Bacteriophage T4

Secondary article

Christopher K Mathews, *Oregon State University, Corvallis, Oregon, USA*

Analysis of molecular events following infection of *Escherichia coli* by bacteriophage T4 has revealed some of the most fundamental principles of biology, including the chemical nature of genes and mechanisms of gene expression. In addition, the phage has provided researchers with virus-coded proteins that are continuing to serve as reagents to help fuel the biological revolution that began in 1953 with Watson and Crick and is continuing today.

Article Contents

Introduction

Research on bacteriophages provided our earliest insights into the chemical nature of viruses, as well as illuminating general biological concepts, such as the identity of deoxyribonucleic acid (DNA) as the carrier of genetic information, the nature of the genetic code, and the existence of messenger ribonucleic acid (mRNA). The discovery of bacteriophages is generally credited to independent studies by Edward Twort in 1915 and Felix d'Herelle in 1917 (Stent, 1963). However, the river of information that has flowed from phage T4 did not approach flood stage until three decades later, when Demerec and Fano isolated from the sewage system in Brooklyn, New York, seven closely related phages whose common host was *Escherichia coli*. These phages, which were named T1 through T7, had desirable properties for laboratory study – a host bacterium that is easy to culture, rapid adsorption to host cells, a high probability of productive infection, and abundant and rapid growth within the host. The so-called T-even phages (T2, T4 and T6) were particularly closely related.

In the late 1940s Max Delbrück, a physicist attracted to biology, suggested that researchers in the field, each of whom was studying a separate phage–host system, should join forces and study one bacteriophage, namely T4, and its close relative T2. Almost all workers in the fledgling area of molecular biology were persuaded by the force of Delbrück's personality, and this greatly accelerated the pace of research. Highlights of the 1950s and early 1960s included the Hershey–Chase demonstration that genes are composed of DNA, Delbrück's discovery of recombination in T4, Benzer's fine-structure gene mapping, Luria's discovery of host-induced modification and restriction, Streisinger's demonstration of the validity of the genetic code, Cohen's discovery that phage infection creates new metabolic pathways in infected cells, the observations of Volkin and Astrachan, and of Brenner and colleagues, that led to Spiegelman's demonstration of messenger RNA, the analysis of viral assembly pathways *in vitro* by Edgar and Wood, and many other fundamental contributions (Cairns *et al*., 1966). Another crucial advance was the discovery of conditional lethal mutants, derived from both temperature-sensitive and suppressible nonsense mutations (Epstein *et al*., 1963), which mapped in virtually all T4 genes whose products were essential to growth, thus permitting geneticists and biochemists to join forces in defining all of the processes in the viral life cycle.

Perhaps the only drawback to the choice of T4 as the benchmark system for virus research was the fact that it is a virulent phage, which always initiates a cycle of lytic growth upon infection. Concurrent investigations in France focused upon bacteriophage λ, and led to the discovery of lysogeny, in which infection can lead to a long-term nonlethal relationship between virus and host. Analysis of this temperate bacteriophage generated insights into DNA replication, transcriptional regulation, latent virus infection mechanisms and genetic recombination, which rival in importance the lessons taught by T4.

Classification of Bacteriophage T4

Many different criteria have been used for taxonomic classification of viruses: host range, morphology (virion size and shape, presence or absence of an envelope, capsid symmetry), nature of the genome (DNA or RNA, single strand, negative strand, etc.), physical properties, replication strategies (lytic or lysogenic, etc.), geographic distribution, pathogenicity, etc. The International Committee on Taxonomy of Viruses has developed a universal system of virus taxonomy, based upon the premise that almost all existing viruses of humans, domestic animals and economically important plants have been isolated and classified. The taxonomy includes orders, families, subfamilies, genera and species, much like the taxonomy used for cellular organisms. According to this taxonomy, T4 and its cousins T2 and T6 are classified in the family *Myoviridae* – double-strand DNA phages with contractile tails. The genomes of these viruses show about 85% homology by analysis of DNA heteroduplexes. Subsequent to the isolation of the seven T phages, a large number

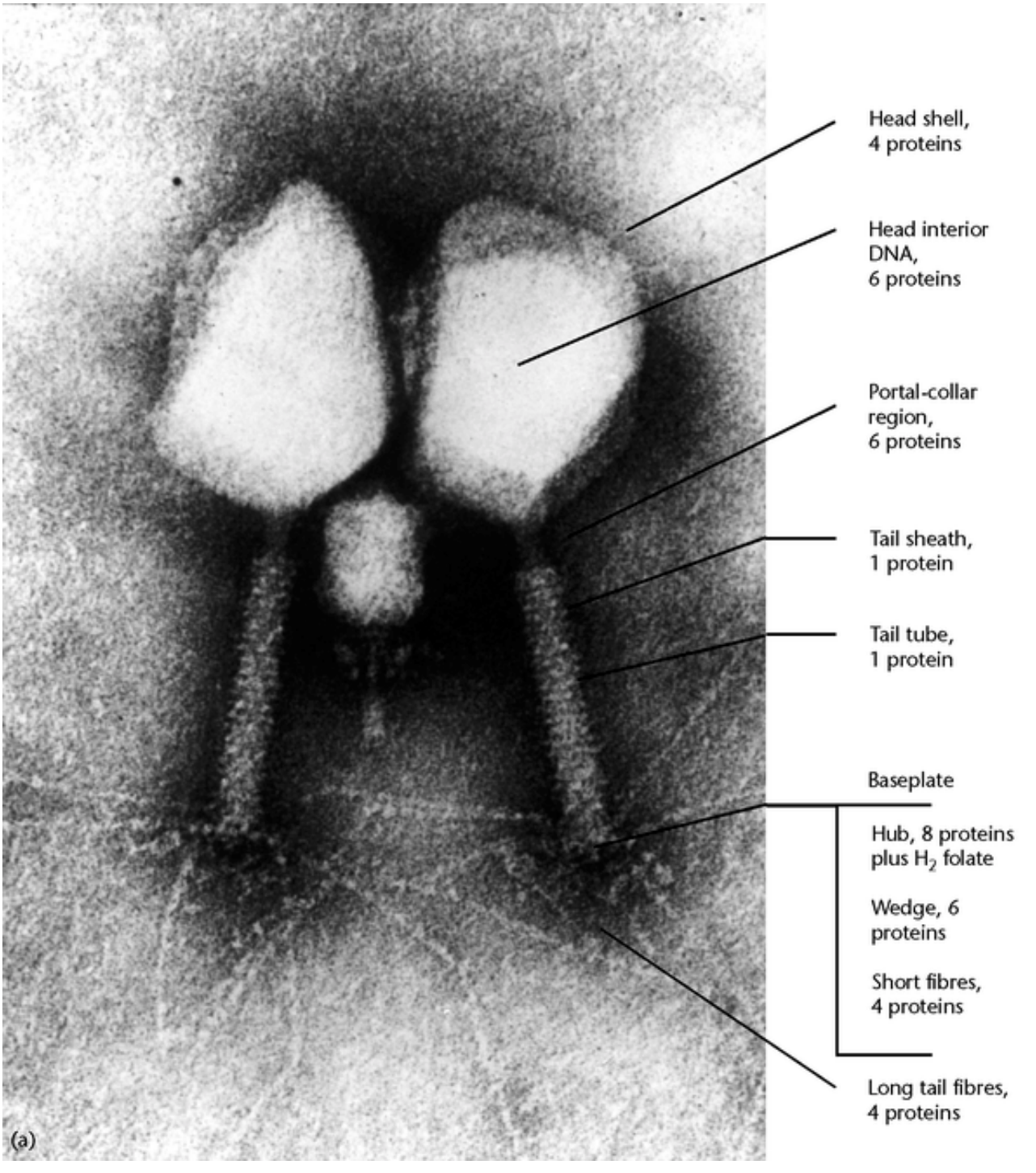

Figure 1 Structure of the T4 particle, as revealed by (a) electron microscopy (× 500 000) and (b) protein analysis of the separate substructures. The gene products in each substructure are identified, either by number or letter denoting the gene. Reprinted with permission from Eiserling FA and Black LW (1994) in: Karam JD (ed.) *Molecular Biology of Bacteriophage T4*, p. 209. Washington, DC: American Society for Microbiology.

of closely related phages, the RB series, was isolated (Russell and Huskey, 1974).

Propagation and Preparation of Bacteriophage Stocks

T4 has desirable properties for isolation of large phage stocks, including a short latent period (i.e. rapid growth), a high burst size and a high plating efficiency. This means that phage growth is rapid following infection of a host cell, the number of phage particles produced in a single infection is high, and a phage particle has a high probability of producing a plaque when plated on a Petri dish along with a lawn of susceptible bacteria. Typically one can produce a high-titre T4 stock by transferring phage from a single plaque (10^4–10^6 virions) to a culture of *E. coli* in the early phase of exponential growth ($1–5 \times 10^7$ colony-forming cells per millilitre), in a minimal medium of glucose and inorganic salts, or an enriched medium containing amino acids in addition to glucose or glycerol. Phage released from those few cells that become infected in turn infect other cells in the still-growing culture. After two or three cycles of phage infection, reproduction and release by lysis, essentially all of the cells in the culture have become infected, and final phage titres may be as high as 5×10^{11} plaque-forming units per millilitre of culture. By these means, one can easily prepare a stock containing 10^{14} phage particles, all of them derived from the phage in a single plaque, which in turn originated from a single T4 virion. Thus, large amounts of genetically homogeneous virions can readily be isolated.

Typically, the first step in purifying T4 is addition of a few drops of chloroform to the lysate. This brings about lysis of any infected cells that have not yet ruptured, and it sterilizes the lysate, if long-term storage is contemplated. Early purification steps usually involve differential centrifugation. If a large lysate is being worked up, phage can be precipitated by addition of polyethylene glycol or ammonium sulfate to the lysate; one then decants the supernatant, reducing the volume of fluid that must be centrifuged. Stocks are usually treated with deoxyribonuclease, to digest any unpackaged DNA; the treatment is not harmful to mature virions. If desired, the virus stock can be highly purified by centrifugation to equilibrium in a caesium chloride gradient, followed by dialysis to remove the caesium chloride. T4 stocks at all stages of purity can be stored at 4°C for months with negligible decline in titre, or for years with only modest loss of infective units.

Physical and Chemical Properties of the Virion

The T4 virion is composed of approximately equal weights of DNA and protein. Whereas the genome comprises but a single molecule of double-stranded DNA, the protein make-up of the virus is extremely complex. **Figure 1** shows an image of the virus as reconstructed from electron micrographic images. A distorted icosahedral head, which contains the DNA genome, is attached by a neck, surrounded by a collar with whiskers, to the tail, which contains an inner core or tube, surrounded by a contractile sheath containing 144 identical subunits. Each sheath subunit contains one molecule of bound adenosine triphosphate (ATP), whose cleavage evidently drives contraction of the tail sheath in the early stages of infection.

At the base of the tail is the complex baseplate, consisting of two separate substructures: six peripheral wedges surrounding a central core. Attached to the baseplate are two sets of fibres, long and short. The protein composition of each substructure is indicated on the figure by the use of gene numbers or letters. Thus, the 20 faces of the icosahedral head contain the products of genes 23, *hoc* and *soc*, while each vertex contains a molecule of gp24; gp stands for gene product so gp24 is the product of gene 24. Similarly, the short tail fibres contain gp11. The detailed analysis of each substructure was made possible by the existence of conditional lethal mutations in each of the numbered genes. Thus, for example, infection of *E. coli* with a phage bearing an unsuppressed *amber* mutation in gene 18 is abortive, because no completed tails are formed. However, morphologically normal heads are formed, and these can be purified for analysis by means similar to those described above for complete virions. By these means, all of the major substructures have been purified and analysed with respect to gene product composition and stoichiometry.

As might be expected from the structural complexity of T4, the phage is genetically complex as well, and its linear double-stranded DNA genome is large, about 168.9 kilobase pairs (kbp) in length. However, DNA molecules purified from virions are significantly larger, about 172 kbp in length. The reason for this apparent discrepancy is that the genome is terminally redundant; about 2% of the genome is repeated at each end of the molecule. In addition, the genome is circularly permuted with respect to base sequence; different DNA molecules vary with respect to starting and ending points. As pointed out by Streisinger *et al.* (1967), this virion genome arrangement results from a replication strategy designed to protect the ends of linear chromosomes. Because of the requirement that DNA replication initiate by addition of nucleotides to a 3′ primer terminus, leading-strand replication can proceed to the end of a linear DNA template, but lagging-strand replication cannot. Thus, in the absence of a strategy to prevent chromosome loss, a little bit of the 5′ end of a linear DNA molecule will be lost in each replication event. Eukaryotic cells protect against this loss by adding telomeres to chromosome tips, while most bacteria and many phages use a circular genome that simply contains no problematic ends. The strategy adopted by T4 involves recombination among newly replicated DNA molecules, yielding highly branched molecules that

include concatemers, or end-to-end linear DNA aggregates up to 20 genome-equivalents in length. A DNA head-packing mechanism that recognizes a particular length and cuts at regular intervals of about 1.02 genome lengths as the DNA is packaged into heads generates the observed genome arrangement, as suggested by Scheme 1 where the alphabet represents the genome.

This replication scheme explains what was initially a puzzling observation; when the first detailed genetic map was obtained in 1963 from recombination analysis, the map was circular. Circular permutation of the genome explains how a circular map can be generated from a linear chromosome.

The T4 DNA molecule is distinctive also in its nucleotide structure. As shown in 1952 by Wyatt and Cohen, T-even DNA contains no cytosine, but instead contains a modified pyrimidine, 5-hydroxymethylcytosine, or HMC. This modified base pairs with guanine, as does cytosine in most other DNAs. HMC is further modified, by formation of glycosidic bonds between the hydroxymethyl groups and C1 of glucose. The glucosylation patterns of T-even and RB phages vary; in T4 every HMC residue is glucosylated, with 70% of the glycosidic bonds in the α configuration and 30% in the β. By contrast, T2 and T6 DNAs both contain some nonglucosylated HMC residues and some diglucosylated residues (the diglucosyl unit is gentiobiose).

Subsequent investigation by Cohen revealed the process by which these DNA base modifications occur: infection stimulates the synthesis of new enzymatic activities that carry out hydroxymethylation at the deoxyribonucleoside monophosphate level. This was a singularly important finding, because it suggested that viruses encode new enzymes that participate in propagation of the virus and destruction of the host cell, and that these enzymes could be sites for specific chemotherapeutic attack in viral disease treatment. Later, the glucosyl modification reactions were found to occur at the polynucleotide level, with virus-coded enzymes transferring glucose from uridine diphosphate glucose to hydroxymethyl groups in polymerized DNA. The function of glucosylation is to protect phage DNA from a still-uncharacterized enzyme system of *E. coli*, which recognizes nonglucosylated HMC-containing DNA and initiates its breakdown. The hydroxymethyl modification is a protective device for phage against its own activities. After infection, T4 uses RNA polymerase of the host cell to transcribe its own genes. However, T4 modifies *E. coli* RNA polymerase so that it can only transcribe HMC-containing DNA and cannot act upon cytosine DNA. Moreover, as one source of nucleotides for its own DNA replication, T4 encodes nucleases that degrade the host-cell genome. These enzymes degrade any DNA that contains cytosine, but any HMC-containing DNA is impervious to attack.

Replication and Life Cycle

The biological parameters of the T4 infectious process can be described in the context of a one-step growth curve (**Figure 2a**). In this experiment, aliquots of an infected cell culture are plated at various times after infection on a lawn of susceptible bacteria. Either one phage particle or one infected cell gives rise to one plaque (an 'infective centre'). Typically, lysis begins at about 25 min after infection at 37°C, when the infective centre titre rises rapidly, as each infected cell releases 100 or more progeny phage particles, each capable of initiating plaque formation. The intracellular development of phage can also be followed by treating each sample with chloroform before plating, a treatment that releases all mature intracellular virions that have been formed and allows each one to initiate plaque formation. By 30 min after infection, all of the infected cells have lysed, and there is no further increase in infectious phage titre.

A striking feature of the curve is the 'eclipse period', an interval of about 12 min during which no infective phage particles can be recovered from infected cells. What events occur during this period to prepare the cell for the rapid accumulation of mature intracellular virus particles, which begins at 12 min and continues until lysis begins? **Figure 2b** summarizes the timing of the major metabolic events occurring after the initiation of infection at time zero. The production of infectious viral progeny is the result of exquisitely timed biochemical steps that occur after adsorption of phage to the cell wall initiates the process, including a sequential pattern of gene activation, DNA replication, the synthesis of structural proteins and viral substructures, their assembly into virus particles and their release by cell lysis. **Figure 3** summarizes these events pictorially and sets the stage for our metabolic discussion.

Adsorption and penetration

Infection begins with molecular recognition between the tips of the long tail fibres and receptors in the outer layer of the bacterial cell surface – the lipopolysaccharide in B strains of *E. coli*, and the OmpC outer membrane protein in K strains. This is followed by repositioning of the phage so that the baseplate sits above a site of adhesion between inner and outer membranes of the bacterium. Short tail fibres extending from the baseplate (gp11; see **Figure 1**) then pin the phage to the cell, via binding to the outer membrane polysaccharide. These events trigger a conformational change in the structure of the baseplate, shown by electron microscopy to change a hexagonal structure to a six-pointed star with a large central hole. One of the baseplate proteins, gp5, has lysozyme activity and digests polysaccharide linkages below the baseplate. Next, the tail sheath subunits undergo a structural change resulting in contraction of the sheath, which in turn drives the tail core through

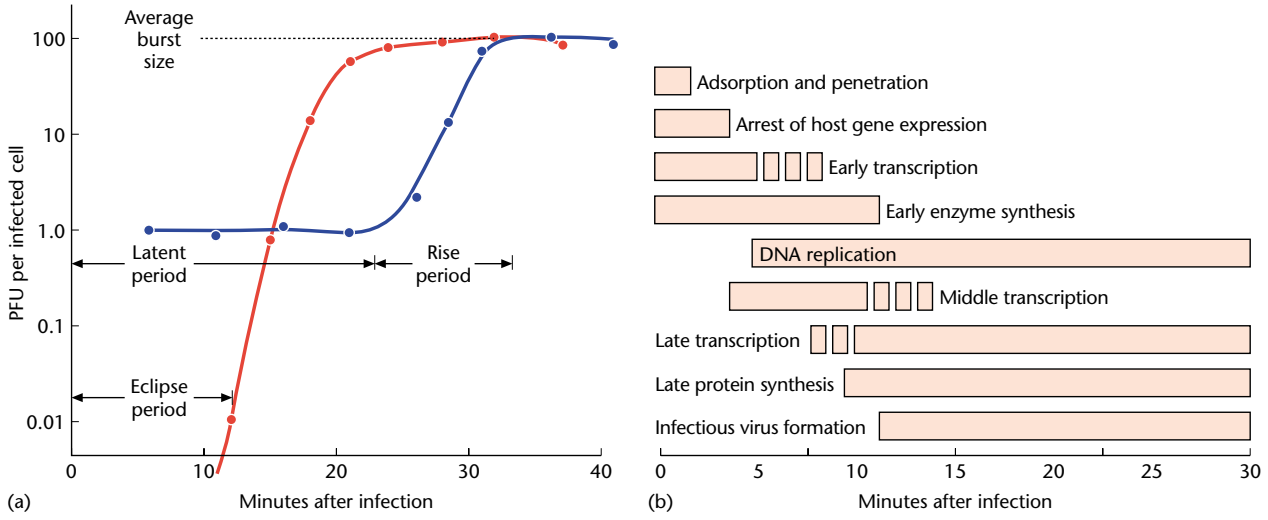

Figure 2 (a) A one-step growth curve for bacteriophage T4 infecting *E. coli* B at 37°C. Blue circles, infective centres; red circles, intracellular phages; see text for details. PFU, plaque-forming unit. (b) A chronology of major events in T4 infection. The precise times of expression of genes within the major temporal classes vary somewhat with individual genes, as shown. Under some conditions, lysis may occur later than 30 min after infection; under those conditions, replication, transcription, translation and viral assembly continue until lysis. Reprinted with permission from Mathews CK (1994) in: Karam JD (ed.) *Molecular Biology of Bacteriophage T4*, p. 2. Washington, DC: American Society for Microbiology.

the newly enlarged hole in the baseplate and the outer membrane of the cell. DNA is rapidly ejected through this tube into the cell interior, in a process aided by the transmembrane proton gradient and establishment of a voltage-dependent ion-selective channel.

Early metabolic events

Immediately after DNA injection, several events result in the arrest of essentially all host-cell macromolecular synthesis – DNA replication, transcription and protein synthesis. RNA polymerase of the host cell is recruited to transcribe a set of T4 genes called 'early'; these are genes

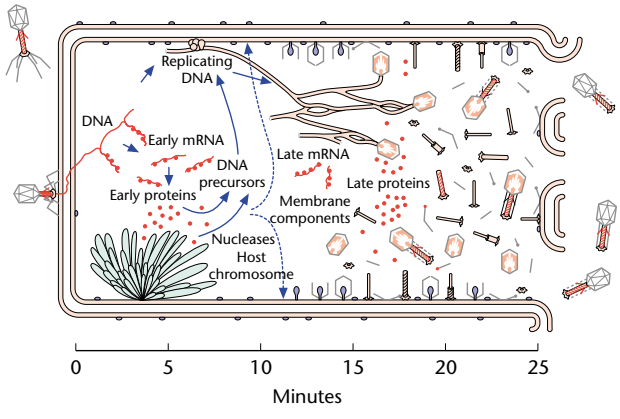

Figure 3 Overview of the T4 developmental programme. Created by FA Eiserling and reprinted with permission from Mathews CK (1994) in: Karam JD (ed.) *Molecular Biology of Bacteriophage T4*, p. 5. Washington, DC: American Society for Microbiology.

transcribed by unmodified *E. coli* RNA polymerase. Among these genes is *alc*, whose protein product interacts with RNA polymerase and renders it incapable of transcribing cytosine-containing DNA. This is a critical event in the arrest of host-cell transcription. This also leads, along with the action of other T4 proteins, to unfolding of the host-cell nucleoid, making the bacterial chromosome accessible to degradation by nucleases encoded by the phage and which, as noted earlier, selectively degrade DNA that contains cytosine instead of the modified base HMC. This provides a pool of 10–20 genome-equivalents per cell of precursors to viral DNA.

However, because each cell will ultimately produce about 200 viral gene copies, it is evident that most of the raw material for DNA synthesis must come from *de novo* deoxyribonucleotide synthesis. In fact, there is a massive increase in the rate at which infected cells synthesize DNA precursors, brought about by synthesis of several virus-coded enzymes of deoxyribonucleoside triphosphate (dNTP) synthesis. These include novel activities such as those that replace cytosine in DNA with HMC, but also increases in preexisting activities, including ribonucleotide reductase and thymidylate synthase. These latter activities result from expression of phage genes to form enzymes readily distinguishable from their host-cell counterparts that catalyse the same reactions. The phage-coded enzymes form a multienzyme complex, which facilitates the multi-step reaction pathways leading to dNTPs and their delivery to viral DNA replication sites.

DNA replication

During the first few minutes after T4 infection, the cell prepares for phage DNA replication, which will occur at rates up to tenfold higher than those seen in uninfected *E. coli*. Deoxycytidine triphosphate (dCTP) is broken down to deoxycytidine monophosphate (dCMP) by phage-coded enzymes, and then modified to give 5-hydroxymethyl-dCTP, the substrate for polymerization of HMC. Other phage-coded enzymes are synthesized, and they begin to work on the ribonucleotides that accumulate after arrest of host-cell ribosomal RNA (rRNA) and transfer RNA (tRNA) synthesis. These ribonucleotides are reduced to deoxyribonucleotides and converted to dNTPs.

At about 6 min after infection the earliest phage DNA replication can be detected (**Figure 2b**). Initiation of T4 DNA replication occurs by two distinct mechanisms. First, at several distinct origin sites, *E. coli* RNA polymerase transcribes a few nucleotides, yielding RNA primers that can be extended by the phage-specific replication complex. Second is a mode of recombination-dependent initiation. As noted earlier, recombination is crucial to the replication strategy for the linear T4 genome; mutants defective in recombination functions can initiate DNA replication, but infection is abortive, because concatemeric replication intermediates cannot be formed. Recombinational DNA replication initiation involves invasion of a duplex by a homologous single-strand DNA 3′ end, with extension from that 3′ end directed by a template DNA strand of the invaded duplex. By this means, recombination can occur at the ends of identical newly replicated DNA molecules, or it can form a branched structure if the recombination event involves two DNAs with permuted sequences. Eventually, highly branched structures with concatemeric arms are formed, and in the several minutes before DNA packaging into phage heads begins, all of the replicating DNA in a cell may be present as one gigantic branched molecule, 100 times larger than the DNA packaged into an individual virion.

Much of our understanding of the mechanism of replicative DNA chain elongation has come from analysis of T4 DNA replication *in vitro*, driven by purified T4 proteins (Alberts and Miake-Lye, 1992). The laboratories of Alberts and Nossal showed that double-stranded DNA replication can be driven at rates and fidelity comparable to values determined *in vivo* by a mixture of seven proteins: gp43 (DNA polymerase), gp45 (sliding clamp), gp44 and gp62 (clamp loading complex), gp32 (single-strand DNA-binding protein), gp41 (helicase) and gp61 (primase). These classic studies, which came to fruition in the mid-1970s, have guided research on many other biological systems and have revealed a common plan, used by both prokaryotes and eukaryotes, for replicating duplex DNA.

Temporal control of transcription

The T4 life cycle, like that of most DNA viruses, is controlled in large part by temporal regulation of transcription (**Figure 2b**). T4 genes are classified by their time of transcription – early, middle or late. All T4 genes are transcribed by RNA polymerase of the host cell, but phage-specified modifications of the host's RNA polymerase direct the enzyme to different promoters at different times. Within the first few minutes, unmodified *E. coli* RNA polymerase transcribes genes with promoters similar to those for bacterial genes. Early gene products include proteins that replace the *E. coli* σ subunit and adenosine diphosphate (ADP)-ribosylate the α subunit, changes that direct the enzyme to different promoters. Later still, two additional phage proteins (gp33 and gp55) associate with RNA polymerase, preparing it for late transcription. Late genes are preferentially transcribed from replicating DNA, and there is an additional requirement for the polymerase-accessory proteins used also in DNA replication – gp44, gp45 and gp62. These latter proteins have been characterized as a 'mobile enhancer', which tracks along DNA and directs modified RNA polymerase to the correct targets (Herendeen *et al*., 1992).

Morphogenesis and DNA packaging

Whereas much of the T4 propagation programme is directed by temporal control of gene expression, assembly of the particle involves simultaneous gene expression, but an obligatory sequence of gene product interactions. With few exceptions, the myriad structural proteins of the viral particle are synthesized as the products of late genes. The different substructures – heads, tails and tail fibres – are assembled by independent morphogenetic pathways. This was established from analyses of the assembly-defective mutants described by Epstein *et al*. (1963). For example, a gene 23 mutant could not form phage heads, but abortively infected cells produced normal tails and tail fibres. These observations were used by Edgar and Wood as the basis for an *in vitro* assay system that revealed the sequential order of gene product interactions. The assay involved complementation; for example, mixing an extract of gene 23 mutant phage-infected bacteria, which made no heads, with a similar extract from a gene 18 mutant phage infection, where completed tails were not formed, led to the assembly *in vitro* of infectious particles, because the defect in each extract was complemented by the normal function expressed in the other. Application of this method allowed investigators to define the complete assembly pathways for heads and the several tail subassemblies: fibres, baseplate cores and wedges, sheathed tail and tail collar. One of the most important problems, still not completely solved, involves the mechanism of DNA packaging into the head and the process by which about 1.02 genome-equivalents of DNA is recognized as a 'headful'. DNA enters a head

precursor, the prohead, through one vertex, assisted by the products of genes 16 and 17. Expansion of the prohead and an ATP-dependent process involving gp17 are both involved in pulling DNA into the head.

Lysis

The final step in phage production is lysis, with release of progeny phage to the medium (**Figure 2a**). Normally this coincides with the metabolic death of the infected cell. Action of the *t* gene product somehow disrupts the inner membrane, allowing the lysozyme encoded by gene *e* to reach the outer membrane, where it attacks glycosidic bonds in the lipopolysaccharide, allowing osmotic forces to rupture the weakened cell wall. Although this process sounds simple, it is still quite poorly understood. For example, reinfection of a previously infected cell causes 'lysis inhibition', a process that prolongs the metabolic lifetime of the infected cell and permits continued phage production for many hours, with eventual phage yields approaching 1000 per cell. Equally mysterious, *r* mutants of T4 display a rapid-lysis phenotype, because they cannot establish lysis inhibition. The *r* mutants map in three genes, *rI*, *rII* and *rIII*. The *rII* gene in particular has been used in genetic analyses of recombination, complementation and mutagenesis. Despite the extensive attention this gene has received, the biochemical function of the *rII* gene product is still quite obscure.

Ecology of Bacteriophage T4

Like all other bacteriophages, the ecological distribution of T4 is dictated by the distribution of its host. Because *E. coli* normally inhabits the intestinal tract of humans, it is not surprising that the seven T phages were isolated from a sewage treatment plant. What may be more surprising is the diversity of different T4-related phages isolated from similar habitats. In 1964, Rosina Berry isolated several dozen such phages from municipal sewage treatment plants on Long Island, not far from Cold Spring Harbor (Russell and Huskey, 1974). These phages were all distinct by various biological criteria, but were related to T4 immunologically and by the apparent presence of HMC in their DNA. While it is not clear how such great diversity arose in a limited range of habitats, the RB phages have been useful to T4 workers. For example, the T4 DNA polymerase encoded by gene 43 has not been crystallized despite numerous attempts. However, gp43 of the closely related phage RB69 has been crystallized and its structure determined, thereby providing useful insights into the structure of the corresponding T4 enzyme.

Other T4-related phages have been isolated in hospitals, zoos and treatment plants from locations as diverse as Denver, Colorado; Tübingen, Germany; and Tbilisi, Georgia. As pointed out by Kutter *et al*. (1996), the natural hosts in the wild for these phages may be bacteria other than *E. coli*.

Evolution of Bacteriophage T4

As a virulent phage, T4 is limited in its ability to exchange genetic material with its host. Related to this, perhaps, much of the genome of T4 and its relatives is unique, or at least distinctive. As of 1996, only 42 of the nearly 300 open reading frames in the T4 genome had shown significant similarity to any sequences deposited in GenBank (Kutter *et al*., 1996). Much of the sequence similarity is in coding sequences for enzymes of DNA replication and DNA precursor synthesis, suggesting that these genes may have been recruited from cellular sources and retained because they increased the ability of the phages to reproduce during their very short life cycle.

Probably the most interesting question about evolution of T4-related phages is the evolutionary origin of the three introns found in T4 genes: *td*, the structural gene for T4 thymidylate synthase; *nrdB*, which encodes the small subunit of aerobic ribonucleotide reductase; and *nrdD*, the large subunit of anaerobic ribonucleotide reductase. T4 was the first prokaryote found to contain introns, and the discovery provoked a lively debate over whether introns have an ancient or recent evolutionary origin for T4. The 'introns-ancient' theory, for example, postulates that prokaryotic introns were once far more abundant than they are today, that they are being lost through evolution, and that introns found in phages today are molecular fossils. Belfort (1991) has argued that phage introns are largely uninformative for this question, partly because T4 is just as closely related to eukaryotes as it is to bacteria, suggesting that it shares ancestry with both. The T4 introns have characteristics of mobile genetic elements; each intron encodes a 'homing' endonuclease, capable of assisting in the insertion of an intron into a new gene. Although this tends to support an 'introns-recent' evolutionary model, in which introns arose by horizontal gene transfer, Belfort contends that evidence is equally strong for ancient and recent origins for introns.

Applications of Bacteriophage T4: Molecular Biology and Phage Therapy

In addition to providing the biological system of choice for fundamental investigations of DNA replication, transcription mechanisms, mutagenesis, recombination and morphogenesis, T4 has provided researchers with reagents that have applications in both basic and applied research. A partial list includes polynucleotide kinase, widely used for

5′ end-labelling of DNA (but an enzyme of still unknown function in T4 biology); DNA ligase, for years the only available enzyme capable of blunt-end joining in gene cloning; the ultraviolet repair enzyme encoded by gene *denV*, used as a reagent for quantitative analysis of thymine dimers; and gp32, used to coat single-strand DNA molecules for electron microscopic analysis. A multiple mutant of T4 that contains cytosine instead of HMC in its DNA has found wide use as a transducing phage for strain construction in *E. coli*, and somewhat less wide use as a cloning vector.

Currently, phage researchers are returning to an application that generated widespread attention in the early years of phage research, namely, their use as antibacterial agents in treating infectious diseases. For years the conventional wisdom held that phages had received a thorough trial as antibiotics in the 1920s and 1930s, and that the trial had failed. Currently that conventional wisdom is being challenged. Kutter (1997) has pointed out that the early therapeutic trials failed for reasons that could be predicted and dealt with in light of today's knowledge. For example, with infection by multiple bacterial species, multiple phages need to be administered. Moreover, to minimize development of phage-resistant strains, treatment with mixtures of phages could be advised, such that resistance could arise only from rare multiple independent mutational events. Although it is not clear that T4 will be in the vanguard of the re-emerging interest in phage therapy, the molecular understanding revealed through several generations of T4 research is clearly lighting the way for some promising potential applications of phages in treating infectious disease.

References

Alberts B and Miake-Lye R (1992) Unscrambling the puzzle of biological machines: the importance of the details. *Cell* **49**: 1–12.

Belfort M (1991) Self-splicing introns in prokaryotes: migrant fossils? *Cell* **64**: 9–11.

Cairns J, Stent GS and Watson J (eds) (1966) *Phage and the Origins of Molecular Biology*. Plainview, NY: Cold Spring Harbor Laboratory Press.

Epstein RH, Bolle A, Steinberg C *et al*. (1963) Physiological studies of conditional lethal mutants of bacteriophage T4. *Cold Spring Harbor Symposia on Quantitative Biology* **28**: 375–392.

Herendeen DR, Kassavetis GA and Geiduschek EP (1992) A transcriptional enhancer whose function imposes a requirement that proteins track along DNA. *Science* **256**: 1298–1303.

Kutter EM (1997) Phage therapy: bacteriophages as antibiotics. Unpublished manuscript posted on the Internet. URL: [http://www.evergreen.edu/user/T4/PhageTherapy/Phagethea.html]

Kutter E, Gachechiladze K, Poglazov A *et al*. (1996) Evolution of T4-related phages. *Virus Genes* **11**: 285–287.

Russell RL and Huskey RJ (1974) Partial exclusion between T-even bacteriophages: an incipient genetic isolation mechanism. *Genetics* **78**: 989–1014.

Stent GS (1963) *Molecular Biology of Bacterial Viruses*. San Francisco: Freeman.

Streisinger G, Emrich J and Stahl MM (1967) Chromosome structure in phage T4, III. Terminal redundancy and length determination. *Proceedings of the National Academy of Sciences of the USA* **57**: 292–295.

Further Reading

Cohen SS (1968) *Virus-induced Enzymes*. New York: Columbia University Press.

Edgar RS and Wood WB (1966) Morphogenesis of bacteriophage T4 in extracts of mutant-infected cells. *Proceedings of the National Academy of Sciences of the USA* **55**: 498–505.

Karam JD (ed.) (1994) *Molecular Biology of Bacteriophage T4*. Washington DC: American Society for Microbiology.

Mathews CK, Kutter EM, Mosig G and Berget P (eds.) (1983) *Bacteriophage T4*. Washington DC: American Society for Microbiology.