17 Replication Strategies of Small and Medium-sized DNA Viruses

CHAPTER

- DNA viruses express genetic information and replicate their genomes in similar, yet distinct, ways
- PAPOVAVIRUS REPLICATION
- Replication of SV40 virus – the model polyomavirus
- The replication of papillomaviruses
- THE REPLICATION OF ADENOVIRUSES
- Physical properties of adenovirus
- The adenovirus replication cycle
- REPLICATION OF SOME SINGLE-STRANDED DNA VIRUSES
- Replication of parvoviruses
- DNA viruses infecting vascular plants
- The single-stranded DNA bacteriophage ΦX174 packages its genes very compactly
- QUESTIONS FOR CHAPTER 17

DNA viruses express genetic information and replicate their genomes in similar, yet distinct, ways

Given that DNA is the universal genetic material of cells, it is not particularly surprising that viruses utilizing DNA as their genome comprise a significant proportion of the total number of known viruses. It also is not particularly surprising that such viruses will often use a significant proportion of cellular machinery involved in decoding and replicating genetic information encoded in double-stranded (ds) DNA, which is, after all, the stuff of the cellular genome.

While it might be expected that all viruses with DNA genomes would follow a generally similar pattern of replication, this is not the case. Indeed, viruses with DNA genomes utilize as many variations on a general replication strategy as do RNA viruses. There are both naked and enveloped DNA-containing viruses, and a number of DNA viruses encapsidate only a single strand of DNA. One group of animal viruses utilizing DNA as genetic material replicates in the cytoplasm of eukaryotic cells, and some DNA viruses infecting plants contain multipartite genomes. A major and extremely important group converts RNA into DNA while a related group converts RNA packaged in the virion into DNA as the virus matures!

While one can make useful generalizations concerning the replication of DNA viruses (indeed,

one must if the material is to be readily mastered), it is wise to treat such generalities as only basic guides. Thus, viruses of eukaryotic cells that replicate using the nucleus express their RNA using cellular transcription machinery, but bacterial DNA viruses as well as at least one group of insect DNA viruses (the baculoviruses) encode one or a number of novel RNA polymerases or specificity factors to ensure that only viral mRNA is expressed following infection. Similarly, the cytoplasmic-replicating DNA genome-containing poxviruses of eukaryotes encode many enzymes involved in transcription and mRNA modification.

Many DNA viruses use DNA replication enzymes and mechanisms that are generally related to the processes seen in the uninfected cell, but there is one major complexity when DNA replication of viruses is considered. This is the fact that while all viral DNA replication requires a primer, many groups do not utilize RNA primers! Thus, one of the basic tenets of the process outlined in Chapter 13 is violated.

Viruses with linear genomes face a major problem that also affects the replication of cellular chromosomal DNA. This "end problem" derives from the fact that the primer for DNA replication must be able recognize short stretches of the viral genome—either through base pairing or through specific DNA-protein interactions. Consider the problem for discontinuous strand DNA synthesis as shown in Fig. 13.1 in Chapter 13. When the primer anneals to the very 3′ end sequences of the template, DNA replication can proceed 5′ to 3′ down to the next fragment. But how is the primer to be removed and replaced with DNA? There is no place for a new primer to anneal upstream of this last gap to be filled. This situation means that the viral genome would have to become shorter every time it replicated and would rapidly disappear!

Different linear DNA viruses have evolved different means to overcome this end problem. Herpesviruses and many bacterial DNA viruses have genomes with repeated sequences at their terminals so that the viral genome can become circular via a recombination event following infection. Thus, even though the virion DNA is linear, replicating viral DNA in the cell is either circular or joined end to end in long **concatamers**. These structures are then resolved to linear ones when viral DNA is encapsidated.

Adenovirus, on the other hand, has solved the problem by using a primer that is covalently bound to a viral protein that binds to the viral DNA's end. Further, adenovirus DNA proceeds only continuously; there is no discontinuous strand synthesis.

Small single-stranded (ss) DNA viruses, like parvoviruses, have solved the problem by encoding a complementary repeat sequence at the end that allows the genome to form a "hairpin loop" at the end; thus, the end of the molecule is not free. A similar solution is seen in the genome structure of poxvirus. Like chromosomal DNA, this linear DNA genome is covalently closed at its ends. Thus, in effect, replication just proceeds "around the corner" onto the complementary strand.

Another "general" strategy found in the replication of nuclear-replicating eukaryotic viruses and many bacteriophages is the establishment of infections where the viral genome remains in lifelong association with its host. Such a process has tremendous evolutionary advantages to any pathogen, but again, the specifics of the process in terms of mechanism differs greatly between the groups.

Given these variations, it is important to describe the basic processes of DNA virus replication in a logical way, and this is perhaps best done by consideration of how much cellular function and cellular transcriptional machinery are needed for productive replication. This is roughly correlated with overall size of the viral genome. The usefulness of such a grouping is that the viruses in each group share certain similarities in their replication strategies. Equally important, they share similarities in the way they can alter cells during the replication process. Such alterations can have profound and far-reaching effects on the host's health.

The discussion of three unrelated families of viruses infecting eukaryotic cells in this chapter follows this, admittedly flawed but convenient, organizational strategy. The unifying features of these viruses are that they replicate in the nucleus of the host cell, and each strictly relies on one or

another related function found in actively replicating animal cells for their successful propagation. Two other families of viruses, one infecting plants and the other bacteria, are included to demonstrate some of the strategies viruses can utilize to ensure that their DNA genomes are as physically compact as possible.

PAPOVAVIRUS REPLICATION

The term *papovavirus* stands for "*pa*pilloma, *po*lyoma, *va*cuolating" viruses. Actually, members of the group fall into two distinct families: the papillomaviruses and polyomaviruses. These two groups are similar regarding icosahedral capsids, circular genomes, and the ability to remain associated with the host for long periods, as well as their requirement to specifically alter cell growth in the host cell's response to neighboring cells for virus replication. They differ in genome size and in many details of host cell specificity. One unusual structural feature of the polyomavirus capsid is that although it is an icosahedron, the capsid subunits do not form hexon and penton arrays as is normal for such a structure (see Chapter 5). Rather, all 60 pentameric subunits are equivalent and can assemble in an asymmetrical fashion to form the capsid. This is shown in Fig. 17.1a; the actual properties of the capsid proteins that allow this unusual packing strategy are unknown to date.

Replication of SV40 virus – the model polyomavirus

The polyomaviruses have genomes of approximately 5000 base pairs. Capsids are made up of three proteins, usually called VP1, VP2, and VP3. Polyomaviruses can cause tumors in animals and can transform the growth properties of primary cells in culture, especially the cells from animals different from the virus's natural host (see Chapter 10). Polyomaviruses also stay persistently associated with the host, often with little evidence of extensive pathology or disease. Although these viruses kill the cells in which they replicate, this process is slow. In keeping with the requirement for extensive cellular function during replication, there is no virus-induced general shutoff of host function.

One widely studied polyomavirus is murine polyomavirus (Py), originally isolated from wild mice and named for its ability to cause many types of small tumors in some strains of newborn mice. Another widely studied polyomavirus is SV40 virus, which was originally named *simian vacuolating agent 40*. SV40 virus originally was found as a contaminant of African Green monkey kidney cells (AGMK) in which poliovirus was being grown for vaccine purposes. Early recipients of the Salk polio vaccine got a good dose of the virus, but no pathology has been ascribed to this, at least to the present time.

Whereas Rous sarcoma virus (a retrovirus) had been known to cause tumors in chickens since the early part of this century, the fact that its genome is RNA made understanding of its mechanism of oncogenesis out of the reach of molecular biologists working in the 1950s and 1960s. Indeed, major progress awaited the discovery of reverse transcriptase by Howard Temin and David Baltimore in 1970. In contrast, the fact that the DNA-containing SV40 and mouse polyomaviruses cause tumors in the laboratory provided a model for the study of the process that could be exploited with the techniques available at the time. The study of these viruses essentially launched the molecular biological study of carcinogenesis and eventually led to the discovery of tumor suppressor genes and their important role in regulating cell growth and division.

Its importance in fundamental research in oncogenesis, ease of manipulation in the laboratory, and convenient genome size have contributed to SV40 virus's status as, arguably, the most extensively studied of all DNA viruses. While Py and SV40 replication differ in some important features, the overall strategy is the same. Two human polyomaviruses, BK and JC, are known, and a third is suspected to exist but has not been rigorously identified. The BK and JC viruses are closely related to SV40, and are thought to be spread by the respiratory route. Primary infection occurs in children

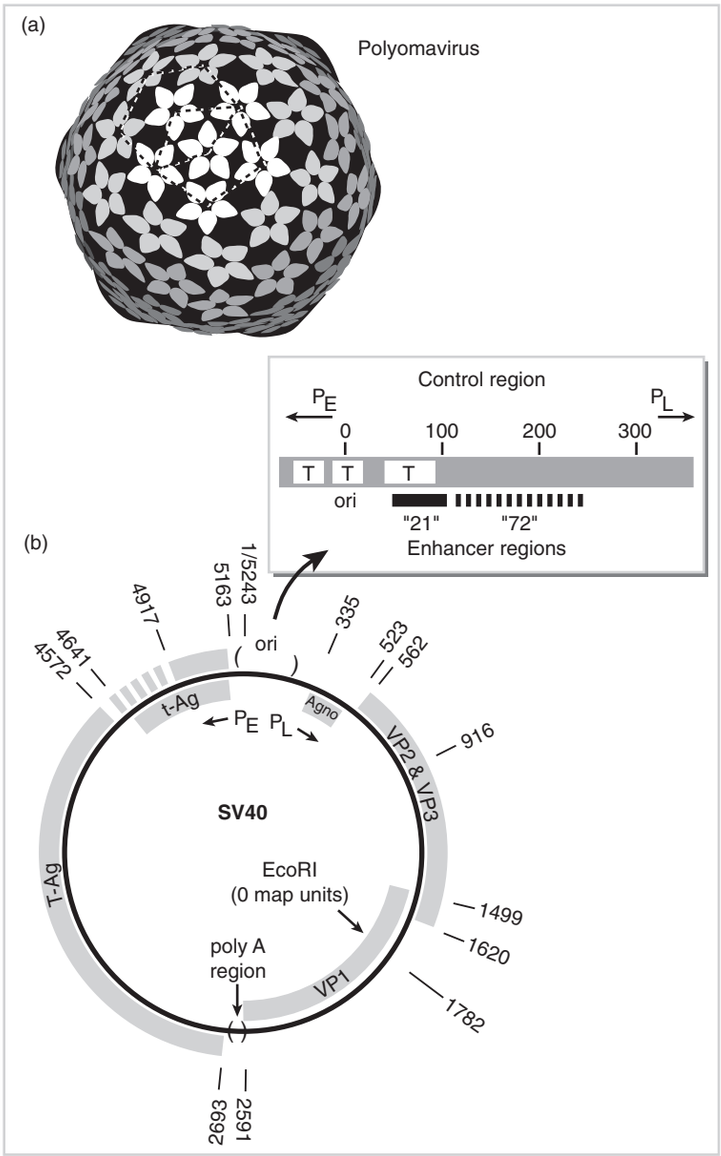


Fig. 17.1 Polyomavirus and the genetic and transcript map of SV40 virus. *a*. The 60 pentameric subunits of the capsid proteins are arranged in an unusual fashion so that the packaging of individual capsomers is not equivalent in all directions. The drawing is based on computer-enhanced analysis using the electron microscope and x-ray diffraction methods (see Chapter 5) published by Salunke et al. (*Cell* 1986;46:895–904). The 5243 bp dsDNA genome is condensed with host cell histones and packaged into the 45 nm diameter icosahedral capsid. *b*. The early and late promoters, origin of replication, and bidirectional cleavage/polyadenylation signals are shown along with the introns and exons of the early and late transcripts. A high-resolution schematic of the approximately 500 bp control region with the early and late promoters is also provided. Two early promoter enhancers, one containing the 21 bp repeats and the other containing the 72 bp repeats, are shown. The origin of replication (ori) is situated between the enhancers and the early promoter, and the three binding sites for large T antigen (T) are indicated. *c*. A higher-resolution schematic of the processing of early viral mRNAs. Splice sites, translational reading frames, and other features are indicated by sequence number. Details are described in the text. Note that the 3′ end of the pre-mRNA occurs just beyond the early polyadenylation site (2590) that is situated in the 3′ transcribed region of the late pre-mRNA. *d*. A higher-resolution schematic of the processing of late viral mRNAs. Splice sites, translational reading frames, and other features are indicated by sequence number. Details are described in the text. Note that the 3′ end of the pre-mRNA occurs just beyond the late polyadenylation site (2650) and is situated in the 3′ transcribed region of the early pre-mRNA. (T-Ag, large T antigen; t-Ag, small t antigen.)

with little obvious pathology. In the United States, most children are infected with BK virus by the age of 5 to 6 years, and the only signs of infection may be a mild respiratory illness. Infection with JC virus occurs somewhat later, with most children being infected between the ages of 10 and 14 years.

Resolution of infection is complete in children with normally functioning cell-mediated immunity. Despite resolution, the virus persists for the life of the individual—one primary site of persistence being the kidney from which BK virus can be periodically shed. In addition, JC virus can be recovered from brain biopsy specimens. While this persistence has no clinical manifestations in the healthy individual and is thought to be the result of viral genomes persisting in an inactive state in non-dividing, terminally differentiated cells, immunosuppression by HIV infection or prior to organ transplantation can lead to severe consequences. In immune-compromised individuals, JC

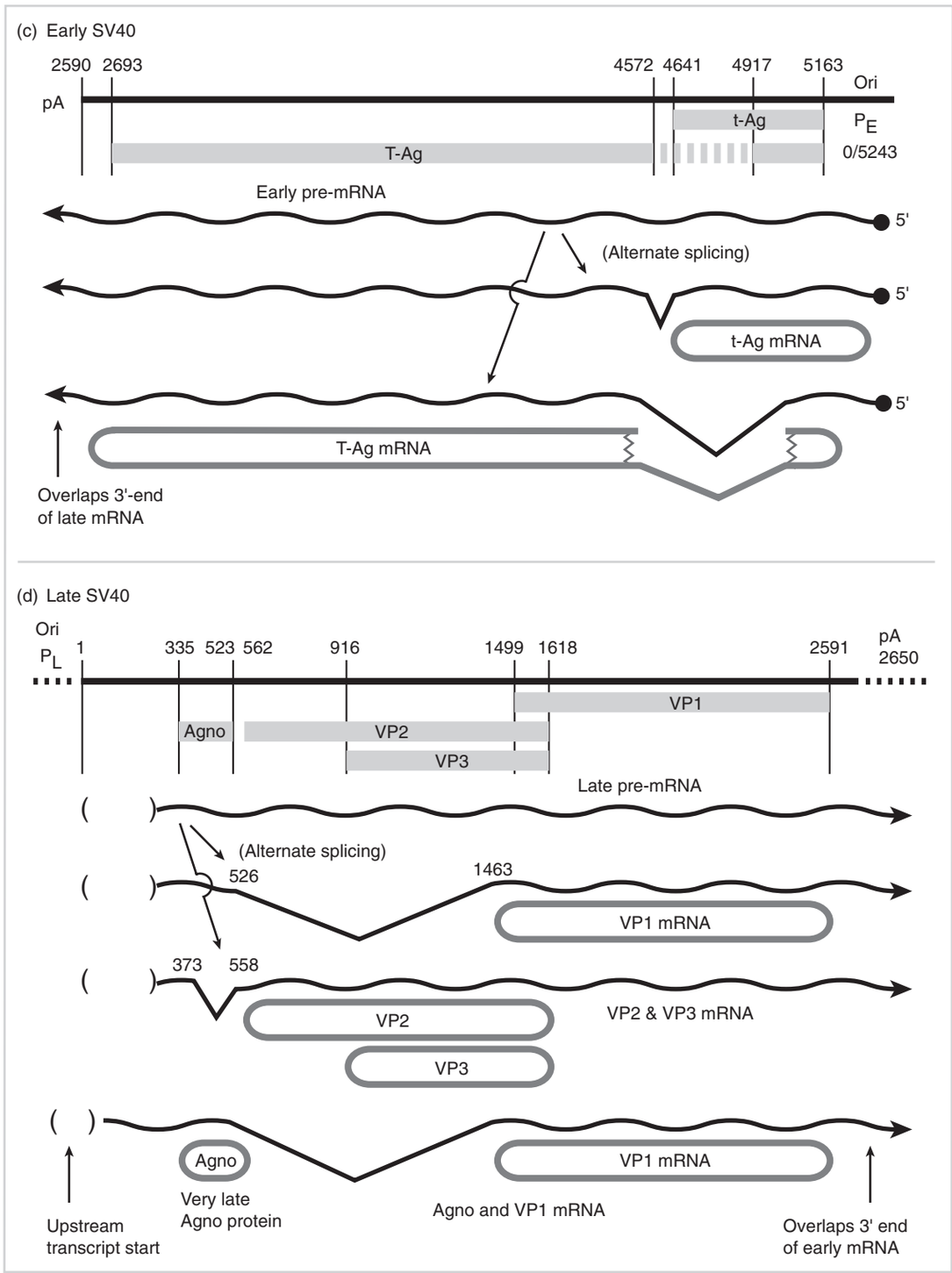


Fig. 17.1 *Continued*

virus is associated with a rare progressive destruction of neural tissue in the **CNS** (progressive multifocal leukoencephalopathy [PML]). This neuropathology is the result of the fact that transcription of the JC virus RNAs can take place in oligodendrocytes (but not other cells) in the adult brain, but just what aspect of immune-suppression such as that engendered by HIV infection reactivates this dormant virus is unknown. While not as firmly established, it is pretty certain that BK virus infections can lead to urinary tract pathologies in immune-compromised individuals.

The exact sequence of human JC virus isolated from individuals in various parts of the world varies enough to allow its use as a genetic population marker. Extensive studies on natural isolates show that individual variants are strongly associated with individual ethnic and racial population groups, and their movements throughout the world can be traced by the occurrence of specific virus variants. This means that the virus has been associated with the human population for an extremely long time, and that variants have arisen as populations have diverged.

The pattern of infection of young animals followed by virus persistence and shedding is quite characteristic of the infection of laboratory strains of mice with murine polyomavirus. One notable difference between the pathology of this virus and that of SV40, JC, and BK viruses, however, is that infections of suckling mice can lead to the formation of tumors, hence, the name polyomavirus. Genetic studies suggest that a major factor in the ability of the murine virus to cause tumors is the presence of specific endogenous retroviruses in the laboratory mouse strains, and while there is some suggestive evidence that human polyomaviruses can be associated with tumors, definitive evidence of causation is lacking.

The SV40 genome and genetic map

The SV40 virus genome contains 5243 base pairs, and its map showing essential features is displayed in Fig. 17.1b. The genome is organized into four functional regions, each of which is discussed separately.

The control region This region covers about 500 bases and consists of the origin of replication, the early promoter/enhancer, and the late promoter. The sequence elements in this region overlap to a considerable extent, but the bases specifically involved with each function can be located precisely on the genome. This has been done by making defined mutations in the sequence and analyzing their effects on viral genome replication and on expression of early and late genes.

The early promoter region contains a TATA box and enhancer regions (noted by 72-base and 21-base repeats). Surprisingly, the late promoter does not have a TATA box, and late mRNA initiates at a number of places within a 60 to 80 base region. The multiple start sites for late mRNA expressed from this "TATA-less" promoter was one of the early clues that the TATA box functions to assemble transcription complexes at a specific location in relation to mRNA initiation. It is not clear exactly what substitutes for the TATA box in the late promoter, but it is thought that transcription complexes can form relatively readily throughout the region.

The origin of replication (ori) is about 150 base pairs in extent and contains several elements with a sequence critically linked by "spacers" whose length but not specific sequence is important in function of the origin. The ori elements have some dyad symmetry; that is, sequence of the far-left region is repeated in the inverse sense in the far-right region. This symmetry is thought to have a role in allowing the DNA helix to "melt" at the origin, facilitating the entry of replication enzymes to begin rounds of DNA replication. The general process was described in Chapter 13.

The early transcription unit The SV40 genome's early region is shown in high resolution in Fig. 17.1c. It is transcribed into a single mRNA precursor that extends about halfway around the genome, and contains two open translational reading frames (ORFs). The single early pre-mRNA transcript can be spliced at one of two specific sites (i.e., the pre-mRNA is subject to alternative

splicing—see Chapter 13, especially Fig. 13.7). If a short intron is removed, an mRNA is generated that encodes a relatively small (approximately 20 kd) protein (**small t antigen**), which has a role in allowing the virus to replicate in certain cells.

A slightly smaller (and more abundant) mRNA is generated by splicing out a larger intron in the pre-mRNA. This removes a translation terminator that terminates the small-t-antigen ORF. The smaller (!) mRNA encodes the **large T antigen** (approximately 80 kd). The large T antigen has a number of functions, including the following:

1 activation of cellular DNA and RNA synthesis by binding to the cellular growth control gene products named Rb and p53. This binding stops these control proteins from keeping the cell contact inhibited. This function causes the infected cell to begin a round of DNA replication;
2 blockage of apoptosis that is normally induced in cells where p53 is inactivated at inappropriate times in the cell cycle;
3 binding to the SV40 ori to initiate viral DNA replication;
4 shutting off early viral transcription by binding to regions in and near the early promoter;
5 activating late transcription; and
6 playing a role in virion assembly.

The late transcription unit Late mRNA is expressed from a region extending around the other half of the genome from the late promoter; this is shown in Fig. 17.1d. The late region contains two large ORFs that encode the *three* capsid proteins. Part of the expression of late proteins, then, requires alternate splicing patterns, just as is seen with the generation of early mRNA. Splicing of a large intron from the primary late pre-mRNA transcript generates an mRNA that encodes the 36 kd major capsid protein (VP1). A small amount of mRNA is generated by splicing a small intron near the 5′ end of the mRNA, allowing the first ORF to be translated into the 35 kd VP2 protein.

The third capsid protein, VP3, is also expressed from the same mRNA encoding VP2 by utilization of an alternative translation initiation site. Ribosomes sometimes "miss" the first AUG of the 5′ ORF in the mRNA expressing VP2. When this happens, the ribosome initiates translation at an AUG in phase with the first one but downstream, producing the 23 kd VP3 protein. Thus, one mRNA encodes both VP2 and VP3, depending on where the ribosome starts translation. This "skipping" does not violate the general rule that a eukaryotic ribosome can only initiate a protein at the 5′ ORF, as the first AUG is not seen and thus is in the operational leader sequence of the mRNA.

There is a fourth late protein expressed from the late region, but this is only seen very late in infection. This basic protein, the "agnoprotein," is encoded in a short ORF upstream of that encoding VP2. Very late in infection, some mRNAs are produced by initiation of transcription farther upstream than at earlier times, and these can be translated into this protein. The role of this product is not fully understood, but it may be involved in allowing the virus to replicate in certain cells that are normally nonpermissive for viral replication. This is termed a *host-range* function.

The polyadenylation region About 180 degrees around the circular SV40 genome from the ori/promoter region lies a second *cis*-acting control region. It contains polyadenylation signals on both DNA strands so that transcripts transcribed from both the early and late regions terminate in this region. It is notable that the polyadenylation signals for the mRNAs are situated such that the early and late transcripts have a region of 3′ overlap. This can lead to the generation of dsRNA during the replication cycle, with attendant induction of interferon in infected cells (see Chapter 8).

Productive infection by SV40

Productive infection by SV40 in its normal host can be easily studied in cell culture using monkey kidney cells. The replication cycle is quite long, often taking 72 hours or more before cell lysis and release of new virus occur. One reason for this "leisurely" pace is that the virus is quite dependent on

continued cellular function during most of its replication. The virus replicates efficiently in cultured cells that are actively dividing either because they have not yet reached confluence or because the cells are growth transformed and not subject to contact inhibition of growth. (The basic growth properties of cultured cells are discussed in Chapter 10.)

While the virus replicates efficiently in replicating cells, it also is able to replicate well in cells that are under growth arrest by virtue of T-antigen expression early in infection. Manifestations of this ability provide many useful insights into the nature of the cell's ability to control and regulate its own DNA replication, and led to the discovery of the tumor suppressor genes p53 and Rb discussed in a following section.

Virus attachment and entry The replication cycle of SV40 is outlined in Fig. 17.2. Virions interact with a specific cellular receptor. This leads to receptor-mediated endocytosis, and the partially uncoated virion is transported in the endocytotic vesicle to the nucleus where viral DNA is released.

The association of viral genomic DNA with cellular chromosomal proteins is a common feature in the replication of all the animal viruses discussed in this chapter. SV40 DNA is associated with histones and other chromosomal proteins when it is packaged into the virion. It remains associated with chromosomal proteins upon its entry into the nucleus. This means, in effect, that viral DNA is actually presented to the cell as a small or "mini"-chromosome. Essentially then, the cell's transcriptional machinery recognizes the viral chromosome and promoters therein merely as cellular genes waiting for transcription.

Early gene expression Early gene expression results in formation of large quantities of large T antigen mRNA, and smaller amounts of small t antigen mRNA. The amounts of protein synthesized are roughly proportional to the amount of mRNA present. The small t antigen contains the same N-terminal amino acids as does large T antigen because of the way early pre-mRNA is spliced into the two early mRNAs, as shown in Fig. 17.1c. The splice-generating mRNA that encodes the T antigen removes a translation stop signal. In contrast, the splice in the t-antigen mRNA is beyond the ORF, and thus does not affect protein termination. Generation of two proteins with major or minor differences in function but with a shared portion of amino acid sequence is quite common with many viruses. It is very important in the expression of adenovirus proteins.

The role of T antigen in viral DNA replication and the early/late transcription switch As outlined in the preceding section, T antigen alters the host cell to allow it to replicate viral DNA. The T antigen also binds to the SV40 ori to allow DNA replication to begin, *and* to shut off synthesis of early mRNA. Each round of DNA replication requires T antigen to bind to the origin of DNA replication and initiate a round of DNA synthesis. DNA replication then proceeds via leading and lagging strand synthesis using cellular enzymes and proteins as described in Chapter 13. Since the SV40 genome is circular, there is no end problem, and the two daughter circles are separated by DNA cleavage and ligation at the end of each round of replication. This resolution of the interlinked supercoiled DNA molecules into individual genomes is mediated by cellular enzymes, notably topoisomerases and resolvases. The process is illustrated in Fig. 17.3. It is important to note that association of the daughter DNA genomes with cellular histones is not shown in the figure, but this association is necessary for the virus to be efficiently encapsidated.

While DNA replication proceeds, the relative rate of early mRNA synthesis declines owing to accumulation of increasing amounts of large T antigen in the cell, which represses synthesis of its own mRNA by binding at the ori and early promoter. While the relative amount of early mRNA declines in the cell at late times, its production never entirely ceases because there is always some template that has not yet bound large T antigen available for early mRNA expression.

At the same time that this versatile protein is modulating and suppressing its own synthesis, it activates transcription of late pre-mRNA from replicating DNA templates. Late transcripts have

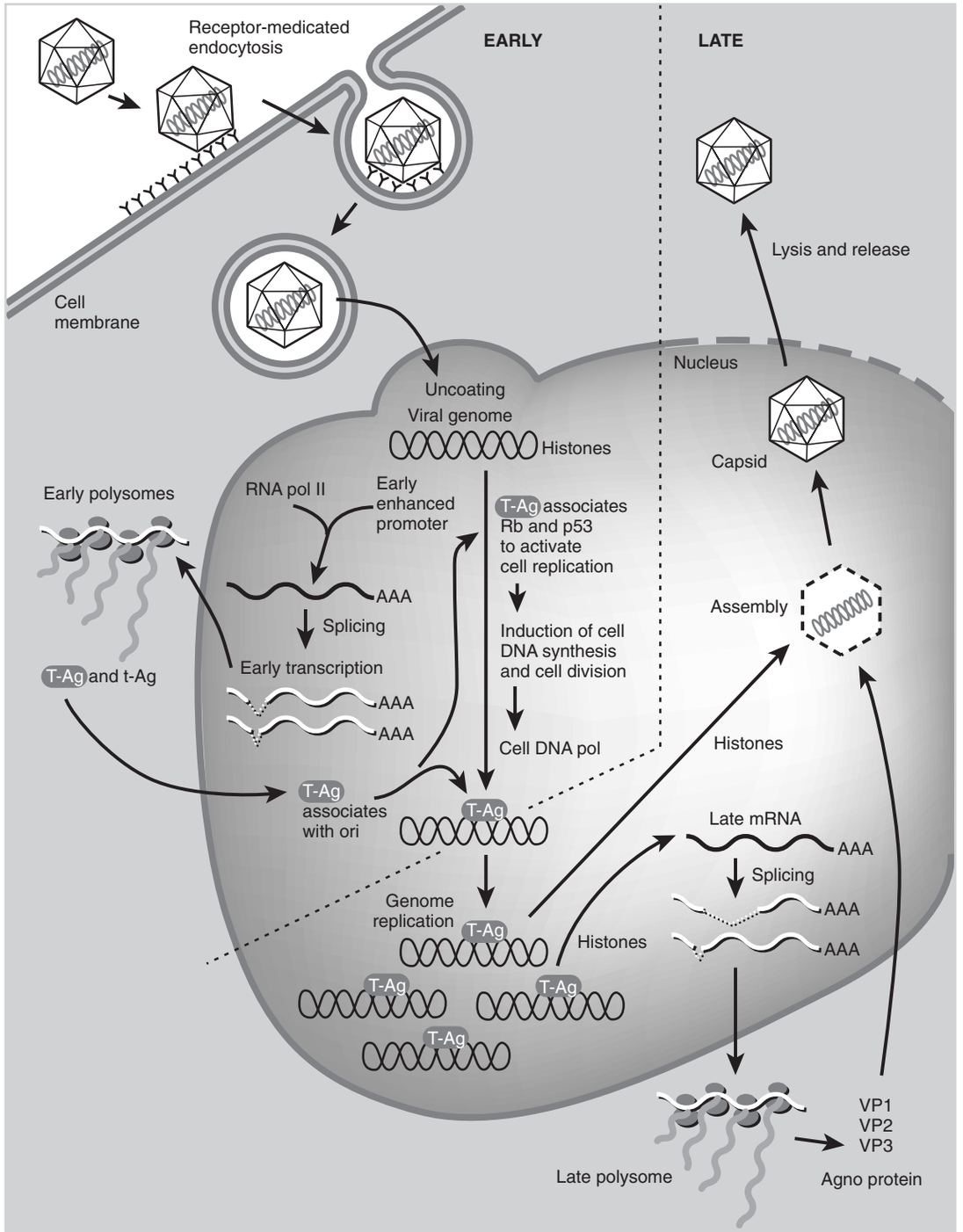


Fig. 17.2 The replication cycle of SV40 virus in a permissive cell. The replication is divided into two phases, early and late. During the early stages of infection, virus attaches and viral genomes with accompanying cellular histones are transported to the nucleus via receptor-mediated endocytosis. RNA polymerase II (pol II) recognizes the enhanced early promoter, leading to transcription of early pre-mRNA, which is processed into mRNAs encoding small t (t-Ag) and large T antigen (T-Ag). These mRNAs are translated into their corresponding proteins. Large T antigen migrates to the nucleus where it carries out a number of functions, including inactivation of the cellular growth control proteins p53 and Rb, and binding of the SV40 origin of DNA replication (ori). Viral DNA replication takes place by the action of cellular DNA replication enzymes, and each round of DNA replication requires large T antigen to bind to the ori.

As genomes are replicated, the late stage of infection begins. High levels of large T antigen suppress the expression of early pre-mRNA and stimulate expression of late pre-mRNA. This is processed into two late mRNAs; the smaller encodes both VP2 and VP3 while the larger encodes VP1. At very late times, some transcripts are expressed and can be translated into the small agnoprotein. Viral capsid proteins migrate to the nucleus where they assemble into capsids with newly synthesized viral DNA. Finally, progeny virus is released by cell lysis.

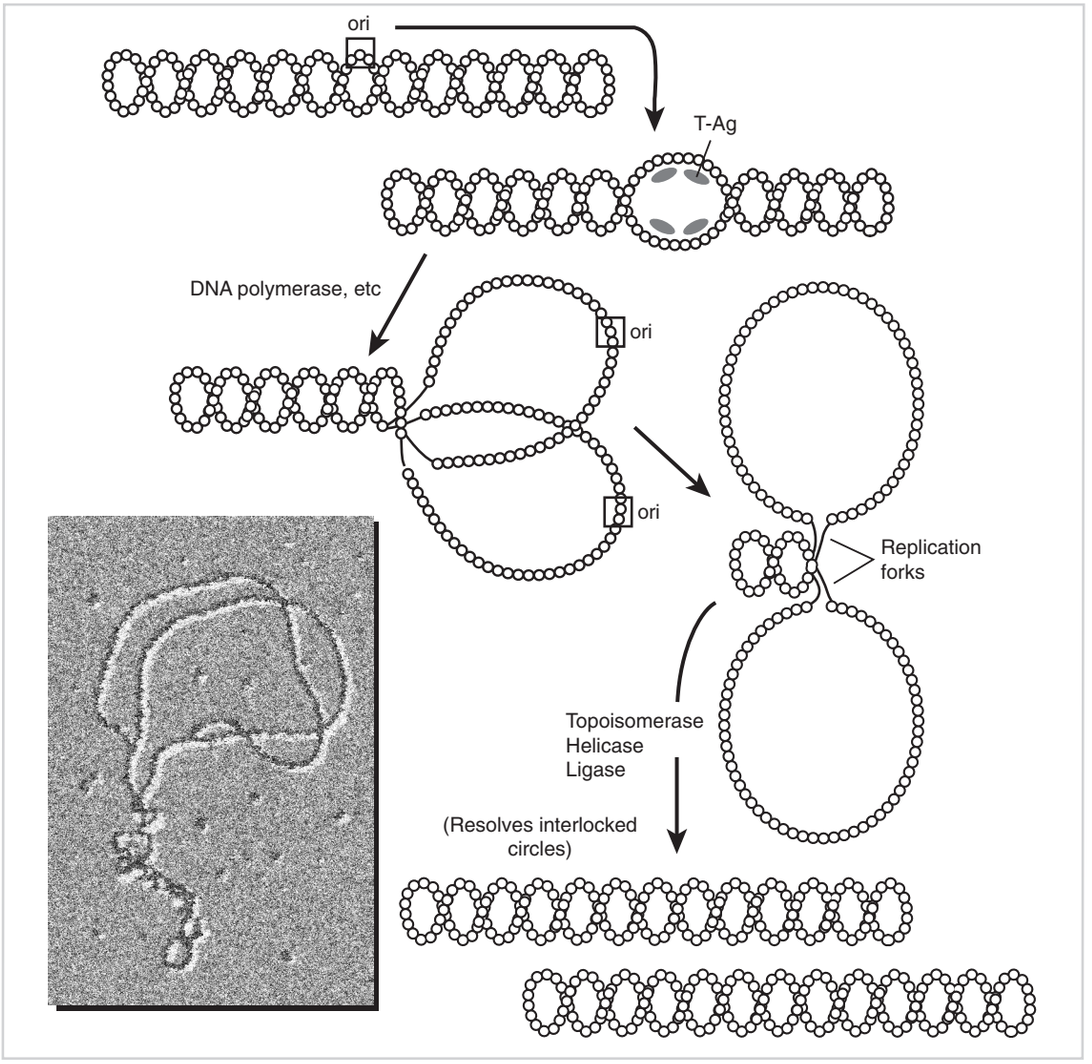


Fig. 17.3 The replication of SV40 DNA. The closed circular DNA has no end problem, unlike the replication of linear DNAs. Structures of the replication fork and growing points are essentially identical to those in replicating cellular DNA, and use cellular DNA replication enzymes and accessory proteins. Replication results in the formation of two covalently closed and interlinked daughter genomes that are nicked and religated into individual viral genomes by the action of cellular topoisomerase and other helix-modifying enzymes. (T-Ag, large T antigen; ori, origin of replication.)

heterogeneous 5′ ends, and as noted previously, very late in infection, the start of late mRNA transcription shifts to a point upstream of that previously used and the **agnogene** protein (the agnoprotein) can be encoded and translated from a novel subset of late mRNAs.

Abortive infection of cells nonpermissive for SV40 replication

Relatively early in the study of polyomavirus replication, infection of cells derived from a species other than the natural host of SV40 was discovered to result in an abortive infection where no virus was produced. Despite this, virus infection was shown to stimulate cellular DNA replication and cell division, and study of this phenomenon provided early important models for the study of car-

cinogenesis. While such abortive infections may be purely a laboratory phenomenon, the information derived from them provided an important foundation for understanding the pathogenesis of papovaviruses in their natural hosts and viral oncogenesis.

In rodent (and some other nonprimate) cells, SV40 virus can infect and stimulate cellular RNA and DNA synthesis by expressing the large T antigen. As noted, this viral protein inactivates at least two cellular tumor suppressor or growth control genes (p53 and Rb). The role of such **oncogenes** in controlling cell growth is briefly touched on in Chapter 10, and is discussed in more detail in Chapter 21.

The two proteins in question (p53 and Rb) have two basic functions. First, they mediate an active repression of cell division by binding to and thus inactivating cellular proteins required to initiate such division. Second, levels of the free proteins above a critical level induce apoptosis (programmed cell death, see Chapter 10) in the cells that escape repression and begin to divide.

As in the early phase of productive infection, in the first stages of infection of the nonpermissive cells, large T antigen binds to p53, thus displacing active replication-initiation proteins. But since the p53 is not free, there is no induction of apoptosis. These are the same steps that occur in the early stages of productive infection; however, viral DNA cannot be replicated in the nonpermissive cells. This failure is due to the inability of T antigen to interact effectively with one or more of its other cellular targets important in the early phases of infection. In this abortive infection, the cells in which T antigen is expressed do not die, but they replicate even while in contact with neighboring cells; this process is shown in Fig. 17.4.

The continued stimulation of cellular DNA replication by expression of viral T antigen can lead to continual cell replication (i.e., transformation). Stable transformation will require the viral genome to become stably associated with cellular DNA by *integration* of viral DNA into the cellular genome. Such viral DNA replicates every time the cell replicates, and thus keeps the cell transformed.

The integration of viral DNA into a host cell chromosome is not a function of T antigen or any other viral product. Indeed, most abortively infected cells will divide for a round or so until the viral DNA is lost, and then they will revert to their normal growth characteristics. This is sometimes termed **transitory (transient or abortive) transformation**.

The integration of viral DNA into the host cell is an entirely random recombinational event and occurs at sites where a few bases of the circular viral DNA can anneal to a few bases of chromosomal DNA. This must be followed by breakage and religation of the chromosome with the incorporated viral DNA. Obviously, this does not occur very frequently, but if a large number of cells are abortively infected with the polyomavirus in question and one or more integrate the viral chromosome and continue to express T antigen, those cells will form a focus of transformation.

Such a focus is a clump of transformed cells growing on the surface of a culture dish of contact inhibited cells. These foci can be counted and are subject to similar statistical analyses as are plaques formed by productive infection. Some typical foci of transformation are shown in Fig. 10.5.

The replication of papillomaviruses

Cell transformation by SV40 appears to be a laboratory phenomenon, and tumors caused by polyomaviruses can be thought of as dead-end artifacts of virus infection. In such infections, persistence appears to be due to the stability of histone-associated viral genomes in nonreplicating cells marked by occasional episodes of low level viral replication as a result of immune crisis or other events that lead to changes in the transcriptional environment of the host cell. In contrast, a related group of viruses, papillomaviruses, follow a natural replication scheme in their host that requires the formation of tumors, albeit benign ones, in their replication cycle. In this strategy of virus replication, persistence is a consequence of the continued replication of cells bearing viral genomes!

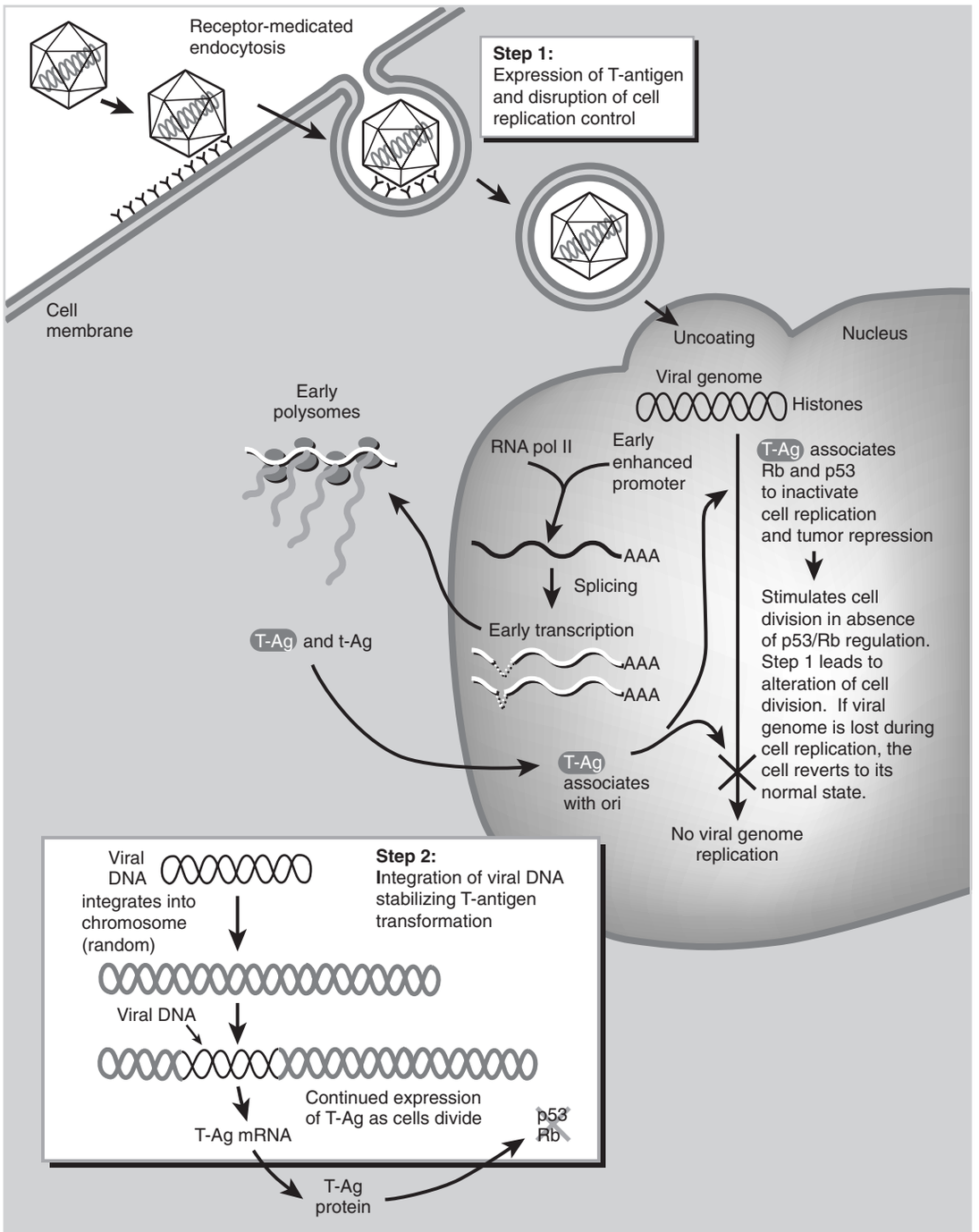


Fig. 17.4 Representation of the two steps in transformation of a nonpermissive cell by SV40. The infection begins as described in Fig. 17.2 and early mRNA is expressed into early proteins. The infection is abortive in that DNA replication and late gene expression cannot occur in the nonpermissive cell. Still, the large T antigen (T-Ag) is able to interfere with cellular growth control (tumor suppressor) proteins, leading to cell replication. Stable transformation requires a second step, the integration of the viral DNA. This is a random (stochastic) occurrence with SV40, and integration is random throughout the genome. A similar path is followed in the transformation of nonpermissive cells by other polyomaviruses. (t-Ag, small t antigen.)

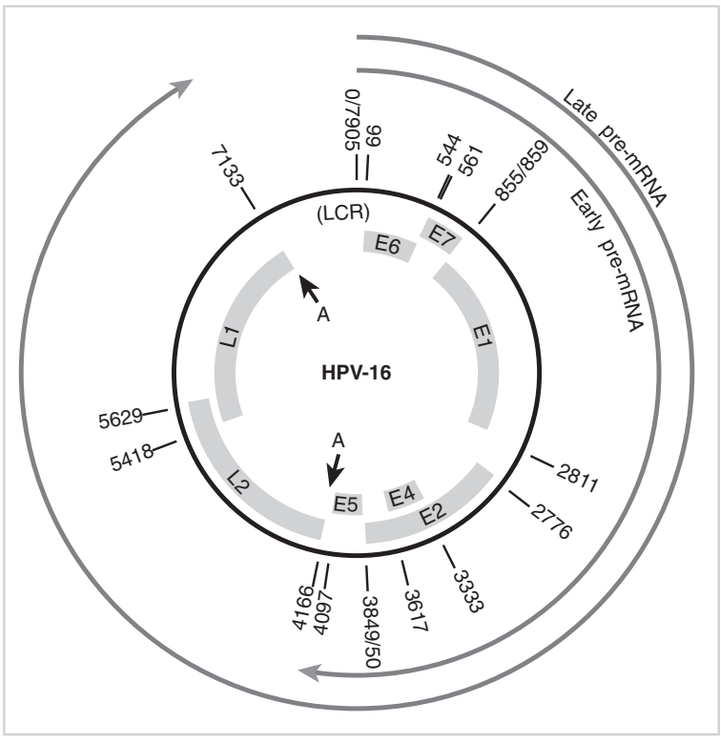


Fig. 17.5 The human papillomavirus (HPV)-16 genome. The 7 kbp circular genome contains a number of translational reading frames that are expressed from spliced mRNAs. Unlike the related polyomaviruses, papillomaviruses encode all proteins on the same DNA strand. The actual details of mRNA expression also appear to differ among different papillomaviruses. For example, HPV-16 has only one known promoter, which appears to control expression of both early and late transcripts. The locations of cleavage/polyadenylation signals for early and late transcripts are shown. All mRNAs appear to be derived by splicing of one or two pre-mRNAs. The characterization of transcripts has required heroic efforts of isolating small amounts of RNA from infected tissue, generating cDNA clones by use of reverse transcriptase and polymerase chain reaction, and then sequence analysis. This is necessary because many are present in very small amounts in tissue and the virus does not replicate in cultured cells. The transcripts shown are three of nine that have been fully characterized, and it can be expected that others are also expressed. The region marked "LCR" encodes both the constitutive (plasmid) origin of replication and an enhancer. Location of the vegetative origin of replication is not known. Specific details of papillomavirus replication are described in the text.

Papillomavirus replication combines some aspects of both the abortive and productive schemes just discussed. These viruses cause warts or papillomas, and there are many different serotypes, with most showing no antigenic cross-reactivity with each other. Infections with most papillomavirus serotypes are completely benign (although irritating or occasionally painful), but some can be spread venereally, leading to persistent genital infections, especially in females. Statistical analyses comparing the incidence of cervical carcinoma and the patterns of persistent infection by some of these papillomaviruses (including human papillomaviruses 16 and 18 [HPV-16 and HPV-18]) demonstrate a highly significant correlation despite the fact that only a small number of individuals actually get the disease. Thus, these viruses are clearly human cancer viruses.

The HPV-16 genome

The circular genome of HPV-16 is shown in Fig. 17.5. It is about 7200 base pairs long and is vaguely reminiscent of that of SV40 except there are many more early ORFs. Note that the region marked "LCR" corresponds to the promoter/origin region of SV40. Since the replication of papillomaviruses is difficult to study in cultured cells, a full characterization of the splicing patterns and transcripts expressed during infection has been and continues to be a very laborious effort. It requires analysis of DNA copies made of viral RNA using retrovirus reverse transcriptase, followed by cloning of the cDNA copies. Polymerase chain reaction (PCR) amplification of cDNA for direct sequence analysis also has been used. General methods for such analysis are covered in Chapters 11, 12, and 14.

Sequence analysis of the bovine papillomavirus genome and the transcripts expressed indicates that early and late transcripts are expressed from a single or limited number of early and late promoters as pre-mRNAs. While the extensive splicing of pre-mRNAs is reminiscent of infections with polyomaviruses, papillomaviruses differ in that early and late promoters appear to be found in several regions within the genome.

Virus replication and cytopathology

Formation of a wart by infection with papillomavirus is outlined in Fig. 17.6. It involves virus entering the basement epithelium of specific tissue (the skin in the case of warts). The virus expresses early genes that induce cells to replicate their DNA rather more frequently than would an uninfected epithelial cell. Thus, one set of early functions is analogous to those of SV40 T antigen. But in marked contrast to SV40 replication in permissive cells, papillomavirus DNA remains in the nucleus of the infected cell as an *episome* or "mini"-chromosome where it can replicate when cell DNA replicates, but it does not replicate to the high numbers seen in viral DNA replication of a productive infection.

Such cell-linked replication is often termed **plasmid-like replication**. It involves the interaction of cellular DNA replication with a specific ori in the virus (ori-P) that acts like an origin of cellular DNA replication and is subject to similar control. This ori-P can be the same as or different from the origin of productive DNA replication (lytic origin or ori-L).

As the cells are stimulated to divide, they differentiate, and as they differentiate, they change their function and begin to produce keratin and other terminally differentiated gene products. At some point in this terminal differentiation, some of these cells become fully permissive for high levels of viral DNA replication and late gene expression to generate capsid proteins. Such cells produce new virus while they die. Since this phenomenon is highly localized, and the virus infection normally just speeds up normal terminal differentiation of the epithelial cells, a benign wart is formed.

For HPV-16 and HPV-18, this growth enhancement is known to be a function of the actions of proteins encoded by the E5, E6, and E7 gene products that associate with and inactivate normal functions of the p53 and Rb proteins in a manner analogous to large T antigen activity in SV40. Presumably, chronic infection of cervical epithelium with either of these viruses can (rarely) generate a true cancer cell by further mutations of other control circuits in the cell. This oncogenic transformation is coincident with integration of papillomavirus DNA into cellular DNA, and it is speculated that oncogenesis involves a process similar to the transformation stabilization seen in abortive SV40 infection of the appropriate nonpermissive cell.

In such a transformed cell, no virus is produced, so formation of the cancer can be looked at as a dead-end accident induced by the continued stimulation of cell division caused by the virus's persistent infection. As these transformed cells continue to divide, they accumulate mutations that eventually allow them to spread to and invade other tissues, and form disseminated tumors (**metastasis**). In the case of benign warts in the skin and elsewhere, either inactivation of the p53 and Rb proteins is not so profound, or the stimulated cells are so close to death in their terminally differentiated state that they cannot become cancerous.

THE REPLICATION OF ADENOVIRUSES

The adenoviruses comprise a large group of icosahedral, nonenveloped viruses of humans and other mammals. In humans, they generally are associated with mild flu-like respiratory diseases, but some serotypes also are associated with gastrointestinal upsets. While adenoviruses are not at all closely related to the papovaviruses, they share with them a long replication cycle due to the need to stimulate and utilize many cellular functions to carry out virus replication. They also share the ability to transform cells in the laboratory via abortive infection. Also like the papovaviruses, adenovirus replication involves extensive splicing of a limited number of pre-mRNAs. The usage of alternative splicing sites leads to the expression of a nearly bewildering number of partially overlapping mRNAs encoding related proteins.

Despite these similarities to papovaviruses, there are striking differences in the details of replication and in the organization and replication of the viral genome. The relatively mild course of adenovirus infection, and some convenient properties in manipulation of the virus, make it an attractive candidate for use as a therapeutic agent (see Chapter 22).

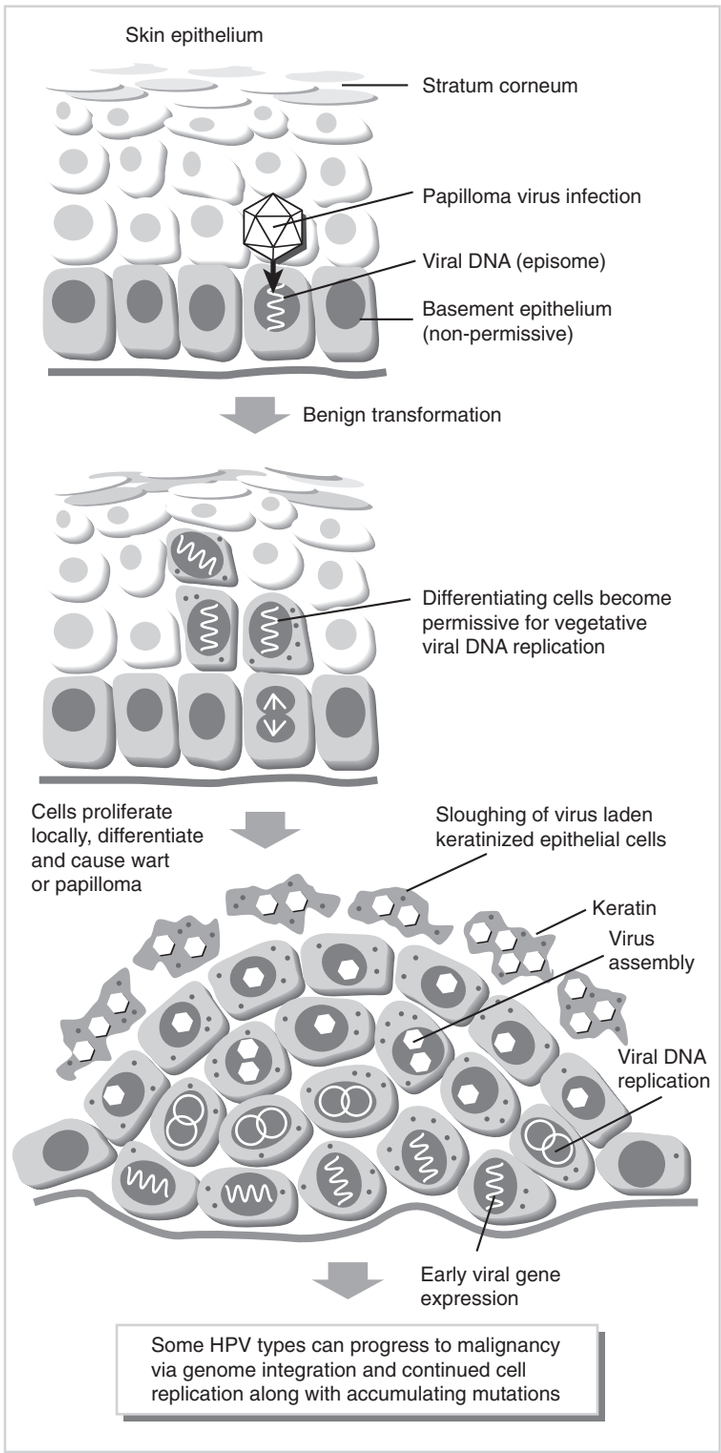


Fig. 17.6 The formation of a wart by cell proliferation caused by infection of basement epithelial cells with human papillomavirus (HPV). Early gene expression leads to stimulation of cell division and terminal differentiation. This results in late gene expression and virus replication in a terminally differentiated, dying cell, which produces large quantities of keratin.

Physical properties of adenovirus

Capsid structure

Adenoviruses have complex icosahedral capsids whose proteins are not present in equimolar amounts (see Fig. 11.5), with projecting spikes or *fibers* at the 12 vertices (pentons). The viral genome is encapsidated with core protein that acts a bit like histone to provide a chromatin-like structure that is condensed in the interior of the nucleocapsid.

The adenovirus genome

The genome of adenoviruses is linear with specific viral protein (*terminal protein*) at the 5′ ends. The genome is about 30,000 base pairs, and the sequence at the genome's end (100–150 base pairs, depending on virus serotype) is inversely repeated at the other end. This is the ori for viral DNA.

The genome map with location of the many transcripts expressed during infection is shown in Fig. 17.7. The genome is divided into 100 map units; therefore, each map unit is 300 base pairs. Transcript location is complicated by complex splicing patterns and the presence of a number of promoters. There are four early transcription "units" termed E1 through E4; each of these contains at least one promoter and polyadenylation signal. A single late promoter produces five "families" of late mRNAs, and there is also an unusual RNA called "VA" that is transcribed by the action of host cell RNA polymerase III (pol III).

The adenovirus replication cycle

Early events

Adenovirus enters the cell via receptor-mediated endocytosis in a manner analogous to that of papillomaviruses. Cellular receptors interact with the virion fiber proteins to initiate infection. Adenovirus DNA with a specific terminal protein bound to each 5′ end is released into the nucleus where it associates with cellular histones. In order to initiate gene expression, adenoviruses must stimulate the infected cell to transcribe and replicate its genes. This is accomplished by expression of the spliced mRNAs encoding the immediate-early (or "pre-early") gene E1A and E1B protein "families." The promoters for these are enhanced and can act in the cell in the absence of any viral modification (like the SV40 early promoter). The E1A gene products block the ability of the p53 and Rb growth suppressor genes to suppress cell division, while one or several E1B proteins inhibit apoptosis in the stimulated cell. Thus, these two proteins work in concert in a manner similar to that of polyomavirus large T antigen.

Stimulation of the infected cell's transcriptional machinery leads to expression of the four early pre-mRNAs that are spliced in various ways to produce early proteins, including a DNA polymerase protein (140 kd pol), a terminal protein, and a 72 kd DNA-binding protein (DBP). The latter shuts off most early promoters, but the E2 region is not shut off because a second promoter becomes active at times when 72 kd DBP is at high levels. Interestingly, the major late promoter is "on" early in infection, but only the L1 region is expressed as mRNA because all transcripts are terminated at the polyadenylation signal at 40 map units. This termination is due to the inhibition of splicing downstream of the L1 region through binding of cellular splice factors. Further, late transcripts downstream of L1 are not transported from the nucleus.

Adenovirus DNA replication

Adenovirus genome replication takes place via an unusual mechanism that involves formation of ssDNA as intermediates; the process is shown in Fig. 17.8. Adenovirus DNA replication

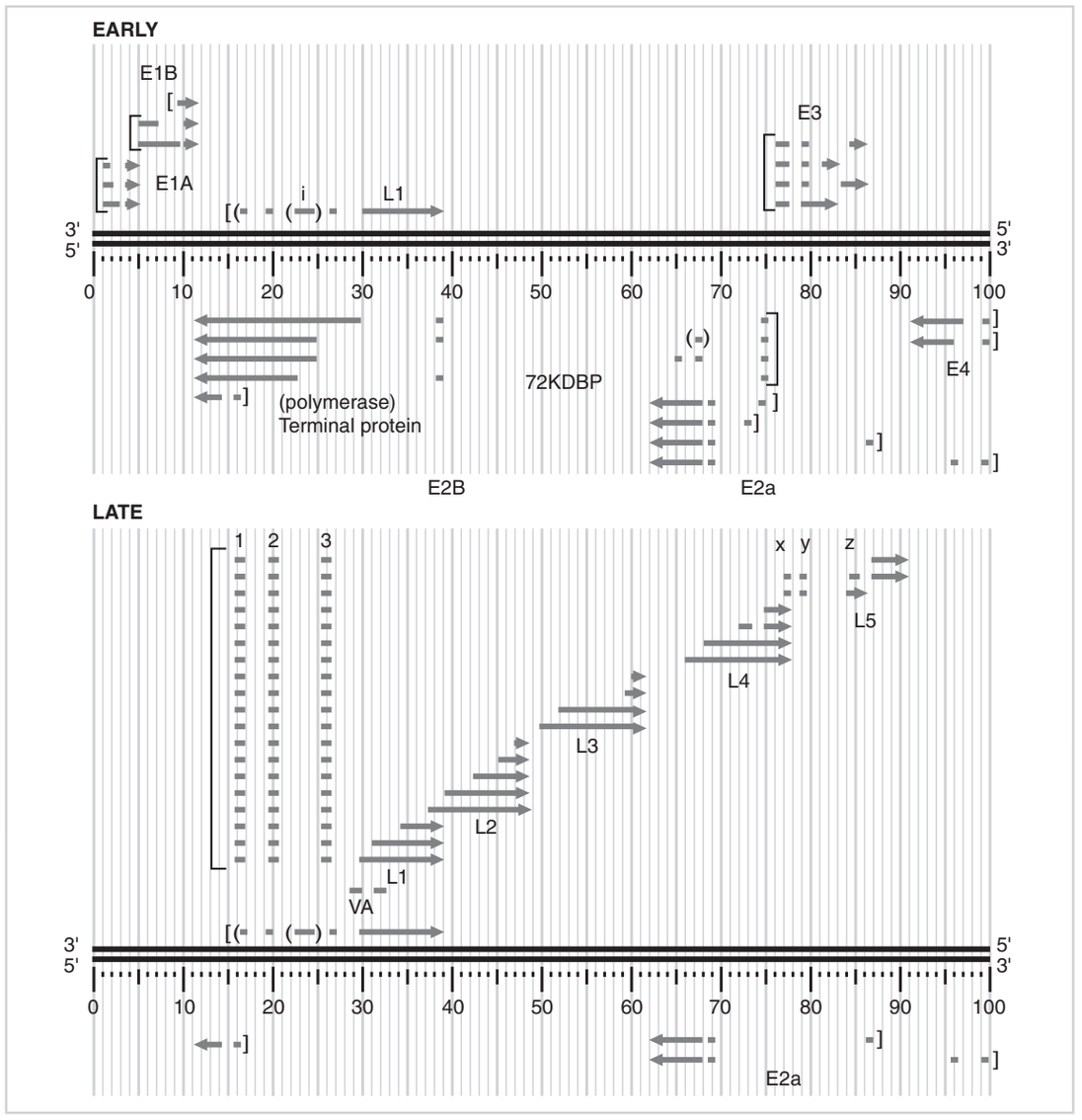


Fig. 17.7 The genetic and transcription map of the 30 kbp adenovirus genome. There are three kinetic classes of transcripts. The E1 transcripts are controlled by enhanced promoters and require no modification of the host cell because some functions of their expression are similar to those of T antigen in SV40 virus replication. These functions include stimulating cellular transcriptional activity and cell replication. Early in infection, only early transcripts are expressed. These include mRNAs encoding viral DNA polymerase and terminal proteins. There are a number of early promoters and transcription units. The E2 transcription unit also has a 72 kd DNA-binding protein (72KDBP) that shuts off early transcription. Two primary transcripts, E2A and E2B, are expressed from the same E2 promoter. The mRNA for the DNA-binding protein continues to be expressed late because there is a second promoter upstream of the E2 promoter that is not shut off by the DNA-binding protein. The major late promoter at map position 15 is always "on," but polyadenylation and splicing patterns change markedly as infection proceeds. Late in infection, the late transcription unit extends to one of five polyadenylation signals and differential splicing results in generation of a myriad of late mRNAs encoding structural proteins as well as proteins involved in host cell modification and virus maturation.

begins at either or both ends of the DNA and uses as a primer an 80 kd precursor of the 50 kd viral genome-bound terminal protein. The large priming terminal protein is proteolytically cleaved to the smaller terminal protein found in capsid-associated genomes during packaging. This is the only known instance where DNA replication initiates without a short RNA primer. However, the terminal protein does contain a covalently bound cytosine residue from which DNA replication

proceeds. Note that replication utilizes the adenovirus-encoded DNA polymerase and is continuous—there are no short Okazaki fragments seen. The process can liberate the other strand as ssDNA, which can become circular by association of inverted repeat sequences at the end, and replication proceeds. Thus, adenovirus DNA replication can proceed via two routes shown in Fig. 17.8. If DNA synthesis initiates at both ends of the genome about the same time, type I replication occurs. If only one end of the genome is used to initiate a round of DNA synthesis, then type II replication occurs.

Late gene expression

With the increase in levels of early 72 kd DBP, much early gene expression shuts off (see Fig. 12.14). At the same time, E4 protein interferes with the inhibition of splicing downstream of the L1 region, effectively this results in *polyadenylation site usage* changes so that transcription from the major late promoter generates transcripts covering as much as 24,000 bases. Differential polyadenylation and splicing generate the five families of late mRNAs that are translated into the structural proteins that will make up the capsids. Other late proteins alter aspects of cellular structure and metabolism to ensure efficient virus assembly and release. In addition to altering splicing patterns, some species of E4 protein actively mediate the transport of late mRNA from the nucleus to the cytoplasm.

VA transcription and cytopathology

The complex interaction between human adenovirus infection and the host cell requires that the cell remain functional for a long period following infection. This precludes extensive virus-induced shutoff of host cell function; hence, virus-induced cytopathology is slow and cell death takes a long time. During this period, the cell can mount defenses against viral gene expression such as the induction of interferons, cellular gene products that can render neighboring cells resistant to virus infection (see Chapter 8). The human virus gets around this problem by synthesis of **VA RNA**, which is a short, highly structured RNA molecule that interferes with the cell's ability to produce interferon. This VA RNA is expressed via cellular RNA pol III, which is the same polymerase used to transcribe cellular amino acid transfer RNAs (tRNAs). Interestingly, while the human Epstein-Barr herpesvirus expresses an analogous transcript (see Chapter 18), suggesting that this is an important feature in virus-mediated immune evasion, a number of adenoviruses of domesticated animals do not express a homologue to VA RNA.

A second aspect of the interaction between adenovirus and the host is reminiscent of papillomavirus replication. Adenovirus remains associated with the host for long periods of time as a persistent infection, especially in the epithelium of the adenoidal tissue and the lungs. The virus infects basement cells, but initiates DNA replication and viral assembly only in terminally differentiated cells. The virus actually induces an acceleration of apoptosis of these differentiated cells. One apparent advantage of eliminating dying infected cells is that more room is made available for the differentiation and growth of basement cells. This provides a ready and continuing source of cells in which the virus can initiate new rounds of replication. This stimulation of apoptosis presumably occurs because the relative levels of E1A and E1B are different in critical cells as compared to cells in which apoptosis is blocked by the latter viral protein.

Transformation of nonpermissive cells by adenovirus

As with SV40, infection of nonpermissive cells by at least some adenovirus types can lead to cell transformation and tumor formation. While there is currently no evidence for any involvement of adenovirus infection in human cancers, transformation seems to be accomplished by mechanisms very similar to those outlined for papovaviruses. Indeed, under some conditions, adenovirus gene products can substitute for early papovavirus gene products in mixed infections.

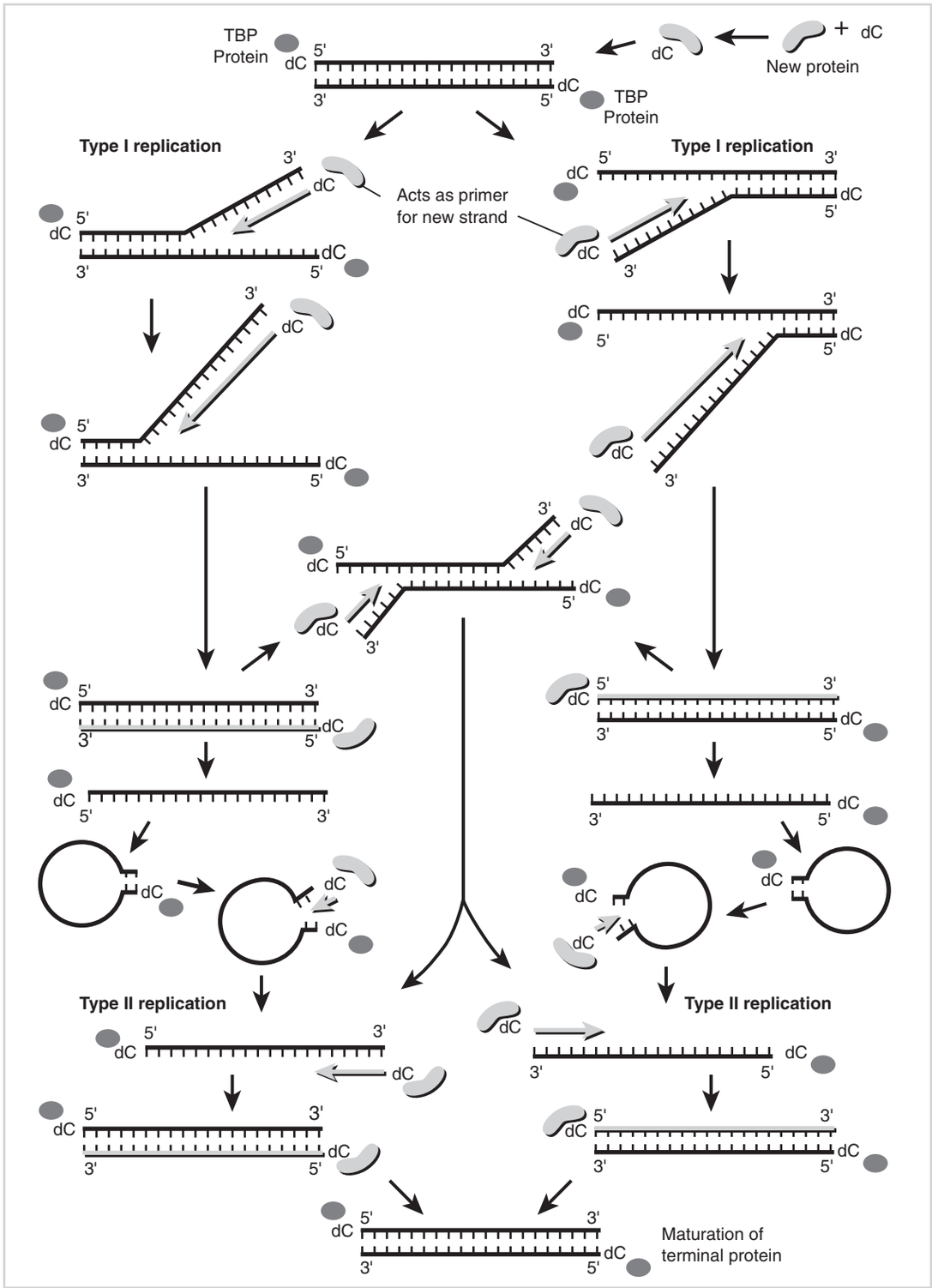


Fig. 17.8 Adenovirus DNA replication. The 5′ ends of the viral genome have 50 kd terminal proteins bound to them. Adenovirus does not have discontinuous strand synthesis, and exhibits other features that are at variance with the general scheme for viral DNA replication outlined in Chapter 14. Of major interest is the fact that there is no discontinuous strand synthesis. The process is marked by the accumulation of a large amount of single-stranded DNA (unusual in eukaryotic DNA replication). Further, the initial priming event requires the first nucleotide of the new DNA strand being covalently bound to the 80 kd precursor of the 50 kd terminal protein. Following complete second strand synthesis, the precursor end proteins are proteolytically cleaved to form the mature terminal proteins. (TBP, Precursor to terminal binding protein.)

REPLICATION OF SOME SINGLE-STRANDED DNA VIRUSES

With many plant viruses, and some animal and bacterial ones, a relatively small capsid size provides some advantages. With plant viruses, this advantage is tied to the limitations of virus capsid size that can "fit" in pores of the plant's cell wall. The advantages for animal and bacterial viruses are less clear, but must exist.

Replication of parvoviruses

The parvoviruses are very small, nonenveloped, icosahedral viruses. Two of the three known groups infect warm-blooded animals while the third group has members that infect insects. The parvovirus capsid diameter is 26 to 30 nm, significantly smaller than the polyomaviruses even though the viral genome is approximately 5 kb long. The virus is able to package the genome into such a small virion because the virus encodes only a single DNA strand. Interestingly, many parvoviruses can package the DNA strand of sense either opposite to mRNA or equivalent to mRNA in equal or nearly equal numbers. This means that the packaging signals utilized by the virus to encapsidate the genome must occur on both strands—this is probably through the interaction of the unique end structures of both strands with capsid proteins.

The genome of adeno-associated virus, a typical parvovirus, is shown in Fig. 17.9. It encodes two protein translational reading frames that are expressed by a variety of transcripts. The first reading frame encodes nonstructural protein involved in replication, and the second encodes the capsid protein. The genome ends contain 120 to 300 bases of inverse repeated sequences so that they can form hairpin loops in solution and in the infected cell's nucleus. These terminal hairpins serve as primers for initiation of DNA replication, and since they are repeated at the ends of both (+) and (–) sense DNA strands, both can serve as templates for DNA replication.

Parvovirus replication is absolutely dependent on the host cell undergoing DNA replication. Thus, the virus can only replicate in actively replicating cells. Despite this, and unlike papovaviruses and adenoviruses, parvovirus has no ability to stimulate cell division via the action of a viral-encoded protein. This inability results in a very tight restriction of virus replication in the host's dividing cells, especially cells of the immune system. This can be devastating to young animals and parvovirus infection of dogs is a major problem in kennels. Parvovirus infection can also be very destructive to actively growing cells in adult animals. For example, **feline panleukopenia**, a disease characterized by destruction of the immune system, is a significant pathogen of domestic cats, and is caused by a parvovirus.

Upon infection, the ssDNA is converted into full dsDNA by cellular DNA repair enzymes following its entry into the nucleus. The double-stranded viral DNA template is transcribed into a number of 3′-coterminal transcripts from one of three viral promoters just 5′ of the transcript starts. Some of these transcripts are spliced, so each translational reading frame is translated into several proteins of related sequence. As noted previously, viral genome replication can only take place in cells in which there is active cellular DNA replication (i.e., in the S phase of cell division). The viral replication enzyme is involved in cleavage of the covalently closed replicating viral DNA into single-stranded genomic DNA and has no polymerase activity.

Dependovirus DNA integrates in a specific site in the host cell genome

One major group of parvoviruses, the **dependoviruses**, is usually found associated with active infections of adenoviruses and occasionally, with herpesviruses. The human parvovirus, adeno-associated virus (AAV), is a well-characterized example. It is entirely dependent on coinfection with adenoviruses or herpesviruses in humans. While the dependoviruses can be grown in culture in

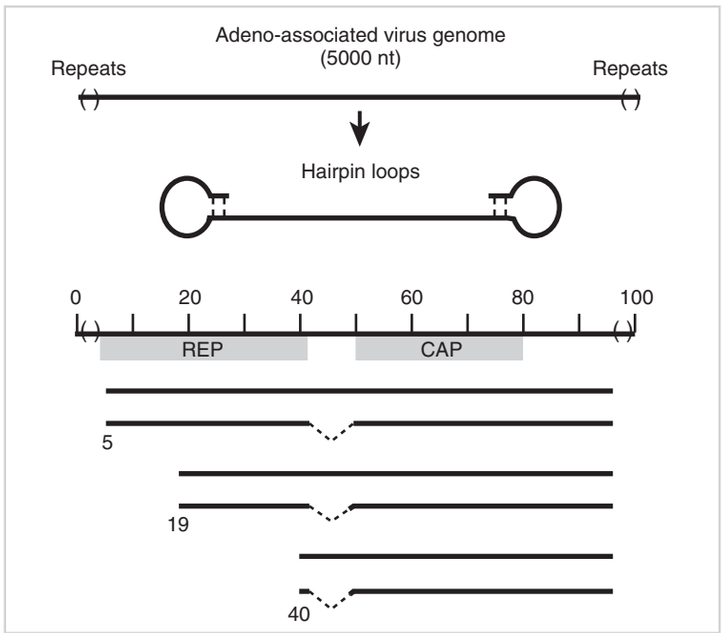


Fig. 17.9 The 5000 nucleotide (nt) linear genome of adeno-associated virus (AAV). This ssDNA has repeated sequences on both ends that allow it to form a "hairpin" structure. This serves as the template for conversion into dsDNA by cellular enzymes. Cellular enzymes also mediate replication of the viral genome. Three families of coterminal mRNAs are expressed from the three AAV promoters; the genome encodes replication proteins and a capsid protein but depends on cell replication for its ability to replicate its genome. This cellular replication is induced by a helper virus such as adenovirus in the animal, but the virus can replicate in cultures of some actively replicating cells. Other groups of parvoviruses, such as minute virus of mice (MVM), are able to replicate in some actively replicating cells of their natural host.

fetal cells or following proper chemical stimulation of some adult host cells, they depend on the adenovirus or herpesvirus helpers to stimulate the cell in such a way that they can divide. Thus, like viroids, these viruses are parasitic on other viruses.

The dependence on a helper virus might be expected to be a great impediment to virus replication for AAV, but this is overcome in part by its ability to integrate into chromosome 19 of the host when it infects a cell in the absence of the helper. The integrated viral DNA allows AAV to remain latent in host tissue for long periods of time, but to "reactivate" if and when that cell is infected with a virus that can act as a helper.

Integration takes place at short stretches of homologous sequences within a region of several hundred bases in the host chromosome. While it allows the viral genome to remain associated with the host for long periods, integrated viral DNA serves as a biological "time bomb" — ready to replicate and kill the cell when it is infected with the appropriate helper. Since the replication of AAV interferes with the efficiency of replication of the helper virus, it may be that this process has the ultimate effect of limiting infection of the helper, thus providing some benefit to the host!

Parvoviruses have potentially exploitable therapeutic applications

The strict requirement for actively replicating cells, and the competition between AAV and adenovirus and herpesvirus infections, suggest that such viruses might be exploitable as antiviral or anticancer agents. Laboratory studies showed this to be feasible. For example, breeds of laboratory mice have high occurrences of certain tumors. Infection of young mice with minute virus of mice (MVM), a murine parvovirus, results in a significant increase in the animal's life span and fewer occurrences of tumors at young ages! It should be clear, however, that an effective application of such a result to human cancers is not a straightforward undertaking.

Another potential use for parvoviruses stems from their ability to integrate in a specific site in the chromosome. This integration is mediated by the hairpin loop ends of the viral genome, and may be useful in designing viral vectors for delivering genes into cells.

DNA viruses infecting vascular plants

While DNA viruses infecting vascular (i.e., "higher") plants might be expected to display genetic variability equivalent to that seen within animal and bacterial viruses, they do not appear to do so. The reason for this is that plant viruses must traverse a relatively thick and dense cell wall to approach and breach the plant cell's plasma membrane. Although at least one algal virus can insert its genome like bacterial viruses inject genomes, apparently the dimensions of the vascular plant's cell wall preclude this accommodation. This results in the viruses of higher plants having a strict limitation on the size of their genomes, and although such viruses are not fully characterized, they may require a significant number of cellular functions for replication.

Geminiviruses

One group of viruses that infects plants have single-stranded, covalently closed circular DNA genomes and are packaged into unusual twinned capsids. These "twin" capsid structures give the group its genus name, *Geminivirus* (from the Latin word *geminae*, for "twins"). The number of genes encoded and their arrangement on the genome distinguishes the three major groups of these viruses. Two of the groups encapsidate the same genome in both of the twinned capsids; thus, they have a monopartite genome. In contrast, the third group contains a bipartite genome, and the two different genomic segments are packaged separately in each of the capsid halves. Rather astonishingly, one geminivirus isolated from bananas contains capsids bearing eight distinct genomic segments. How the virus accomplishes the rather remarkable feat of packaging different genomic segments into different subcapsids is an open question.

Representatives of geminiviruses include maize streak virus (a monopartite genome) and tomato golden mosaic virus (a bipartite genome). The genome (2.7–3.0 kb) organization of the geminiviruses has ORFs oriented in both directions around the circle, much like the papovaviruses. Since geminiviruses are single stranded, the input genome strand must be converted into dsDNA following infection, in order to obtain the appropriate template for transcription of mRNA whose translational reading frames are antisense to the virion DNA.

The geminiviruses are transmitted from plant to plant by leafhoppers or white flies. The virus can remain in the insect for long periods, but unlike the classic arboviruses, geminiviruses do not replicate in their insect vectors. Replication and transcription of these viruses take place in the nuclei of infected plants, using a rolling circle scheme. The exact function of the gene products predicted from sequence analysis has not been determined. Therefore, it is not yet possible to say which of the viral proteins might be specifically involved in this DNA replication.

The single-stranded DNA bacteriophage ΦX174 packages its genes very compactly

The gene packaging of bacteriophage ΦX174 suggests that genomic size compression offers distinct advantages in the prokaryotic world, also. This icosahedral virus has a structure very similar to that of adenovirus, but with shorter fibers. It contains a circular ssDNA genome approximately 3.4 kb long. Upon infection of a bacterial cell, the ssDNA genome is converted into dsDNA. This has been termed the *replicative intermediate* or *replicative form (RF)*, but is quite unlike the complex ribonucleoprotein complex with this name seen in the replication of ssRNA viruses.

Viral-encoded mRNA expression, protein synthesis, and genome replication occur following patterns that are generally simple examples of the more complex replication programs of DNA-containing bacteriophages described in Chapter 19. A striking demonstration of the extent this virus has gone to compress its genome comes from examination of its genetic map, shown in Fig. 17.10. The virus encodes nine distinct genes, but where one might expect about 200 to 300 bases

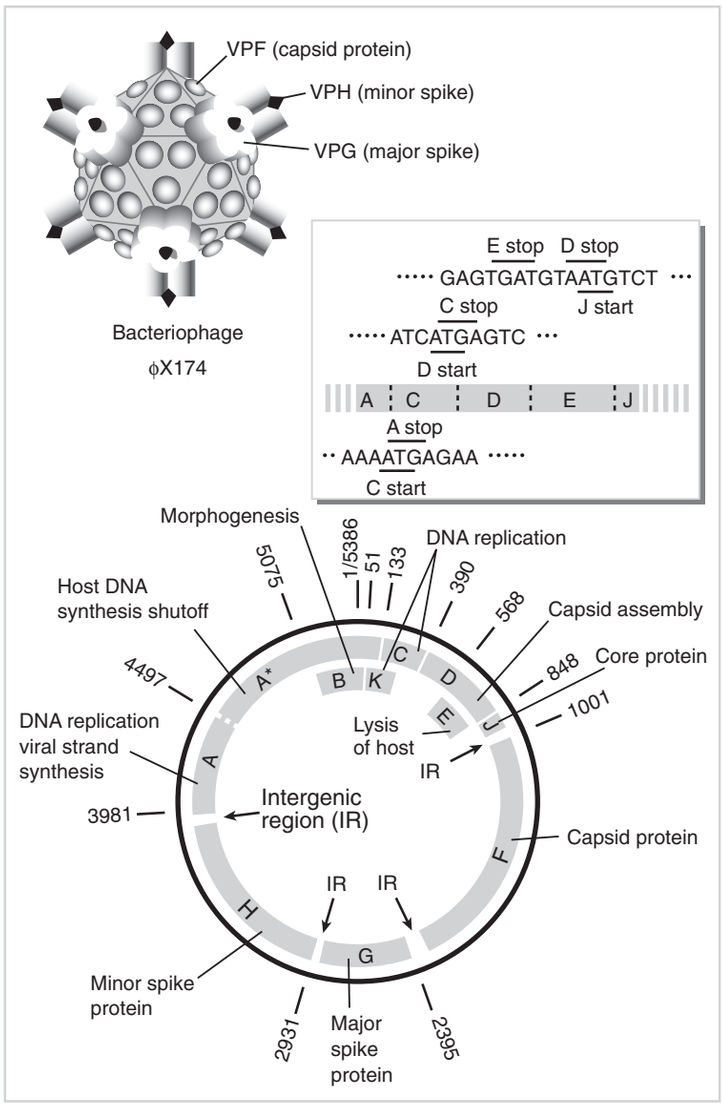


Fig. 17.10 The capsid structure and compressed genome of bacteriophage ΦX174. The capsid is made up of three proteins: major capsid, major spike, and minor spike. In all, a total of 10 genes are compressed into 3.4 kb of ssDNA! This is accomplished by very short intergenic regions, and two completely overlapping genes. The functions of the proteins encoded by these genes are listed.

of the DNA sequence to contain nonprotein information, only 36 bases (<1%) of the genome are free of translational reading frames. This arrangement means that all transcriptional control sequences are contained within translational reading frames.

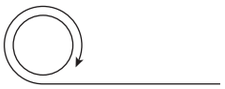
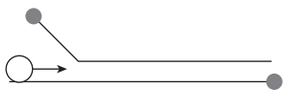
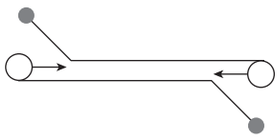
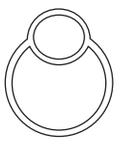
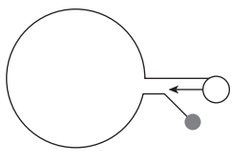
The translation start and stop signals of individual neighboring ORFs often overlap. Further, two genes are *completely* contained within the translational reading frames of other, larger ones. This configuration is accomplished by having the translational frames in different phases, as explained in Chapter 13. While such overlapping genes are found in many viruses, including even the largest ones, such as herpesviruses and poxviruses, ΦX174 has taken this tendency to an extreme.

Such compactness provides some useful advantage to this bacteriophage, but as with all dynamic systems, there is a price. In a viral genome with such overlaps, one base change in a region of overlapping genes can affect *two* rather than one gene function. For this reason, more mutations would be expected to be lethal than is generally seen in viral genomes. This is indeed the case with ΦX174, whose sequence is more strongly conserved during replication than is the case with other DNA viruses, and generation of mutations in this virus for genetic analysis is a laborious task.

Overlapping genes probably result in the virus being less adaptable to host and other changes in its natural environment. This conservatism could have a negative survival value in the prokaryotic world, but the survival of the virus is clear evidence that deleterious effects are compensated by the efficiency of gene packaging.

QUESTIONS FOR CHAPTER 17

1 The drawings in the following table represent possible structures for replicating DNA molecules.

Structure	Possible for adenovirus?
	
	
	
	
	

a Indicate which ones might be found if you examined replicating adenovirus DNA isolated from an infected host cell.

b Adenovirus DNA replication proceeds in two stages. Suppose that you have an in vitro system that allows you to examine features of this synthesis. The reaction mixture has all the required viral and host proteins. Predict the effect of the following modifications on the process of the two stages. Use a "+" sign if the stage will occur normally and a "–" sign if the stage will be blocked by the treatment.

Continued

Modification	First stage	Second stage
Control (no treatment)	+	+
Removal of the terminal protein from both 5′ ends of DNA genome		
Removal of the terminal complementary sequences from one end of the DNA genome		
Prevention of maturation of terminal protein from 80 kd to 55 kd form		

2 Cells that have been infected with adenovirus 2 (Ad2) are treated with the chemicals shown in the accompanying table. In each case, treatment inhibits the production of progeny Ad2 virus in the cell. Briefly (in the space provided) give a reason why the Ad2 life cycle is blocked in each case.

Chemical	Effect on cell	Reason for Ad2 inhibition
NH_4Cl	Blocks acidification of secondary lysosomes and endosomes	
Vinblastine	Disrupts the microtubular cell cytoskeleton	
Emetine	Inhibits protein synthesis	

3 A papilloma (wart) virus enters a cell and does not produce progeny virus; however, episomal DNA is maintained within the cell, and some gene expression occurs. Of which kind of infection is this an example?

4 What are the functions of T antigen during the SV40 infectious cycle?

5 Which of the following about the life cycle of SV40 is false?

a It expresses three transcripts encoding three capsid proteins late.
b The genome contains a specific sequence of nucleotides that acts as a polyadenylation signal for transcripts using either strand of DNA as templates.
c It has specific promoters controlling expression of early and late transcripts.
d It replicates in the nucleus.
e It replicates using mostly cellular enzymes.