 6 (*43*). It is not yet known whether the products of genes 1.7 or 3.5, the other class II genes for which mutants are available, could be responsible for release, but T7 lysozyme, the product of gene 3.5, seems a possible candidate.

There are several indications that T7 lysozyme may have a role in DNA metabolism rather than lysis (*11*). The lysozyme gene lies in the middle of the class II region, among genes involved in the early steps of DNA metabolism. In common with other class II proteins, lysozyme is synthesized between 5 and 15 minutes after infection, an unusual time course for a function concerned with lysis, which begins approximately 25 minutes after infection. Furthermore, the T7 lysozyme does not seem to be required for lysis, since cultures infected with an amber mutant in lysozyme lyse completely, even though no detectable lysozyme is produced (less than 2 percent of the amount produced by the wild type). Lysis is slightly delayed, but this is a property of most mutants in class II genes (*13*). Gene 1 mutants, on the other hand, do not lyse the host even though they can ultimately produce 10 to 20 percent of the amount of lysozyme produced by the wild-type T7 (because a fraction of the host RNA polymerase molecules transcribes past the normal stop signal). The only lysozyme amber mutant available so far grows poorly in the restrictive host; it has a reduced rate of DNA synthesis and produces few progeny. This mutant has not yet been tested for its ability to cause release of host DNA from its "membrane complex."

Normal replication of T7 DNA requires the action of genes 2, 4, and 5 in addition to those discussed above. Mutants with defects in genes 4 or 5 make little if any T7 DNA (*12–14*), and a defect in gene 2 leads to premature cessation of DNA synthesis (*13*). Gene 5 has been shown to specify a DNA polymerase (*46*). The lack of replication in gene 5 mutants suggests that this enzyme may participate directly in replication. So far, no specific functions have been assigned to genes 2 and 4.

Assembly. Functions of class III proteins are primarily, if not exclusively, connected with assembly of the T7 phage particle (*8, 11*). At least 11 different protein chains can be resolved when purified T7 particles are subjected to electrophoresis on SDS-polyacrylamide gels (*11*). These proteins have molecular weights of 13,000 to 150,000 and account for almost half of the coding capacity of T7 DNA. Genes 7, 8, and 10 to 17 specify these 11 proteins, and gene 9 specifies a protein that is found in empty head structures but not in any DNA-containing structures analyzed so far. Thus, genes 7 to 17 all specify proteins found in phage structures.

Analysis of protein compositions, electron microscopy, and complementation of partly assembled particles have been used to determine which proteins make up the head and which the tail of the T7 particle (*8, 11, 13*). Particles containing DNA are produced when mutants in genes 7, 11 to 13, and 17 are grown under restrictive conditions; therefore, none of the proteins specified by these genes is necessary for the production of a filled head. Lysates of wild-type T7 contain unequal amounts of two types of empty head. The minor component contains the proteins specified by genes 8, 10, and 14 to 16, the same proteins which are necessary to make filled heads; the major component contains an additional protein, the one specified by gene 9. Thus, it would appear that the head structure of the T7 particle contains the proteins specified by genes 8, 10, and 14 to 16.

Are the empty heads precursors to filled heads or a by-product? It seems likely that the major species, which contains the protein specified by gene 9, is a precursor to filled heads. In T7 [but not in T4 (*16*) or lambda (*38*)], late proteins are synthesized in normal amounts whether or not phage DNA is synthesized (*11*). A mutant with a defect in gene 5, for example, makes no T7 DNA but makes normal amounts of class II and III proteins; and it also makes a large amount of empty heads. The empty heads made in the absence of T7 DNA synthesis appear identical to the major species of empty head found in wild-type lysates; that is, they contain the proteins specified by genes 8 to 10 and 14 to 16. Virtually none of them lacks the gene 9 protein. It seems likely that gene 9 protein is somehow lost during the filling of the heads, and the empty heads which lack the gene 9 protein may be the by-products of abortive attempts to fill the heads.

Gene 10 specifies the major subunit of the head, a protein having a molecular weight of approximately 38,000

which accounts for more than 60 percent of the mass of the phage particle (*8*). However, mutants with defects in gene 10 also fail to synthesize at least one additional protein, which has a molecular weight of approximately 45,000 (Fig. 2). This protein is also found in the head, but in an amount equal to approximately 10 percent that of the major subunit (*11*). The two gene 10 proteins may be related by a cleavage step, as is found for certain proteins in T4 head assembly (*47, 48*), but such a relationship has not been shown directly.

The tail of T7 contains at least three proteins, those specified by genes 11, 12, and 17 (*8, 11*). Noninfectious, DNA-containing particles from mutants in genes 11 or 12 lack all three proteins, and also lack any visible tail structure when examined in the electron microscope. Particles from mutants defective in gene 17, on the other hand, contain the proteins specified by genes 11 and 12 but not the protein specified by gene 17; and they possess a tail structure which looks similar, but not identical to that of normal T7. These particles can all be complemented in vitro to produce infectious T7 particles (*13*), by procedures similar to those introduced by Edgar and Wood (*49*) in the study of T4 assembly.

It is not clear where the proteins specified by genes 7 and 13 fit into the structure of the T7 phage particle. The DNA-containing particles produced by mutants with defects in genes 11, 12, and 17 have not been carefully tested for the presence of the proteins specified by genes 7 and 13, so it is not known whether one or both may form a part of the tail structure. However, the DNA-containing particles produced by mutants defective in genes 7 or 13 seem to lack only the protein specified by gene 7 or 13, and they do have a tail structure in electron micrographs (*8, 11*). Attempts to complement the gene 7 or 13 particles in vitro have been unsuccessful (*13*). Apparently, the T7 particle can be assembled without the proteins specified by genes 7 and 13, but these proteins cannot be added after assembly is complete.

As was mentioned previously, replication of T7 DNA proceeds through intermediate forms that are subsequently processed to give mature T7 DNA molecules (*39, 40*). Mutants in any of genes 8, 9, 10, 18, or 19 do not produce mature T7 DNA from the intermediate forms (*14, 39, 50*). Thus, maturation of T7 DNA seems to require a minimal head structure (containing the proteins specified by genes 8, 9, and 10) and the functions of genes 18 and 19. This is similar to the situation in lambda and T4, where some kind of head structure seems to be required for maturation of the DNA (*38, 51*). As yet, no specific function has been assigned to either gene 18 or 19.

Summary

Amber mutations or deletion mutations have been found in 25 genes of T7, and the proteins specified by 23 of these genes have been identified by electrophoresis on polyacrylamide gels in the presence of SDS. These genes account for more than 90 percent of the coding capacity of T7 DNA, but there are a few genes in which mutations have yet to be found.

The genetic map of T7 is linear, and all transcription and translation proceed from left to right. Three classes of T7 genes can be distinguished by their times of expression, and the different classes lie in distinct regions of the T7 DNA: class I takes up the left-most 20 percent, class II the next 20 to 25 percent, and class III the right-hand 55 to 60 percent. Class I proteins are the first to appear after infection, followed by class II and III proteins. Class II proteins may appear slightly earlier than those of class III, but the main distinction is that the synthesis of class II proteins stops midway through the infection whereas class III proteins continue to be synthesized until lysis.

After infection, the class I genes are transcribed by the RNA polymerase of the host. A stop signal located approximately 20 percent from the left end of the T7 DNA normally prevents the host RNA polymerase from continuing down the molecule to transcribe the class II and III genes. One of the class I genes, gene 1, specifies a new RNA polymerase which transcribes the class II and III genes.

Something is known about the function of most of the T7 genes. Class I and II genes seem to be concerned mainly with the control of gene expression and with DNA metabolism, that is, breakdown of host DNA and synthesis of T7 DNA. The proteins of the mature phage particle are all specified by class III genes, as are some proteins involved in the assembly of the phage particle.

The availability of mutants defective in most T7 functions and the prospect of soon having a mutation in every T7 gene makes this phage ideal for the study of molecular mechanisms underlying control of gene expression, synthesis of DNA, genetic recombination, and the assembly of virus particles.

References and Notes

1. M. Demerec and U. Fano, *Genetics* **30**, 119 (1945).
2. M. Delbruck, *Biol. Rev. Cambridge Phil. Soc.* **21**, 30 (1946).
3. D. Fraser and R. C. Williams, *J. Bacteriol.* **65**, 458 (1953).
4. S. B. Dubin, G. B. Benedek, F. C. Bancroft, D. Freifelder, *J. Mol. Biol.* **54**, 547 (1970).
5. K. D. Lunan and R. L. Sinsheimer, *Virology* **2**, 455 (1956).
6. D. A. Ritchie, C. A. Thomas, Jr., L. A. MacHattie, P. C. Wensink, *J. Mol. Biol.* **23**, 365 (1967).
7. A. D. Hershey, E. Burgi, L. Ingraham, *Proc. Nat. Acad. Sci. U.S.* **49**, 748 (1963); H. Ris and B. L. Chandler, *Cold Spring Harbor Symp. Quant. Biol.* **28**, 1 (1963); L. A. MacHattie and C. A. Thomas, Jr., *Science* **144**, 1142 (1964).
8. F. W. Studier and J. V. Maizel, Jr., *Virology* **39**, 575 (1969).
9. J. V. Maizel, Jr., in *Fundamental Techniques in Virology*, K. Habel and N. P. Salzman, Eds. (Academic Press, New York, 1969), p. 334.
10. A. L. Shapiro, E. Vinuela, J. V. Maizel, Jr., *Biochem. Biophys. Res. Commun.* **28**, 815 (1967).
11. F. W. Studier, unpublished observations.
12. R. Hausmann and B. Gomez, *J. Virol.* **1**, 779 (1967).
13. F. W. Studier, *Virology* **39**, 562 (1969).
14. ——— and R. Hausmann, *ibid.*, p. 587.
15. A. S. Sarabhai, A. O. W. Stretton, S. Brenner, S. Bolle, *Nature* **201**, 13 (1964); A. O. W. Stretton and S. Brenner, *J. Mol. Biol.* **12**, 456 (1965).
16. R. H. Epstein, A. Bolle, C. M. Steinberg, E. Kellenberger, E. Boy de la Tour, R. Chevalley, R. S. Edgar, M. Susman, G. H. Denhardt, A. Lielausis, *Cold Spring Harbor Symp. Quant. Biol.* **28**, 375 (1963).
17. J. S. Parkinson and R. J. Huskey, *J. Mol. Biol.* **56**, 369 (1971).
18. D. A. Ritchie and F. E. Malcolm, *J. Gen. Virol.* **9**, 35 (1970).
19. F. W. Studier and M. N. Simon, unpublished observations.
20. R. W. Davis and N. Davidson, *Proc. Nat. Acad. Sci. U.S.* **60**, 243 (1968); B. C. Westmoreland, W. Szybalski, H. Ris, *Science* **163**, 1343 (1969); R. W. Davis, M. N. Simon, N. Davidson, in *Methods in Enzymology*, L. Grossman and K. Moldave, Eds. (Academic Press, New York, 1971), vol. **21**, part D, p. 413.
21. Y. Masamune, G. D. Frenkel, C. C. Richardson, *J. Biol. Chem.* **246**, 7874 (1971).
22. M. Gellert and M. L. Bullock, *Proc. Nat. Acad. Sci. U.S.* **67**, 1580 (1970).
23. R. W. Davis and R. W. Hyman, personal communication.
24. M. Chamberlin, J. McGrath, L. Waskell, *Nature* **228**, 227 (1970).
25. W. C. Summers and W. Szybalski, *Virology* **34**, 9 (1968).
26. W. C. Summers, *ibid.* **39**, 175 (1969).
27. W. A. Newton, J. R. Beckwith, D. Zipser, S. Brenner, *J. Mol. Biol.* **14**, 290 (1965); C. Yanofsky and J. Ito, *ibid.* **21**, 313 (1966); R. G. Martin, D. F. Silbert, D. W. E. Smith, H. J. Whitfield, Jr., *ibid.*, p. 357.
28. R. B. Siegel and W. C. Summers, *ibid.* **49**, 115 (1970).
29. W. C. Summers, personal communication.
30. ———, *J. Mol. Biol.* **51**, 671 (1970).
31. ——— and R. B. Siegel, *Nature* **228**, 1160 (1970).
32. R. W. Davis and R. W. Hyman, *Cold Spring Harbor Symp. Quant. Biol.* **35**, 269 (1970).
33. R. W. Hyman, *J. Mol. Biol.* **61**, 369 (1971).
34. J. W. Roberts, *Nature* **224**, 1168 (1969).
35. R. L. Millette, C. D. Trotter, P. Herrlich, M. Schweiger, *Cold Spring Harbor Symp. Quant. Biol.* **35**, 135 (1970).

36. H. Bremer and D. Yuan, *J. Mol. Biol.* **38**, 163 (1968); H. Manor, D. Goodman, G. S. Stent, *ibid.* **39**, 1 (1969); J. K. Rose, R. D. Mosteller, C. Yanofsky, *ibid.* **51**, 541 (1970).
37. T. G. Morrison and M. H. Malamy, *Nature New Biol.* **231**, 37 (1971).
38. W. F. Dove, *J. Mol. Biol.* **19**, 187 (1966).
39. R. Hausmann and K. LaRue, *J. Virol.* **3**, 278 (1969).
40. T. J. Kelly, Jr., and C. A. Thomas, Jr., *J. Mol. Biol.* **44**, 459 (1969).
41. L. W. Labaw, *J. Bacteriol.* **62**, 169 (1951); *ibid.* **66**, 429 (1953); F. W. Putnam, D. Miller, L. Palm, E. A. Evans, Jr., *J. Biol. Chem.* **199**, 177 (1952).
42. M. S. Center, F. W. Studier, C. C. Richardson, *Proc. Nat. Acad. Sci. U.S.* **65**, 242 (1970).
43. P. D. Sadowski and C. Kerr, *J. Virol.* **6**, 149 (1970).
44. M. S. Center and C. C. Richardson, *J. Biol. Chem.* **245**, 6285 (1970); *ibid.*, p. 6292; P. D. Sadowski, *ibid.* **246**, 209 (1970).
45. G. D. Frenkel and C. C. Richardson, personal communication; C. Kerr and P. D. Sadowski, *J. Biol. Chem.* **247**, 305 (1972); *ibid.*, p. 311.
46. P. Grippo and C. C. Richardson, *J. Biol. Chem.* **246**, 6867 (1971).
→ U. K. Laemmli, *Nature* **227**, 680 (1970).
48. J. Hosada and R. Cone, *Proc. Nat. Acad. Sci. U.S.* **66**, 1275 (1970); R. C. Dickson, S. L. Barnes, F. A. Eiserling, *J. Mol. Biol.* **53**, 461 (1970).
→ R. S. Edgar and W. B. Wood, *Proc. Nat. Acad. Sci. U.S.* **55**, 498 (1966).
50. R. Schlegel, personal communication.
→ F. R. Frankel, *Proc. Nat. Acad. Sci. U.S.* **59**, 131 (1968); A. G. Mackinlay and A. D. Kaiser, *J. Mol. Biol.* **39**, 679 (1969).
→ M. S. Reid and R. L. Bieleski, *Anal. Biochem.* **22**, 374 (1968).
53. Research was carried out at Brookhaven National Laboratory under the auspices of the U.S. Atomic Energy Commission.

A National Focus of Science and Research

Detlev W. Bronk

The title under which this is written appeared in the November 1922 issue of *Scribner's Magazine* (*1*). It was the subject of an article by George Ellery Hale, then honorary chairman of the National Research Council of which he had been a principal founder. He wrote:

> The stately plan of the City of Washington, conceived by L'Enfant under the personal inspiration of Washington himself, is rapidly assuming material form. The major axis . . . now terminates admirably in the massive Doric temple of the Lincoln Memorial. . . . Flanking it on its left, in the midst of a spacious square facing the Mall, another marble structure, also associated in its origin with the Civil War, is now rising. This is to be the home of the National Academy of Sciences and the National Research Council.
>
> As a center for the many activities organized by the Academy and Research Council the new building will bring together scientific investigators from all parts of the world. It will serve admirably for international scientific bodies when meeting in the United States, and in so far as may prove practicable it will be rendered available for meetings of the many national scientific and technical societies represented in the Research Council.

Now, 50 years later, the building begun by Hale, the creator of the Yerkes and Mount Palomar Observatories and of the California Institute of Technology, is completed. Increased in size for the needs of the present, its purposes and functions are the same. It had its origins in the election of Hale to the National Academy of Sciences 70 years ago.

Soon after his first attendance at a meeting of the Academy in the old red brick building of the Smithsonian Institution, Hale, never inhibited by his youth, wrote to Simon Newcomb, 30 years his senior: "The Academy should have a building where scientists could go whenever they visit Washington and there exchange ideas and develop wider friendships" (*2*). For 20 years Hale's persistent vision inspired his will to begin that which is now completed.

Hale built many monuments, physical and spiritual, throughout the world of science, but this is uniquely his. "It is indeed a temple that we now dedicate," said Gano Dunn, who was Hale's colleague in the creation of the first unit of the building that was being completed in 1924, "a temple that would not stand without the gifted vision and tireless devotion of George Hale. His spirit will be here always, for it was he who chose the words inscribed upon the dome: "Temple to Science, Pilot of Industry, Conqueror of Disease, Multiplier of the Harvest, Explorer of the Universe, Revealer of Nature's Laws, Eternal Guide to Truth" (*2*).

It was on a spring evening such as this that the spirit of Hale's temple of science first enmeshed me many years ago. I still recall that through the open windows came the soft fragrance of the magnolias many of us pleasantly associate with that place. I was impressed by Christian Herter's portrait of the founders, by the paneled dignity of the board room, by Robert Millikan, who presided. Of the proceedings I remember little, only the final question by George Stewart of Iowa: "Why are we here and what is the purpose of the National Research Council?" And then Millikan's prophetic reply, anticipatory of the roles wisely fulfilled by the Academy and its Council in the framework of the modern scientific endeavor.

The latter hours of that evening are vivid memories. Because Millikan was the father of my closest friend, he was to me a fatherly, friendly mentor. Until he and I took to our beds in the old Cosmos Club on Lafayette Square, he regaled me with memories of how the house of the Academy came to be. Hale had hoped that, by the time of the Academy's 50th anniversary celebration in 1913, Andrew Carnegie might have been persuaded to make a gift to the nation of an Academy building "of such a character as to unite the interests of the various scientific institutions" (*2*) of our country, some of which Carnegie himself had founded. But nothing came of his attempts to interest Carnegie in this project. In 1914, Hale again besieged Carnegie thus: "As the clearing house of American science, and its official center in both a national and an international sense, the building would be a contri-

The author was president of the National Academy of Sciences from 1950 to 1962, and is now president emeritus of Rockefeller University, New York 10021. This article is based on an address delivered at the dedication of the completed building of the National Academy of Sciences in Washington, D.C., on 26 April 1971.