
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Chapter 4

# Virus Replication

A virus must undergo the process of **replication** to create new, infectious virions that are able to infect other cells of the body or subsequent hosts. After gaining entry into the body, a virus makes physical contact with and crosses the plasma membrane of a target cell. Inside, it releases and replicates its genome while facilitating the manufacture of its proteins by host ribosomes. Virus particles are assembled from these newly synthesized biological molecules and become infectious virions. Finally, the virions are released from the cell to continue the process of infection.

The seven stages of virus replication are categorized as follows:

1. Attachment
2. Penetration
3. Uncoating
4. Replication
5. Assembly
6. Maturation
7. Release

A mnemonic to remember the stages of virus replication is the sentence "**A PUR**ple **A**pple **M**ight **R**edden." The letters in bold are the first letters of the names of the seven stages in order.

All viruses must perform the seven stages in order to create new virions. Some stages may take place simultaneously with other stages, or some stages may take place out of order, depending upon the virus. This chapter describes the details of what occurs during each stage of viral replication.

## 4.1 ATTACHMENT

A cell interacts with the extracellular world at the plasma membrane, and it is at this location that a virus first makes contact with a target cell. As described in Chapter 3, "Features of Host Cells: Cellular and Molecular Biology Review," the plasma membrane of the cell is composed of a phospholipid bilayer that has numerous proteins protruding from the membrane. These surface proteins have a variety of functions that include transporting ions and molecules, facilitating the binding of one cell to another, or acting as receptors for incoming proteins. The majority of plasma membrane proteins are **glycosylated**, meaning that they have been modified with sugars and carbohydrates. To infect a cell, it is critical that a virus initiates **attachment**—the binding of the virus to the host cell. This interaction is specific: the virus contains a **virus attachment protein** that adsorbs to a **cell surface receptor** on the cell (Table 4.1). The target receptor molecules on the cell surface are normal molecules required for cellular functions that viruses have evolved to exploit, usually glycoproteins or the sugar/carbohydrate residues present on glycoproteins or the plasma membrane. For instance, rhinovirus binds a protein known as intercellular adhesion molecule 1 (ICAM-1), involved in the attachment of one cell to another. Influenza A virus strains bind to the sialic acid sugars found at the ends of cellular carbohydrate chains, and herpes simplex viruses (HSV) reversibly bind to glycosaminoglycans (GAGs), such as heparan sulfate, in order to bind to the herpesvirus entry mediator protein or nectins on the cell surface (Fig. 4.1).

Some viruses also require **coreceptors** to infect cells. HIV initially binds to a protein known as CD4 on the surface of T lymphocytes ("T cells") but requires one of two coreceptor proteins to continue the process of infection. As will be described later in Chapter 11, "Human Immunodeficiency Virus," humans that have a modified version of CCR5, one of these coreceptors, are largely

**TABLE 4.1 Cell Surface Receptors for Attachment of Human Viruses**

| Virus | Cell surface receptor(s) |
|---|---|
| Rhinoviruses | Intercellular adhesion molecule 1 (ICAM-1) (90%), low-density lipoprotein receptor (10%) |
| Poliovirus | Poliovirus receptor (PVR) CD155 |
| Human immunodeficiency virus | CD4 (receptor); CCR5 or CXCR4 (coreceptors) |
| Influenza A virus | Sialic acid |
| Measles virus | CD46, CD150 |
| Herpes simplex virus-1 | Heparan sulfate, HVEM, Nectin-1 |
| Dengue virus | DC-SIGN |
| Hepatitis B virus | Sodium taurocholate–cotransporting polypeptide |
| Human papillomavirus | Heparan sulfate, integrins |

**Essential Human Virology.** http://dx.doi.org/10.1016/B978-0-12-800947-5.00004-1

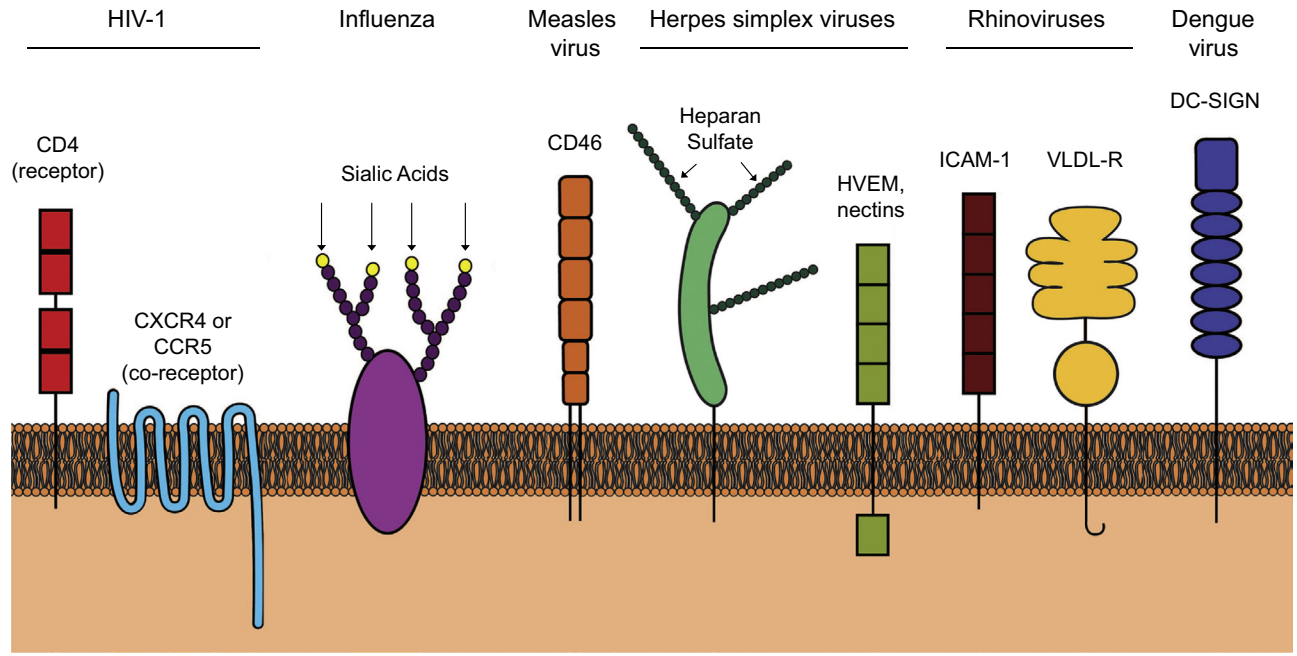


**FIGURE 4.1 Cell surface receptors.** Different viruses use specific cell surface receptors for attachment. HIV-1 requires CD4 as a receptor and chemokine receptors CCR5 or CXCR4 as coreceptors. Influenza viruses bind to terminal sialic acid residues found on cell surface glycoproteins. Laboratory strains of measles virus bind CD46 (although CD150 is also a receptor for the virus). Herpes simplex virus-1 initially binds to heparan sulfate on GAGs in order to specifically bind entry receptors, such as HVEM or nectins. Ninety percent of rhinoviruses use ICAM-1 as a receptor, while 10% use the VLDL receptor. Dengue virus attaches using DC-SIGN. Note the different structures and types of receptors that viruses use for entry. The tropism of the virus is determined by which cells in the body express the cell surface receptor.

resistant to infection with HIV because the virus cannot use the modified CCR5 as a coreceptor and so infection is blocked. Infection of a cell can be prevented if attachment of the virus can be inhibited, and virus attachment proteins are the target of many antiviral drugs in use and in development.

Attachment involves opposing electrostatic forces on the virus attachment protein and the cell surface receptor. The virus attachment protein is located in the outermost portion of the virus, since this is where contact with the cell occurs. The attachment protein protrudes from the envelope of an enveloped virus, whereas nonenveloped viruses have one or more capsid proteins that interact with the cell surface receptor. The viral attachment proteins can extend from the surface of the virion or can be within "canyons" formed by capsid proteins. For example, 90% of human rhinovirus serotypes bind to ICAM-1 on the surface of cells. Instead of binding to the outside of the rhinovirus capsid, the molecule docks into a deep canyon formed by the rhinovirus VP1, VP2, and VP3 proteins (Fig. 4.2A). In contrast, 10% of human rhinoviruses attach to the very low-density lipoprotein (VLDL) receptor. This interaction does not occur in canyons formed by the viral proteins, however. Instead, several VP1 proteins at the vertices of the icosahedral capsid bind to the receptor (Fig. 4.2B). Even if the binding affinity between the VP1 protein and the VLDL receptor is low, the multiple VP1 proteins increase the total binding strength of the interaction. This example also illustrates that different strains of the same virus can take advantage of different cell surface receptors for attachment.

## 4.2 PENETRATION

Following attachment, successful viruses quickly gain entry into the cell to avoid extracellular stresses that could remove the virion, such as the flow of mucus. **Penetration** refers to the crossing of the plasma membrane by the virus. In contrast to virus attachment, penetration requires energy, although this is contributed by the host cell, not the virus.

Several different mechanisms are utilized by viruses to gain entry into a cell (Fig. 4.4, Table 4.2). One of these takes advantage of a normal host process: endocytosis. As described in Chapter 3, "Features of Host Cells: Cellular and Molecular Biology Review," cells are able to import molecules through the process of endocytosis. Receptor-mediated endocytosis occurs when receptors on the cell surface are bound by their ligands and internalized in clathrin-coated pits or caveolae that become endocytic vesicles. Eventually, these vesicles lose their clathrin or caveolin coating and fuse with "early endosomes," slightly acidic vesicles (pH of 6.0–6.5) that become "late endosomes" as their acidity increases (pH of 5.0–6.0). Late endosomes deliver materials to lysosomes, larger vesicles full of digestive enzymes.

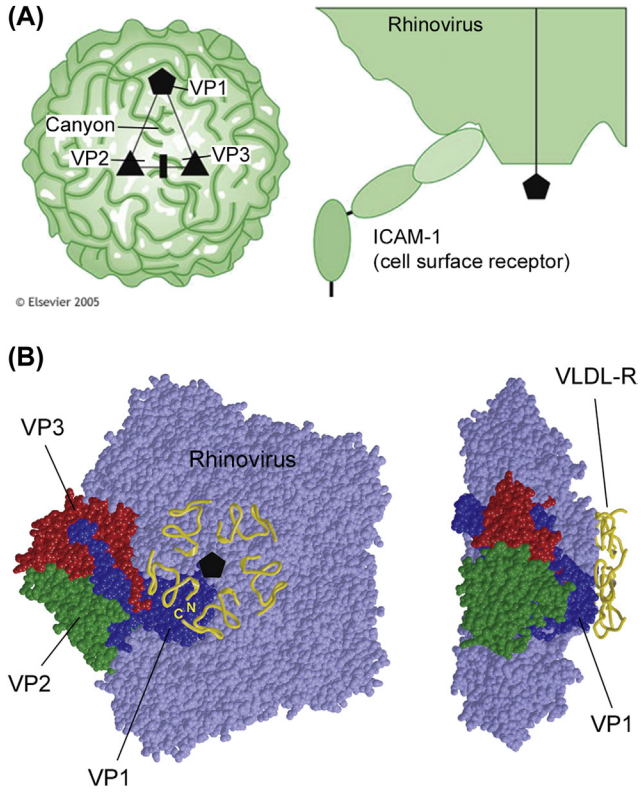


**FIGURE 4.2 Rhinovirus attachment.** (A) 90% of rhinovirus serotypes use ICAM-1 as a cell surface receptor. The cellular protein binds into a canyon formed by capsid proteins VP1, VP2, and VP3. *(Reproduced with permission from Elsevier Academic Press: Alan J. Cann, 2005. Principles of Molecular Virology, fourth ed.)* (B) 10% of rhinovirus serotypes bind the very low-density lipoprotein receptor (VLDL-R). In contrast to ICAM-1 binding, the binding of these rhinovirus serotypes occurs on the fivefold axis at the vertex of the capsid icosahedron, formed by repeating VP1 proteins. This space-filling model shows the surface of the rhinovirus capsid (gray) with one structural unit highlighted, formed by VP1 (blue), VP2 (green), and VP3 (red). The gold molecules represent the VLDL receptors, showing where they bind to rhinovirus protein VP1. *(Reprinted with permission from: Verdaguer et al., 2004. Nature Structural and Molecular Biology. Macmillan Publishers Ltd, 11(5), 429–434.)*

Receptor-mediated endocytosis is commonly used by viruses to penetrate the plasma membrane. As the pH of the endosome drops, the viral proteins change configuration, which allows them to escape from the endosome. Depending upon the virus, this can happen in early endosomes, late endosomes, or lysosomes. Both enveloped and nonenveloped viruses take advantage of receptor-mediated endocytosis to gain entry into the cytoplasm of the cell (Fig. 4.4). Most types of viruses use clathrin-mediated endocytosis to enter the cell, including dengue virus, hepatitis C virus, and reoviruses. A few well-known viruses that infect humans, such as SV40 and papillomaviruses (that cause warts or cervical cancer), use caveolae-mediated endocytosis; this was discovered by using a drug that inhibited the formation of caveolae. Blocking clathrin-mediated endocytosis did not prevent the entry of these viruses into cells. Still other viruses undergo receptor-mediated endocytosis that is independent of both clathrin and caveolin.

Other forms of endocytosis, such as bulk-phase endocytosis and phagocytosis, are also exploited by viruses to enter the cell. In bulk-phase endocytosis, the cell forms a vesicle that engulfs whatever molecules are present in the extracellular fluid, including viruses. **Phagocytosis** is a form of receptor-mediated endocytosis that is used by specialized cells to engulf entire cells. Recently, two large DNA viruses, HSV-1 and mimivirus, were shown to enter cells through phagocytosis-like pathways.

### In-Depth Look: Tropism

Different cells perform different functions within a multicellular organism. As such, not all cells within the body display the same types of cell surface proteins. The **tropism** of a virus refers to the specificity of a virus for a particular host cell or tissue. Viruses will only be able to infect the cells that display the molecules to which their virus attachment proteins bind. Similarly, one reason that certain viruses have a narrow **host range** is because different host species may lack the cell surface proteins that a particular virus uses for attachment. For instance, humans are the only known natural hosts of poliovirus. Because of this, poliovirus has historically been a difficult virus to study because the cell surface receptor it uses for attachment, called CD155 or the poliovirus receptor, is not present in small animal models, such as mice. In 1990, a transgenic mouse strain was engineered to express the human CD155 molecule. These mice were susceptible to infection, whereas the normal nontransgenic mice were not (Fig. 4.3).

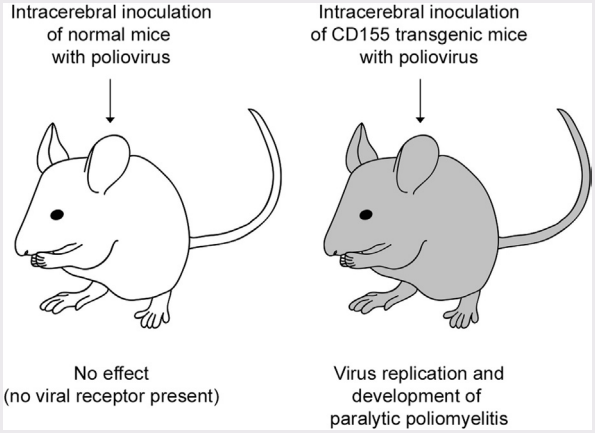


**FIGURE 4.3 Transgenic mice for the human poliovirus receptor CD155.** In 1990, the Racaniello research group generated mice that expressed human CD155, the poliovirus receptor. Poliovirus does not replicate in normal mice because they lack this receptor. When the CD155-transgenic mice were infected intracerebrally with poliovirus, the virus replicated in the brain and spinal cord, causing paralysis. These transgenic mice have proved useful in studying poliovirus infection in a nonprimate animal model. *(Study documented in Ren et al., 2012. Cell 63(2), 353–362.)*

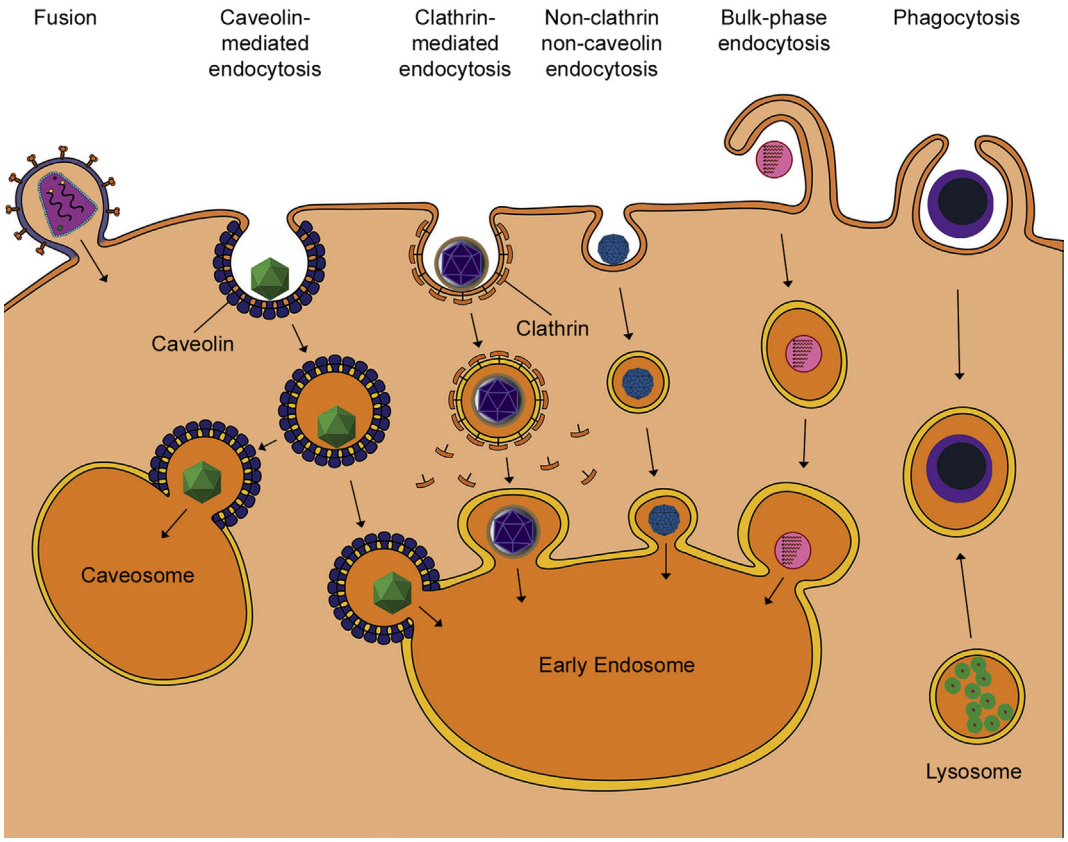


**FIGURE 4.4 Viral penetration into the cell.** Different viruses take advantage of various cellular mechanisms to gain entry into the cell after binding their specific cell surface receptors. Some enveloped viruses undergo fusion, which fuses the viral envelope with the plasma membrane. Both enveloped and nonenveloped viruses take advantage of receptor-mediated endocytosis in caveolin- or clathrin-coated pits to gain entry into the cytoplasm of the cell. Still other viruses undergo receptor-mediated endocytosis that is independent of both clathrin and caveolin. Bulk-phase endocytosis and phagocytosis are also utilized by viruses to gain entry into the cell.

**TABLE 4.2 Methods of Penetration for Select Human Viruses**

| Type of penetration (entry) | Virus examples |
|---|---|
| Clathrin-mediated endocytosis | Dengue virus, hepatitis C virus, reovirus, adenovirus, parvovirus B19, West Nile virus |
| Caveolin-mediated endocytosis | Human papillomavirus, SV40, hepatitis B virus |
| Fusion | HIV, influenza, respiratory syncytial virus, herpes simplex viruses, dengue virus, Ebola virus |

A method of penetration that is used exclusively by enveloped viruses is **fusion**. Fusion of the viral envelope can occur at the cell membrane or within endocytosed vesicles, such as the endosome, and is mediated by the same viral protein that is used by the virus for attachment or by a different viral protein, depending upon the virus. For instance, HIV has a protein known as gp120 that binds to CD4 and one of the two coreceptors for entry, CCR5 or CXCR4. Once this occurs, a different viral protein, gp41, fuses the virus envelope with the cell membrane, releasing the nucleocapsid into the cytoplasm.

**Study Break**
Describe the different ways that viruses can gain entry into the cytosol.

## 4.3 UNCOATING

**Uncoating** refers to the breakdown or removal of the capsid, causing the release of the virus genome into the cell to wherever genome replication and transcription will take place. Uncoating can be separated from or tightly linked with penetration, and viruses achieve uncoating in a variety of different ways (Fig. 4.5). For example, rhinoviruses are taken into the cell by receptor-mediated endocytosis in clathrin-coated vesicles. Within the acidic endosome, the virus expands in size about 4%, and one of the capsid proteins, VP1 (viral

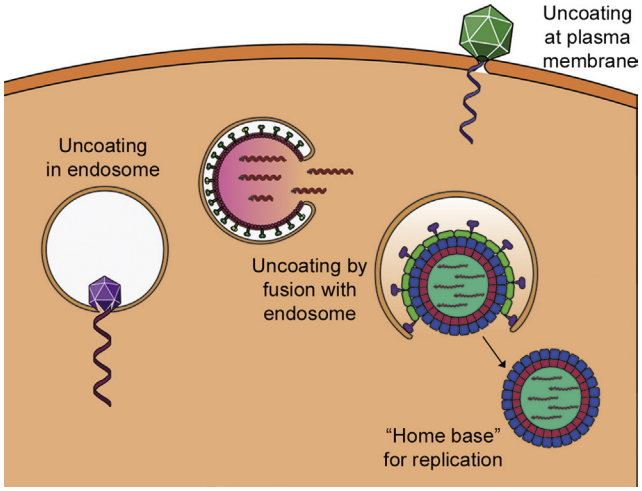


**FIGURE 4.5 Uncoating of virion capsids.** Certain viruses, including rhinoviruses, expand to form pores in the endosome through which the viral genome can escape. Like influenza virus, other viruses induce fusion of the virion envelope with the endosomal membrane, releasing the viral genome. Historically, it has been though that poliovirus capsids do not enter the cell at all: binding of the capsid to the cell surface receptor induces a conformation change that creates a pore in the membrane through which the genome is transported. Many viruses maintain a partially intact capsid in the cytosol that acts as a "home base" for replication, like reoviruses do.

protein 1), forms pores in the endosome that allow the release of the rhinovirus RNA genome. On the other hand, influenza virus has a viral protein known as hemagglutinin (HA) embedded into the virus envelope. HA binds to sialic acid residues found on the surface of respiratory epithelial cells, and penetration occurs through receptor-mediated endocytosis. The low pH of the endosome causes a conformational change in the viral HA protein, revealing a fusion peptide that brings the two membranes close together and fuses the viral envelope with the endosomal membrane. In this case, the HA protein facilitates both attachment and uncoating of the virus. The released viral RNA genome segments are transported to the nucleus and enter through nuclear pores. Other viral capsids, such as those of poliovirus, have been thought to not enter the cell at all: the binding of the poliovirus capsid to the cell surface receptor causes a conformational change in the virion that creates a pore in the cell membrane through which the viral RNA is released into the cytoplasm. In contrast, many viruses remain largely intact after penetration. Reoviruses do not completely uncoat within the cytoplasm, providing a "home base" for genome replication.

Many herpesviruses infect neurons but must replicate in the nucleus, which can be quite a distance from their site of entry at the plasma membrane. After fusion of the viral envelope with the plasma membrane, the intact nucleocapsids of HSV-1 are transported along microtubules to the nucleus. HSV proteins bind to dynein, a host cell protein that "walks" vesicles of cargo along microtubules. At the nucleus, the HSV capsid docks at a nuclear pore and its viral DNA is transported into the nucleus (Fig. 4.6).

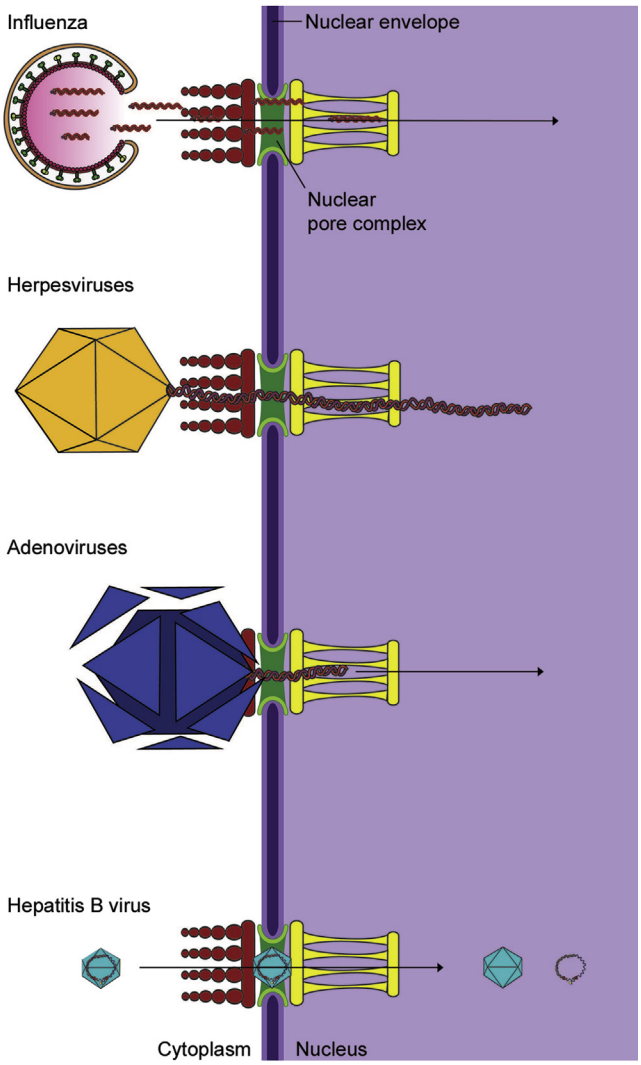


**FIGURE 4.6 Transport of viral genomes into the nucleus.** Several viruses must transport their genomes into the nucleus for viral transcription and/or replication to occur. Influenza genome segments are transported through the nuclear pore into the nucleus. Herpesvirus capsids are transported along microtubules to the nuclear pore, where uncoating occurs. Adenovirus capsids disassemble at the nuclear pore and the viral DNA is transported into the nucleus. Other viruses, including hepatitis B virus, are small enough that the entire capsid might be able to pass through the nuclear pore.

Still other capsids are small enough to pass through nuclear pores: the hepatitis B capsid, with a diameter around 30 nm, may be imported intact through a nuclear pore to uncoat within the nucleus.

## 4.4 REPLICATION

In the same way that our DNA encodes the information to manufacture our proteins, a virus's genome acts as the instructions for the synthesis of virus proteins. To create new virions, the proteins that will be incorporated into the virion are made through expression of viral genes, and the virus genome is

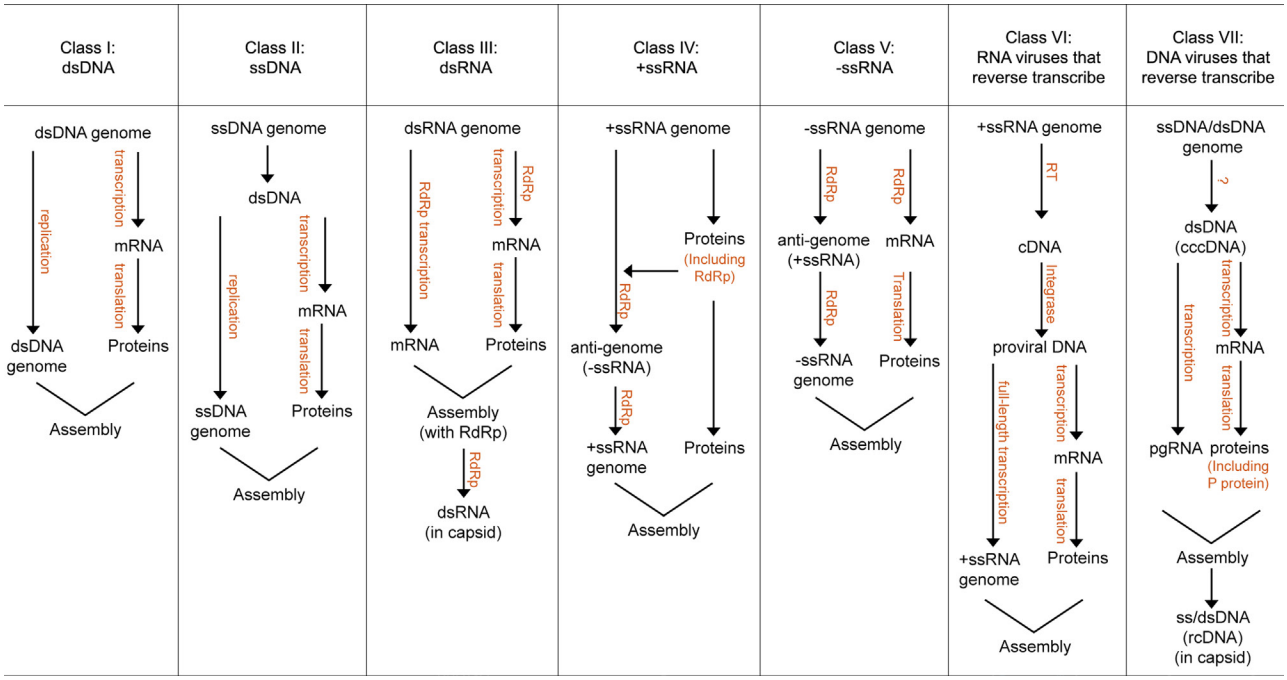


**FIGURE 4.7 The replication strategies of the viruses in the Baltimore replication classes.** All viruses, regardless of their type of nucleic acid genome, must copy the genome and translate proteins in order to assemble nascent virions.

copied through the process of **replication**. The replication strategy of a virus is generally dependent upon the type of nucleic acid genome it contains (Fig. 4.7). Many classes of viruses exploit cellular proteins to replicate their genomes. The location of these proteins will therefore generally dictate where the replication of the viral nucleic acid will take place.

All living organisms have genomes composed of double-stranded (ds) DNA, but viruses are unique in that their genomes can be made of single-stranded (ss) DNA, as well as dsRNA or ssRNA, which can be positive sense or negative sense. The Baltimore classification system, as discussed in Chapter 2, "Virus Structure and Classification," is useful when discussing the general replication strategies of viruses because it categorizes the viruses into seven classes based upon their type of genome:

1. Double-stranded DNA viruses
2. Single-stranded DNA viruses
3. Double-stranded RNA viruses
4. Positive-sense RNA viruses
5. Negative-sense RNA viruses
6. RNA viruses that reverse transcribe
7. DNA viruses that reverse transcribe

As the replication strategies of these classes are discussed, it may be helpful to refer to Table 4.3 and Fig. 2.11 to keep track of which vertebrate-infecting viral families are found in each class.

Viral nucleic acids are found in a variety of configurations. They can be linear or circular, and they can be **segmented** into several smaller pieces within the virion, as occurs with influenza viruses, or **nonsegmented** like rabies virus, containing one molecule of nucleic acid that encodes all necessary genes. Longer molecules are more subject to breaking, but segmented viruses must package all genome segments into a virion for it to be infectious. Regardless of the structure of their nucleic acid, all viruses need to express their viral proteins and replicate their genome within the cell in order to create new virions.

## 4.4.1 Class I: dsDNA Viruses

All living organisms have double-stranded DNA genomes. Viruses with dsDNA genomes therefore have the most similar nucleic acid to living organisms and often use the enzymes and proteins that the cell normally uses for DNA replication and transcription, including its DNA polymerases and RNA polymerases. These are located in the nucleus of a eukaryotic cell, and so all dsDNA viruses that infect humans (with the exception of poxviruses) enter the nucleus of the cell, using the various mechanisms of entry and uncoating mentioned above. Many recognizable human viruses have dsDNA genomes, including herpesviruses, poxviruses, adenoviruses, and polyomaviruses.

Transcription of viral mRNA (vmRNA) must occur before genome replication if viral proteins are involved in replicating the virus genome. In addition, certain translated viral proteins act as **transcription factors** to direct the transcription of other genes. As discussed in Chapter 3, "Features of Host Cells: Cellular and Molecular Biology Review, transcription factors bind to specific sequences within the **promoters** of

**TABLE 4.3 Families of Human Viruses Within Each Replication Class**

| Family | Virus examples |
|---|---|
| **Class I: dsDNA viruses** | |
| Adenoviridae | Adenovirus |
| Herpesviridae | Herpes simplex virus, Epstein–Barr virus, varicella zoster virus |
| Papillomaviridae | Human papillomavirus |
| Polyomaviridae | JC polyomavirus, BK polyomavirus, SV40 |
| Poxviridae | Variola, vaccinia |
| **Class II: ssDNA viruses** | |
| Parvoviridae | Parvovirus B19 |
| Anelloviridae | Torque teno virus |
| **Class III: dsRNA viruses** | |
| Picobirnaviridae | Human picobirnavirus |
| Reoviridae | Rotavirus |
| **Class IV: +ssRNA viruses** | |
| Astroviridae | Human astrovirus |
| Caliciviridae | Norwalk virus |
| Coronaviridae | Human coronavirus |
| Flaviviridae | Dengue virus, yellow fever virus, West Nile virus, hepatitis C virus |
| Hepeviridae | Hepatitis E virus |
| Picornaviridae | Poliovirus, rhinovirus, enterovirus, hepatitis A virus |
| Togaviridae | Eastern equine encephalitis, Chikungunya virus, rubella virus |
| **Class V: –ssRNA viruses** | |
| Arenaviridae | Lymphocytic choriomeningitis virus, Lassa virus, Machupo virus |
| Bunyaviridae | Hantavirus, Crimean–Congo hemorrhagic fever virus |
| Filoviridae | Ebola virus, Marburg virus |
| Orthomyxoviridae | Influenza A virus, influenza B virus |
| Paramyxoviridae | Nipah virus, Hendra virus, measles virus, mumps virus |
| Rhabdoviridae | Rabies virus |
| **Class VI: RNA viruses that reverse transcribe** | |
| Retroviridae | Human immunodeficiency virus-1 and -2 |
| **Class VII: DNA viruses that reverse transcribe** | |
| Hepadnaviridae | Hepatitis B virus |

cellular genes immediately upstream of the transcription start site to initiate transcription of those genes. **Enhancers**, regulatory sequences also involved in transcription, are located farther away from the transcription start site and can be upstream or downstream. dsDNA viruses also have promoter and enhancer regions within their genomes that are recognized not only by viral transcription factors but by host transcription factors, as well. These proteins initiate transcription of the viral genes by the host RNA polymerase II.

Processing of viral precursor mRNA (also known as post-transcriptional modification) occurs through the same mechanisms as for cellular mRNA. Viral transcripts receive a 5′-cap and 3′-poly(A) tail, and some viruses' transcripts are spliced to form different vmRNAs. For example, the genes of herpesviruses are each encoded by their own promoter and are generally not spliced, but the human adenovirus E genome has 17 genes that encode 38 different proteins, derived by alternative splicing of vmRNA during RNA processing.

The dsDNA viruses transcribe their viral gene products in waves, and the **immediate early** and/or **early** genes are the first viral genes to be transcribed and translated into viral proteins. These gene products have a variety of functions, many of which help to direct the efficient replication of the genome and further transcription of the **late** genes that encode the major virion structural proteins and other proteins involved in assembly, maturation, and release from the cell. The replication of the viral genome requires many cellular proteins; having the late genes transcribed and translated after the virus genome has been replicated ensures that the host enzymes needed for replication are not negatively affected by the translation of massive amount of virion structural proteins.

To create new virions, viral proteins must be translated and the genome must also be copied. With the exception of poxviruses, the genome replication of all dsDNA viruses takes place within the nucleus of the infected cell. Eukaryotic DNA replication, also reviewed in more detail in Chapter 3, "Features of Host Cells: Cellular and Molecular Biology Review," is also carried out by DNA polymerases and other proteins within the nucleus. DNA polymerases, whether they are cell derived or virus derived, cannot carry out de novo synthesis, however. They must bind to a short primer of nucleic acid that has bound to the single-stranded piece of DNA, forming a short double-stranded portion that is then extended by DNA polymerase (Fig. 4.8A). Primase is the enzyme that creates primers during cellular DNA replication, and some viruses, such as polyomaviruses and some herpesviruses, take advantage of the cellular primase enzyme to create primers on their dsDNA genomes during replication. Other herpesviruses, such as HSV-1, provide their own primase molecule, although this process occurs less commonly. Still other viruses, such as the adenoviruses, encode a viral protein primer that primes its own viral DNA polymerase (Fig. 4.8B). Cellular DNA polymerases are used by polyomaviruses and papillomaviruses, while all other dsDNA viruses encode their own DNA

polymerases to replicate the viral genome. Many other cellular enzymes and proteins are required for DNA synthesis, and viruses are dependent on these to varying degrees, depending upon the specific virus. The poxviruses are a notable exception to this: they encode all the proteins necessary for DNA replication. In fact, they also encode the proteins needed for transcription of RNA, and so, unlike all other dsDNA viruses, they do not need to gain entry into the nucleus of a host cell for either genome replication or transcription and processing of viral genes, allowing their replication to take place entirely in the cytoplasm.

| Abbreviations | |
|---|---|
| cccDNA | Covalently closed circular DNA |
| cDNA | Complementary DNA |
| DR | Direct repeat |
| dsDNA | Double-stranded DNA |
| dsRNA | Double-stranded RNA |
| IN | Integrase |
| LTR | Long terminal repeat |
| PBS | Primer-binding site |
| pgRNA | Pre-genomic RNA |
| PPT | Polypurine tract |
| rcDNA | Relaxed circular DNA |
| RdRp | RNA-dependent RNA polymerase |
| RT | Reverse transcriptase |
| ssDNA | Single-stranded DNA |
| +ssRNA | Positive-sense RNA |
| –ssRNA | Negative-sense RNA |
| vmRNA | Viral mRNA |

## 4.4.2 Class II: ssDNA Viruses

Viruses with ssDNA genomes infect primarily bacteria and plants, although two families, Anelloviridae and Parvoviridae, infect humans. These viruses are some of the smallest known viruses, with nonenveloped icosahedral capsids of 18–30 nm in diameter, and correspondingly small genomes of 4000–6000 nucleotides. Because they encode only a few genes, they are completely dependent on host cell enzymes for genome replication and transcription.

During replication, the ssDNA genome enters the nucleus of the host cell, where the ssDNA is converted to dsDNA by DNA polymerase during S phase of the cell cycle. This occurs because the ssDNA genome of parvoviruses has "hairpin" ends that fold back and complementarily bind to the ssDNA (Fig. 4.8C). This process, known as **self-priming,**

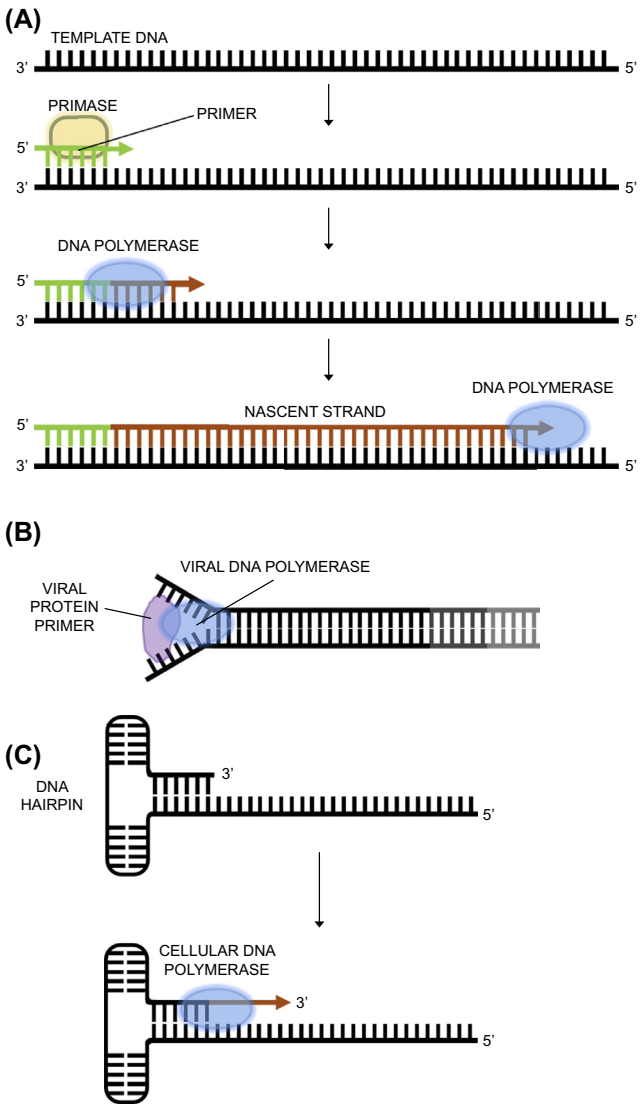


**FIGURE 4.8 DNA priming.** DNA polymerases cannot carry out de novo synthesis and so need a primer in order to replicate DNA. Some viruses take advantage of the cellular primase in order to create primers (A), while other viruses, such as adenoviruses, encode a protein primer that primes its own DNA polymerase (B). In the process of self-priming, the ssDNA genomes of parvoviruses fold back upon themselves to form hairpin ends that act as a primer for host DNA polymerase (C).

creates a primer for DNA polymerase to extend. After the ssDNA genome becomes double-stranded, RNA polymerase II is able to transcribe the viral genes, which are then translated into viral proteins, and DNA polymerase replicates the genome so assembly of nascent virions can occur.

## 4.4.3 Class III: dsRNA Viruses

dsRNA viruses are all nonenveloped and possess icosahedral capsids. They have segmented genomes, and two families of dsRNA viruses infect humans. Viruses in the *Reoviridae* family include rotavirus, so named because

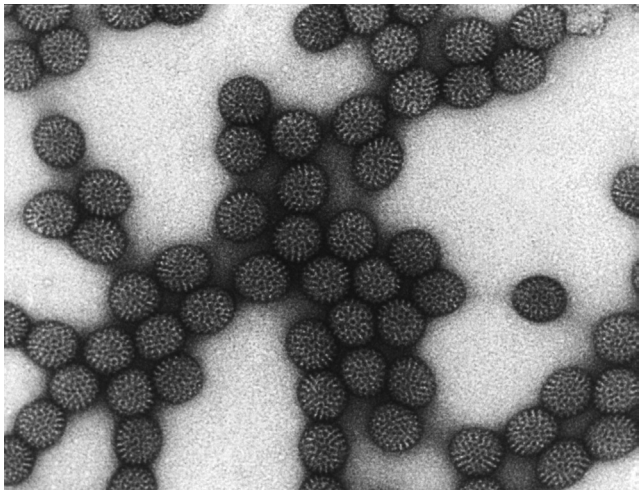

**FIGURE 4.9 Rotavirus virions.** Negatively stained transmission electron micrograph of rotavirus particles at 446,428× magnification. *Micrograph courtesy of the CDC/Dr. Erskine Palmer.*

the virion looks like a wheel (*rota* means "wheel" in Latin; Fig. 4.9). Rotavirus has 11 genome segments and is the major cause of childhood diarrhea. Picobirnaviruses are another family of dsRNA viruses that infect humans, but they are bisegmented, only having two genome segments that together are around 4.2 kb in length (the name of the viral family means "small two-RNA viruses," referring to the two dsRNA genome segments). Human picobirnaviruses have been isolated from diarrhea, although the association of the virus as a cause of a specific disease is currently unclear.

Unlike DNA viruses, viruses with RNA genomes do not usually enter the nucleus of an infected cell. Because they do not have a DNA intermediate, none of the host enzymes involved in DNA replication are required for the replication of the RNA genome. However, RNA viruses must still transcribe vmRNA so that viral proteins can be translated by host ribosomes and new virions can be formed. Cellular mRNA is transcribed by a **DNA-dependent RNA polymerase** called RNA polymerase II. As the name suggests, a DNA-dependent RNA polymerase requires a DNA template to make RNA, so it cannot transcribe mRNA from an RNA genome. Therefore, cells do not contain the enzymes necessary to transcribe mRNA from an RNA template, and so all RNA viruses must carry *or encode* their own **RNA-dependent RNA polymerase** (RdRp) to transcribe viral mRNA (Fig. 4.10A). dsRNA viruses contain an RdRp that is carried into the cell within the virion.

As mentioned in Section 4.3, reoviruses do not completely uncoat within the cytoplasm of the cell, providing a "home base" for transcription. In fact, free viral dsRNA or mRNA is not observed within the cytoplasm of the cell. The RdRp is closely associated with the partially uncoated capsid, which contains pores through which the transcribed

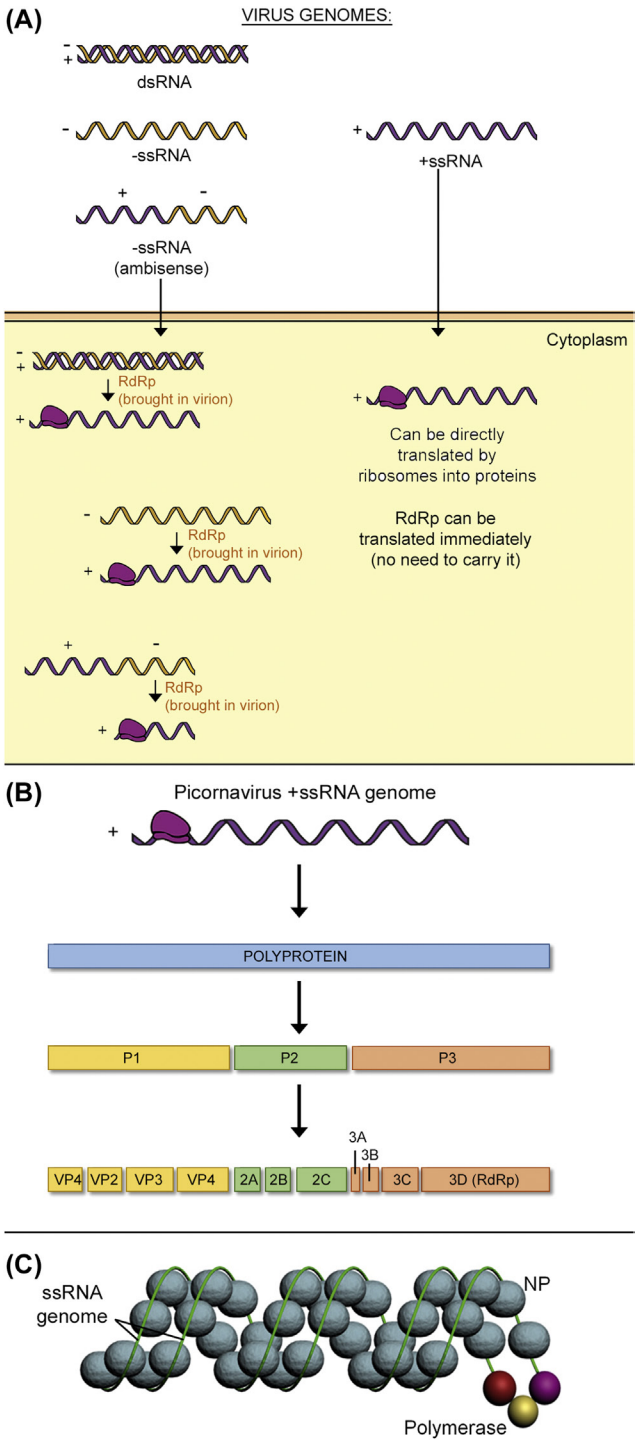


**FIGURE 4.10 Details of RNA virus replication.** (A) Like mRNA, +ssRNA viruses have infectious genomes that can immediately be translated by ribosomes. Consequently, they do not need to bring an RdRp into the cell. RNA viruses with dsRNA, −ssRNA, or ambisense genomes must carry their own RdRp protein into the cell in order for transcription to occur. (B) The +ssRNA picornaviruses encode a single polyprotein that is cleaved several times to create all the proteins necessary for replication. (C) The ribonucleoprotein complex of helical viruses, such as this one from influenza A virus, is composed of the −ssRNA genome, the protecting nucleocapsid protein (NP), and associated proteins, including the RdRp complex. *Graphic from Fig. 1 of Tao et al., 2010. PLoS Pathogens 6(7), e1000943.*

mRNA passes to enter the cytoplasm of the cell, where the mRNA associates with host ribosomes.

Only one of the two RNA strands within the rotavirus genome, the negative strand, is used as a template by the virus's RdRp to transcribe mRNA. The viral mRNA is translated by host ribosomes to produce structural and nonstructural viral proteins. Each genome segment is transcribed into mRNA that is **monocistronic**, meaning that each mRNA transcript encodes one protein only, as is the case with eukaryotic mRNA transcripts. As new capsids are forming, a viral mRNA from each genome segment becomes enclosed within the capsid, along with the RdRp protein. Within the capsid, the RdRp synthesizes along each mRNA transcript just once to create the complementary negative strand, thereby forming the dsRNA genome in the newly formed capsid.

**Study Break**
Describe the difference between a DNA-dependent RNA polymerase, an RdRp, and an RNA-dependent DNA polymerase.

### 4.4.4 Class IV: +ssRNA Viruses

RNA viruses are unique in that their genetic information is encoded using RNA, not DNA. As we have seen, this can occur as dsRNA, but many ssRNA viruses also exist. Viruses with ssRNA genomes that can act directly as mRNA are known as **positive-sense** RNA viruses (abbreviated +ssRNA). Similarly, ssRNA viruses with genomes that are not able to be immediately translated by ribosomes are known as **negative-sense** RNA viruses (abbreviated -ssRNA). Negative-sense RNA must be copied into positive-sense RNA by a viral RdRp before it can be translated by ribosomes (Fig. 4.10A). The terms *positive strand* and *negative strand* are also used interchangeably with these two terms.

+ssRNA viruses are more abundant than any other class of viruses and infect a wide range of host species. They include seven different human viral families, including the coronaviruses, flaviviruses, and picornaviruses, that cause significant disease in humans (Table 4.3). Their abundance indicates that +ssRNA viruses have been very successful evolutionarily.

Because the genome of +ssRNA viruses acts as mRNA, these viruses have genetic information that is **infectious**. Their genomes are translatable by host ribosomes and have 5′-caps (or proteins that act similarly to a 5′-cap) and often contain poly(A) tail sequences at the 3′-end. Experiments that delivered only the genome of poliovirus into the cytoplasm of a cell resulted in new virions being formed, because translation of the genome is the first activity that takes place with +ssRNA genomes. This produces all the viral proteins necessary for orchestrating the remainder of the replication cycle. Where dsRNA viruses must carry an RdRp within the virion, +ssRNA viruses *encode* an RdRp within their +ssRNA genome. The RdRp protein is produced immediately upon entry into the cell by translation of the viral genome. It is important to note, however, that even though the virus encodes its own RdRp protein, cellular proteins are often also required for replication to take place. For example, despite that the poliovirus genome is infectious, it is not replicated when injected into a *Xenopus* frog oocyte (ovum/egg) unless the cytoplasm from a human cell is injected alongside the genome, indicating that at least one human cellular component is required for poliovirus genome replication.

A common characteristic of +ssRNA viruses is that their infectious genome encodes a **polyprotein**, meaning that the genome is translated by ribosomes into a long chain of amino acids that is then cleaved into several smaller proteins. This provides an economical method of deriving several proteins from the translation of only one piece of mRNA. In the case of picornaviruses, the positive-strand genome is translated in its entirety, and then proteases cleave the polyprotein in different locations to create several different proteins (Fig. 4.10B). Alternative cleavage of certain sections results in additional proteins.

Other +ssRNA viruses, such as the togaviruses that include rubella virus, begin by translating only a portion of the +ssRNA genome to create an initial set of proteins that direct the later replication of the genome and translation of other viral proteins. This allows for the creation of "stages" of virus replication, similar to what is observed with immediate early, early, and late gene transcription of certain DNA viruses. Creation of polyproteins also commonly accompanies this method of translation, and termination suppression results in the production of different polyprotein chains. This happens at a low rate (approximately 10% of proteins initially synthesized from the togavirus +ssRNA genome) but results in the generation of important viral proteins, including the viral RdRp.

As the replication proteins of +ssRNA viruses are synthesized, they tend to gather at or within certain membranes in the cell, creating **replication complexes**. For example, the viral proteins of poliovirus remain bound to rough ER (rER) membranes or secretory vesicles. Poliovirus is a nonenveloped virus, so the function of this appears to be to concentrate viral proteins in one location of the cell to better facilitate replication processes. Another reason this may have evolved is to shield the viral ssRNA from intracellular immune responses, discussed further in Chapter 6, "The Immune Response to Viruses."

The viral genome of +ssRNA viruses is used to create a complementary negative strand, the **antigenomic RNA**, that is used as a template to create many copies of the +ssRNA genome. Along with viral protein production,

copying of the viral genome is a necessary step in generating the required elements for creating new virions.

### 4.4.5 Class V: –ssRNA Viruses

In contrast to +ssRNA viruses, negative-sense RNA viruses (–ssRNA viruses) have genomes that do not act as mRNA. Therefore, like their dsRNA counterparts, they must carry an RdRp within the virion into the cell. There exist six –ssRNA virus families that include some of the most well-known disease-causing viruses, including Ebola virus, Marburg virus, measles virus, mumps virus, rabies virus, and influenza virus. These viruses have enveloped, helical capsids and can have segmented genomes (like influenza) or nonsegmented genomes (like rabies virus).

Viruses with –ssRNA genomes generally do not enter the nucleus (although the –ssRNA influenza viruses are a notable exception to this rule). The –ssRNA genomes are not capped and do not have poly(A) tails, because the –ssRNA does not function as mRNA. Instead, the genome must first be transcribed by the viral RdRp into mRNA, which is then translated. –ssRNA viruses have helical nucleocapsids, where the viral RNA is coated with a repeating nucleocapsid protein, termed NP or N. In addition, the viral RdRp and other proteins necessary for transcription also associate with the nucleocapsid protein and viral RNA. Together, the complex of viral RNA and proteins is termed the viral **ribonucleoprotein** complex, because it contains RNA and viral proteins (Fig. 4.10C). Within the cell, the –ssRNA is immediately transcribed into viral mRNAs by the viral RdRp and any other required helper proteins to produce the virus's proteins.

As with +ssRNA viruses, antigenomic RNA is created to act as a template for replication of the genome. In the case of –ssRNA viruses, the complementary antigenomic RNA is +ssRNA. This antigenome is not identical to the positive-sense viral mRNAs produced during infection, however, since viral mRNAs are capped and polyadenylated. At some point during viral replication, a switch occurs so the RdRp drives genome replication over mRNA transcription. This can occur because certain translated viral proteins bind to the –ssRNA genome at sites that would normally stop the polymerase, allowing it to continue copying the entire genome into the antigenome. In some cases, newly translated viral proteins join the RdRp complex to promote genome replication over viral mRNA transcription.

Certain RNA viruses, termed **ambisense** viruses, have genomes that are partially negative sense and partially positive sense. They are still considered within the class of –ssRNA viruses, however, because the positive-sense portion of their genome is not directly infectious: it must first be copied into an antigenome segment that is used to create the viral mRNA. The arenaviruses are the only ambisense viruses that infect humans (although plant viruses in the *Tospovirus* genus within the *Bunyaviridae* family also have an ambisense genome). The arenaviruses, which include Lassa virus and lymphocytic choriomeningitis virus, have two genome segments, a long (L) segment and short (S) segment, which are each ambisense. Positive-sense viral mRNA is transcribed by the RdRp from the negative-sense portions of these segments. A complementary antigenome is also transcribed for each segment, and this is used to create the viral mRNA from the positive-sense portions of the genome segments (Fig. 4.11).

RNA viruses are more prone to mutation than DNA viruses. All polymerases, whether they use DNA or RNA as a template, introduce errors as they incorporate an incorrect nucleotide, but DNA-dependent DNA polymerases have **proofreading** ability: they can remove an incorrectly placed nucleotide and replace it with the correct one. RdRps, on the other hand, do not have proofreading ability. This raises the overall error rate of the enzyme, from 1 error per $10^9$ bases for a DNA polymerase to greater than 1 error per $10^5$ bases for an RdRp, which results in lower enzyme **fidelity**, or accuracy. RNA viruses have some of the highest mutation rates of all biological entities. Mutations generated by RdRps may result in mutated viral proteins and, subsequently, slightly different strains of the virus that may survive better under environmental pressures.

When more than one strain of virus enters a cell, **recombination** can occur. Recombination is the process by which a virus exchanges pieces of its genetic material with another strain of the same virus. This process, which can occur in dsRNA, +ssRNA, or –ssRNA viruses, occurs during genome replication when the RdRp, while copying one RNA genome template, switches to the template of another strain of the virus and continues replicating, thereby creating a hybrid genome that is different from either parent strain. The switching of

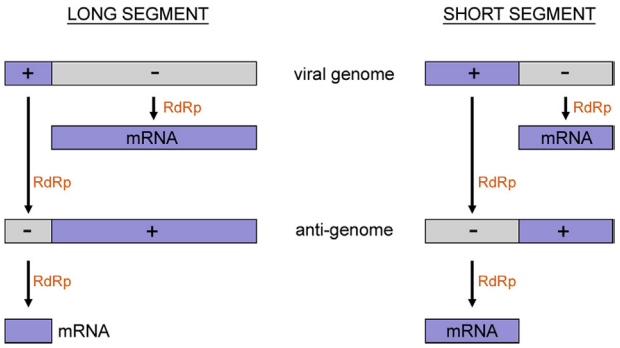


**FIGURE 4.11 Replication of ambisense genomes.** Ambisense genomes are composed of both –ssRNA and +ssRNA. The viral RdRp transcribes the mRNA from the –ssRNA portion. The +ssRNA portion is not directly translatable by ribosomes and must first be transcribed into the antigenome, which has the opposite sense as the ambisense genome. The –ssRNA portion of the antigenome is then transcribed into mRNA by the viral RdRp.

templates has been shown to occur at random sequences or at complementary sequences on the two genome templates that base pair with each other.

Segmented viruses can also undergo **reassortment** when two strains of virus with segmented genomes enter and replicate within the same cell. When the genome segments are copied, segments from one virus may mix with segments from another virus when they are being packaged into new virions, creating a new strain of virus. This can be potentially dangerous when two strains of viruses from different subtypes reassort to create a viral strain that has not previously circulated within the human population. This occurs with influenza A virus and has been the cause of several influenza epidemics, discussed in more detail in Chapter 10, "Influenza Viruses."

**Study Break**

What is the difference between positive-sense RNA and negative-sense RNA? What does this have to do with the replication of different types of viruses?

### 4.4.6 Class VI: RNA Viruses That Reverse Transcribe

The first event that occurs after a +ssRNA or −ssRNA virus enters a host cell is translation or transcription, respectively. **Retroviruses** also have RNA genomes, but must **reverse transcribe** their genome before using host enzymes to transcribe it. In human cells, DNA is used as a template to create mRNA. Retroviruses, on the other hand, encode and carry within their virions an enzyme called **reverse transcriptase** (RT) that is an RNA-dependent DNA polymerase. Reverse transcriptase is able to reverse transcribe the ssRNA genome into a linear strand of double-stranded complementary DNA (cDNA), which is then integrated into a host chromosome.

There exist only a handful of retroviruses, and even fewer that infect humans (Table 4.3). The most well-studied human retrovirus is HIV, the virus that causes AIDS. Within the body, HIV infects and causes the slow decline of T lymphocytes, unarguably one of the most important immune system cells in the defense against pathogens of all kinds. People with HIV are diagnosed with AIDS when the number of T lymphocytes in the blood falls below a certain number, indicating that the person's immune system is severely compromised (hence the origin of "immune deficiency syndrome" in the name of the disease). Without a functioning immune system, people with AIDS often succumb to **opportunistic infections** that healthy people would manage. It is often these opportunistic infections that ultimately cause death in people infected with HIV.

Retroviruses are unlike any other viruses because their genome is diploid, meaning that two copies of the genome are present within the virion. Other viruses are segmented and have their genomes in several segments, but the segments all encode different viral genes. The retrovirus genome is +ssRNA, although it does not serve as mRNA, like the genomes of +ssRNA viruses do. They are also unique among the RNA viruses because their genome will be copied by cellular enzymes, rather than an RdRp.

The retrovirus genome has several different domains that are of importance during viral replication. The two ends of the genome are flanked by redundant sequences, termed R. Inside of the R domain is the U5 (unique to the 5′) and U3 (unique to the 3′) domains on the 5′- and 3′-ends of the RNA, respectively. These domains will end up forming **long terminal repeats** (LTRs) on each side of the cDNA that will be important for **integration** of the HIV cDNA into the host's DNA.

Because the HIV reverse transcriptase enzyme is a target for drugs against HIV, it has been extensively studied and is the paradigm among retrovirus reverse transcriptases. RT is a heterodimer composed of two different-sized polypeptide chains, although one chain is just a slightly shorter version of the other. It is a unique enzyme because it can perform several enzymatic functions. It acts as an *RNA-dependent* DNA polymerase, a *DNA-dependent* DNA polymerase, and it also has RNase H activity, meaning that it is able to degrade RNA when bound to DNA. Although it is a DNA polymerase, it does not have proofreading ability, and is as error-prone as the RdRps, introducing an incorrect nucleotide once every $10^5$ bases. It is also slow, about a 10th of the speed of DNA polymerase.

Only one of the two viral ssRNA strands is reverse transcribed by RT, although recombination can occur if the reverse transcriptase jumps to the other strand of viral ssRNA during reverse transcription. Initiation of reverse transcription requires a primer, which is provided by one of around 100 tRNAs (most of which are specific for lysine or proline) that the virus obtained from the previous host cell. The tRNA partially unwinds and complementary base pairs to 18 nucleotides within the **primer-binding site** (PBS), located toward the 5′-end of the viral RNA (Fig. 4.12). RT binds to the primer and extends its sequence, adding DNA nucleotides complementary to the viral RNA, until it reaches the end. This forms what is referred to as the **negative-strand strong-stop DNA**, because the RT has extended the negative strand and it stops when it reaches the end of the RNA template. The U5 and R sequences are copied in this segment. The RNase H activity of RT degrades the RNA from the RNA–DNA hybrid, leaving the ssDNA U5 and R sequences. Because the R sequence in the DNA is complementary to the R sequence found on the 3′-end of the viral RNA, the two R sequences bind and the RT is able to continue extending the negative strand until the end, where the PBS sequence

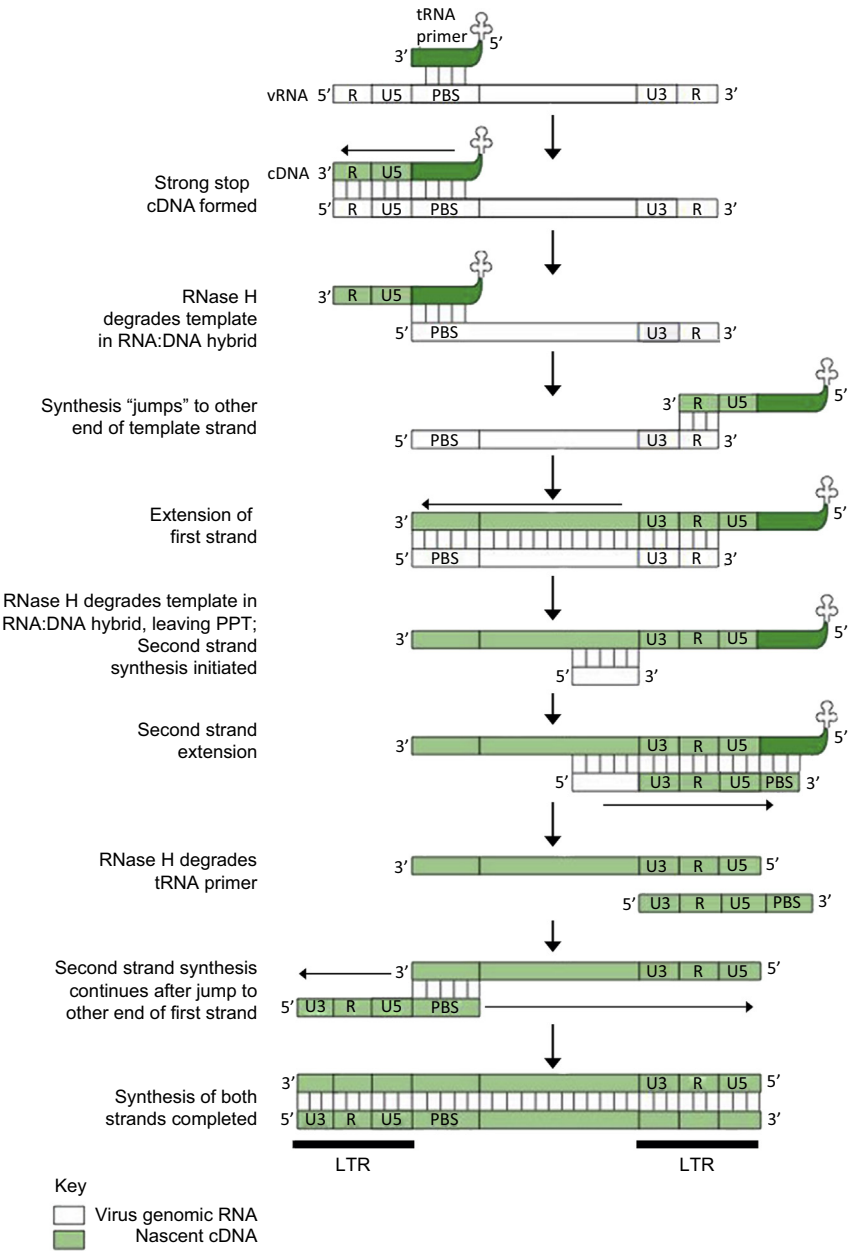


**FIGURE 4.12** **Retrovirus reverse transcription.** Reverse transcription begins with a cellular tRNA binding to the primer-binding site (PBS) in the genomic ssRNA. RT binds to and adds DNA nucleotides to the primer in the 5′→3′ direction, copying the U5 and R sequences and forming the "negative-strand strong-stop cDNA." RT RNase H activity degrades the RNA of the RNA:DNA section. The strong-stop cDNA is transferred to the 3′-end of the RNA and binds it because of the complementary R sites. Synthesis of DNA continues in the 5′→3′ direction, completing the entire negative-sense DNA strand. RNase H activity of RT degrades the RNA template, with the exception of the polypurine tract (PPT), which acts as a primer for the synthesis of the positive-sense DNA strand. RT binds to the PPT and extends it through the U3, R, and U5 domains, completing 18 nucleotides into the tRNA primer to create the new PBS. The tRNA is digested. This "positive-strand strong-stop cDNA" is transferred to the other end of the DNA template; RT completes replication of both strands, creating long terminal repeats (LTRs) on both ends of the double-stranded cDNA. (It is likely that the RNA template circularizes during reverse transcription, which facilitates the "transfer" of the strong-stop cDNAs to the other end of the strand.) *(Image used with permission from Alan J. Cann, 2005. Principles of Molecular Virology, fourth ed. Elsevier/Academic Press, Figure 3.18, Copyright 2005.)*

is found. After being reverse transcribed, the viral RNA is degraded by the RNase H activity of the enzyme, although a small portion resists being digested. This **polypurine tract (PPT)** is composed of purine nucleotides, namely adenine and guanine, and it acts as a primer for RT to begin reverse transcribing the DNA positive strand. This PPT is extended by RT in the 5′→3′ direction, through the U3, R, and U5 domains, until it reaches the tRNA primer, where it finishes the strand by copying 18 nucleotides into the tRNA primer. (You will recall that these 18 nucleotides

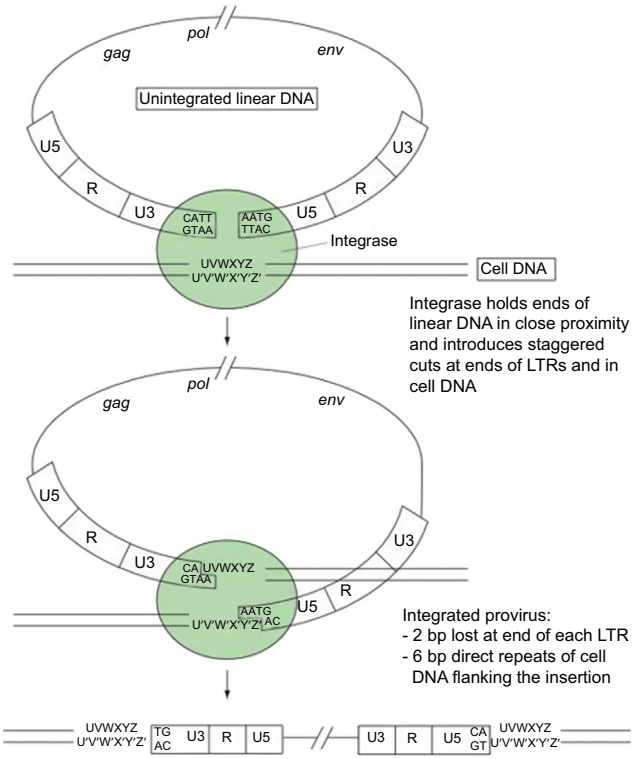


**FIGURE 4.13 Integration of the retrovirus genome into the host chromosome.** Viral integrase removes two base pairs from each end of the proviral DNA, creates a nick in the host chromatin, and joins the proviral DNA to the host DNA. Cellular DNA repair enzymes seal the nick in the sugar–phosphate backbone of the DNA. *Image used with permission from Elsevier/Academic Press: Alan J. Cann, 2005. Principles of Molecular Virology, fourth ed., Fig. 3.20, Copyright 2005.*

of the tRNA initially unwound and bound to the PBS. By copying this segment of the tRNA, another PBS sequence has been created.) This segment is termed the **positive-strand strong-stop DNA**, because it is the positive DNA strand and came to a stop because it reached the end of the viral DNA template.

The tRNA is digested by the RT RNase H, which leaves available the PBS sequence. This anneals to its complementary sequence found at the end of the newly copied negative strand. This acts as the primer for RT to complete the positive strand. Similarly, the negative-strand is extended to the end of the positive-strand strong-stop DNA.

The resulting DNA is double stranded with repeated ends, termed LTRs, that are composed of the U3, R, and U5 domains (Fig. 4.12). These LTRs are important during the integration of this **proviral DNA** into the genome of the host cell. **Integrase** (IN), another necessary retroviral enzyme found within the virion, carries out the process of integration (Fig. 4.13). Having removed two base pairs from each end of the proviral DNA, IN creates a nick in the host chromatin and joins the proviral DNA to the host DNA. DNA repair enzymes within the cell seal the break.

At this point, the integrated viral DNA is like any other cellular gene and will be transcribed by the host RNA polymerase II. Promoter sequences within the U3 region are bound by host transcription factors that are recruited within activated cells, causing the production of several viral mRNAs through splicing and ribosomal frameshifting. The functions of HIV genes will be discussed in more detail in Chapter 11, "Human Immunodeficiency Virus." Full-length mRNA is produced, complete with a 5′-cap and 3′-poly(A) tail, and two copies are packaged into **nascent** (newly formed) virions as the diploid viral genome.

### In-Depth Look: Reverse Transcriptase

Reverse transcriptase was independently discovered in 1970 by two separate groups led by Howard Temin and David Baltimore, who both received Nobel Prize in physiology or medicine in 1975 for their discovery. Reverse transcriptase possesses the activity of three different enzymes in one: an RNA-dependent DNA polymerase, a DNA-dependent DNA polymerase, and RNase H, which degrades the RNA from an RNA:DNA duplex.

The enzyme has been of much utility in molecular biology since its discovery. It was first isolated from viruses but now is produced by bacteria that have been engineered to express the gene. In the laboratory, RT allows for the production of cDNAs from mRNAs. Because cDNA is more stable and easier to manipulate than mRNA, this discovery has been vital in the characterization of gene expression and the study of mRNAs (Fig. 4.14).

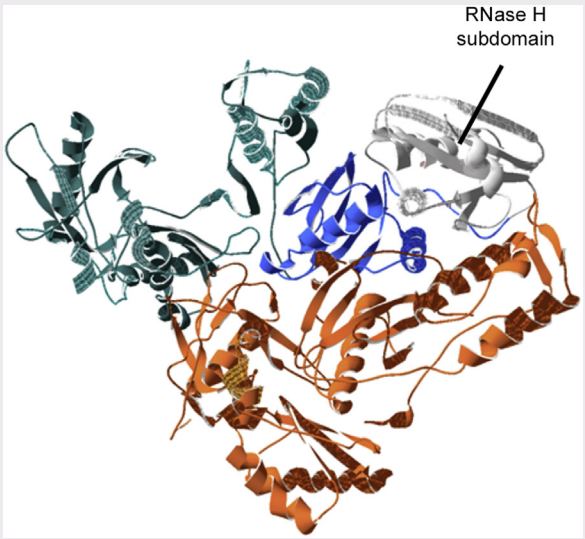


**FIGURE 4.14 HIV reverse transcriptase.** Crystal structure of HIV-1 RT (PDB 1RTD). The p51 subunit is shown in orange; the slightly larger p66 subunit is divided into the fingers (cyan), connection (blue), and RNase H (gray) subdomains. *Image courtesy of Beilhartz et al., 2010. Viruses. 2, 900–926.*

### 4.4.7 Class VII: DNA Viruses That Reverse Transcribe

Retroviruses are RNA viruses that reverse transcribe, but two families of DNA viruses also undergo reverse transcription during their replication within the cell. However, in contrast to retroviruses that undergo reverse transcription as one of the first events in the cell, DNA viruses that reverse transcribe do so at the end of their replication cycle in order to generate their DNA genomes.

Whether RNA or DNA, any virus that reverse transcribes is termed a **retroid virus**. Two DNA virus families fall into this category, *Caulimoviridae*, which infect plants, and the *Hepadnaviridae*, which infect animals. The only human virus within this family is hepatitis B virus (HBV). The virus infects the liver and can cause **hepatitis**, inflammation of the liver. Approximately 5% of infections with HBV become a long-term, **chronic** infection that can lead to liver scarring (cirrhosis) and even liver cancer.

Like other DNA viruses, the genome of hepadnaviruses must also be transported into the nucleus. The DNA genome is partially double stranded and partially single stranded. The complete strand is the negative-sense DNA, and the incomplete strand is positive-sense DNA (Fig. 4.15A). This **relaxed circular DNA** (rcDNA) is transported into the nucleus, where the gapped segment is repaired into double-stranded DNA by a yet unidentified enzyme, completing the **covalently closed circular DNA** (cccDNA). The cccDNA is ligated and maintained as an **episome**, meaning that the circular dsDNA does not integrate into the host DNA but remains as a separate entity within the nucleus.

From the negative strand of episomal cccDNA, host RNA polymerase II transcribes viral mRNAs that leave the nucleus and are translated by host ribosomes to create viral proteins. RNA polymerase II also creates an RNA **pregenome** that leaves the nucleus. Interestingly, it functions as mRNA but also acts as the template for reverse transcription. Within the cccDNA are two identical sequences of 12 nucleotides, termed DR1 (direct repeat 1) and DR2 (direct repeat 2) (Fig. 4.15A). The pregenomic RNA (pgRNA) begins being translated on the negative strand of the DNA at a site upstream of DR1, and RNA polymerase II transcribes the entire negative strand, including the DR2 site. The dsDNA episome is circular, however, and RNA polymerase II continues past its initial starting point, terminating downstream of DR1 at a polyadenylation signal. The result is a long pgRNA that has repeating sequences at both ends that each include a copy of DR1. Part of the repeated sequence folds into a physical structure, known as epsilon ($\varepsilon$), that serves as the initial start of reverse transcription.

The pgRNA and the recently translated P protein are packaged into forming capsids. Similarly to the RT protein of retroviruses, the **P protein** functions as an RNA-dependent DNA polymerase, a DNA-dependent DNA polymerase, and an RNase H, removing RNA from RNA–DNA hybrids. The P protein binds to the ε structure and reverse transcribes a few base pairs from it (Fig. 4.15B). This functions as a primer that binds to the positive-sense pgRNA and reverse transcribes it, creating the complete negative strand of the DNA genome. The RNase H activity of P protein degrades the pgRNA from this strand, leaving a short RNA segment at the end that includes the repeated DR1 sequence. This RNA primer relocates to the other end of the DNA and binds to the identical DR2 sequence there (Fig. 4.15C). P protein extends this positive-sense DNA strand to the end of the negative-sense template. This creates the shorter positive-sense DNA strand within the HBV genome, which bridges the gap between the 5′- and 3′-ends of the negative-sense strand using complementary sequences. This completes genome replication and the creation of the rcDNA.

## 4.5 ASSEMBLY

Viruses are created from newly synthesized components, and to be released from the cell, those components must be collected at a particular site of the cell and undergo **assembly** to form an immature virus particle. In the same way that penetration and uncoating are difficult to separate in the cycle of some viruses, assembly can often occur alongside maturation and release.

The location of virion assembly will depend upon the particular virus. It can take place within the nucleus of the cell, at the plasma membrane, or at a variety of intracellular membranes, such as the Golgi complex. Most nonenveloped DNA viruses assemble their nucleocapsid in the nucleus, since that is the site of genome replication. Viral proteins are imported through nuclear pores to reach the site of assembly. When assembled, most DNA viruses are too large to fit through nuclear pores, however. At this point, some viruses are able to traverse the double-membraned nuclear envelope, while others induce cell lysis or apoptosis to escape the nucleus. On the other hand, viruses with envelopes derived from the plasma membrane usually assemble there.

The nucleic acid genome of a helical virus is protected by repeating capsid proteins. Because of this, capsid proteins can begin wrapping the genome as soon as it is copied (or vice versa, depending upon the virus: the genome can be wrapped around capsid proteins; Fig. 4.16). In contrast, some icosahedral viruses nearly complete the assembly of their capsids before the nucleic acid genome is inserted.

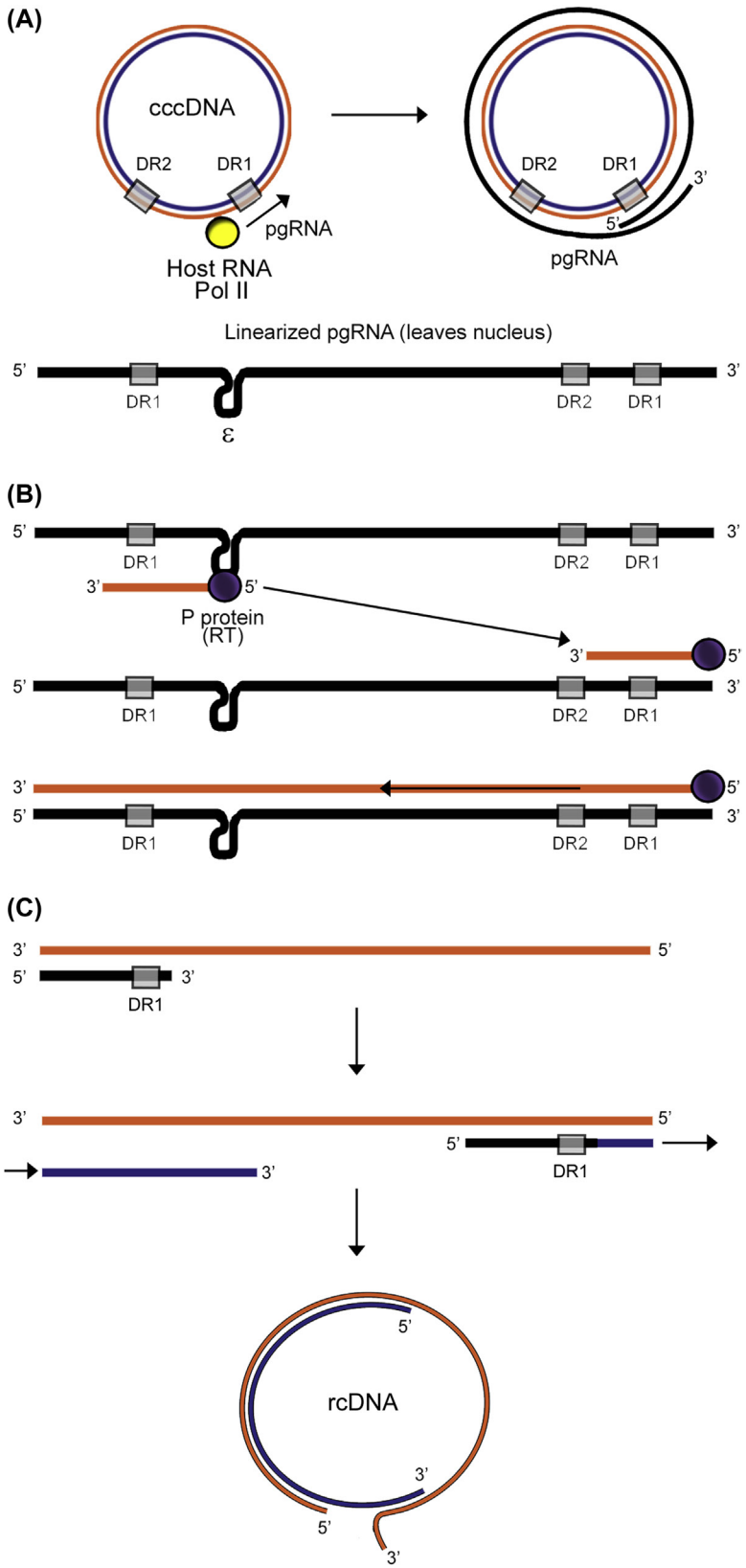


**FIGURE 4.15** **Reverse transcription of the HBV genome.** The HBV genome has identical repeated sequences termed DR1 and DR2. Once the rcDNA is repaired into closed circular DNA (cccDNA), cellular RNA polymerase II transcribes an RNA pregenome (A). It starts upstream of the DR1 site and completes a full circle, ending past the DR1 site where it began. This results in an RNA pregenome that has a 5′-DR1 site and 3′-DR2 and DR1 sites. In the cytoplasm, the pregenomic RNA (pgRNA) is encapsidated. The HBV reverse transcriptase, known as P protein, binds to the ε site and reverse transcribes DNA in the 5′→3′ direction (B). This small piece acts as a DNA primer, which jumps to the other end of the pgRNA, binding to the complementary sequence found there. P protein complex the reverse transcription of the negative DNA strand and degrades most of the RNA:DNA duplex. (C) The remaining RNA segment is transferred to the other end of the negative-sense DNA strand, binding to the complementary DR2 sequence found there. It continues synthesizing DNA in the 5′→3′ direction; because of complementary sequences, the genome circularizes and the P protein completes the synthesis of the shorter positive-sense DNA strand.

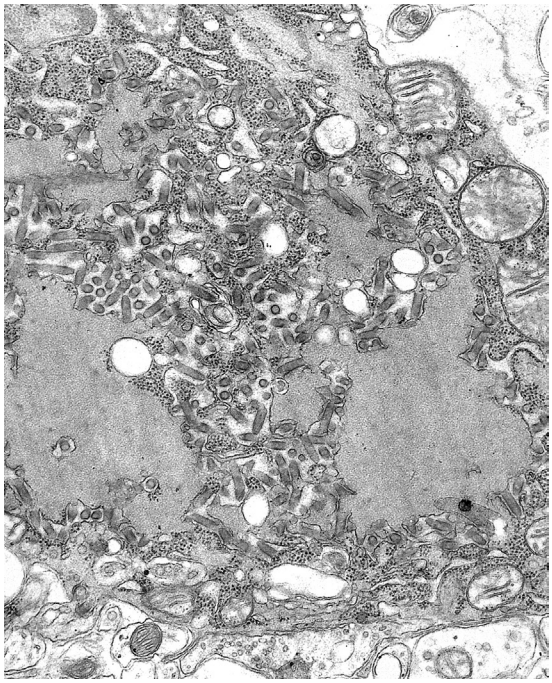

**FIGURE 4.16 Assembly of rabies virions.** This electron micrograph of a cell infected with the rabies virus shows the bullet-shaped capsids assembling within the cytoplasm of the cell. The rabies virus subsequently buds from the plasma membrane, obtaining its envelope. *(Micrograph courtesy of CDC/Dr. Fred Murphy.)*

Spontaneous assembly of the capsid, termed "self-assembly," occurs with the capsid proteins of simple icosahedral viruses, such as the picornaviruses and parvoviruses. The assembly of viruses with more complex architecture is orchestrated by a variety of viral chaperone proteins called scaffolding proteins. Herpesviruses and adenoviruses are examples of large icosahedral viruses that assemble with scaffolding protein assistance.

## 4.6 MATURATION

After the nucleic acid genome and other essential proteins are packaged within the capsid, which was assembled from one or several translated viral proteins, the final steps of virus replication occur: maturation and release. Up to this point, the virion had been in the process of forming, and if the cell were broken open at this point, the virions would not be able to initiate infection of new cells. **Maturation** refers to the final changes within an immature virion that result in an infectious virus particle. Structural capsid changes are often involved, and these can be mediated by host enzymes or virus-encoded enzymes. A good example involves the influenza HA protein. It is involved in attachment to the cell's sialic acid, as described above, and the HA protein is able to bind sialic acid after being glycosylated (via posttranslational modification). However, the HA protein must be cleaved

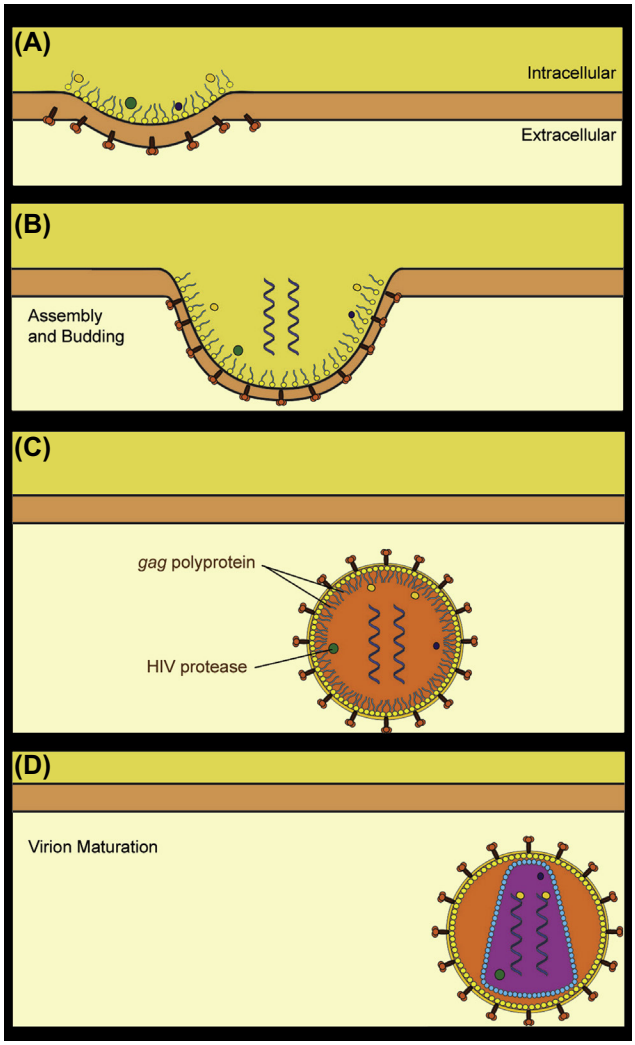


**FIGURE 4.17 Assembly, release, and maturation of HIV virions.** (A) HIV proteins congregate at the plasma membrane of the cell, causing a bud to form in the membrane. The diploid RNA genome is packaged into the assembling capsid (B). The virus is released from the cell membrane, but the Gag polyprotein has not yet been cleaved to separate the capsid and matrix proteins of the virion (C). The HIV protease cleaves the polyprotein, allowing the proteins to complete the infectious virion architecture (D).

into two portions, $HA_1$ and $HA_2$, to become infectious, because although the $HA_1$ portion binds the cell surface receptor, the $HA_2$ portion is what fuses the viral envelope to the endosomal membrane to release the virus into the cytoplasm. This cleavage of HA into $HA_1$ and $HA_2$ is carried out by cell proteases (enzymes that cleave proteins). In contrast, the HIV core particle is composed of proteins encoded by the *gag* gene. The gene is translated into a polyprotein that is cleaved by the viral protease to form the capsid, matrix, and nucleocapsid proteins of the virion. In this case, maturation occurs after the virion has been released from the cell surface (Fig. 4.17) and is required to form an infectious virion. Discussed in

Chapter 8, "Vaccines, Antivirals, and the Beneficial Uses of Viruses," several anti-HIV drugs work by inhibiting the action of the HIV protease, thereby preventing the cleavage of the polyprotein and subsequent formation of an infectious virion.

## 4.7 RELEASE

The final step in the virus replication cycle is **release** of the virion into the extracellular environment, where it can continue the cycle of infection with new cells. Release can occur in several different manners, depending upon the virus. Viruses that obtain their envelope from the plasma membrane generally assemble on the inside layer of the plasma membrane, embedding their envelope proteins into the plasma membrane. As the viral capsid proteins interact, the membrane-associated viral proteins cause the plasma membrane to begin curving around the capsid. This continues until the plasma membrane is completely wrapped around the virus, which leaves the cell. This process is known as **budding** (Fig. 4.17A and B; Fig. 4.18).

Viruses can bud from any of the membrane systems within the cell, including the rER, Golgi complex, or even the nuclear envelope. In this case, the already enveloped virion does not need to bud through the plasma membrane. It generally undergoes exocytosis to leave the cell.

Nonenveloped viruses can also exit the cell via exocytosis. **Lytic** viruses, however, disrupt the plasma membrane and cause the **lysis**, or bursting, of the cell. This releases the nascent virions to infect new cells. Many nonenveloped human viruses are released through cell lysis.

The processes of assembly, maturation, and release are closely linked, but all are required to create progeny infectious virions able to continue the cycle of infection.

## 4.8 VIRUS GROWTH CURVES

In the laboratory, scientists can infect cells with virus to observe the kinetics of the viral replication cycle. **One-step growth curves** are used to study the replication cycle of a virus infection. While at the California Institute of Technology, Emory Ellis and Nobel laureate Max Delbrück devised the "one-step growth curve" using bacteriophages, so named because the viruses replicated simultaneously, all together in one step. To synchronize the infection of many cells at once, a high ratio of virus to cells is used. The **multiplicity of infection** (MOI) refers to this ratio: an MOI of 1 means that 1 virus particle is used per cell for infection, while an MOI of 10 means that 10 virus particles are used per cell. For one-step growth curves, a high MOI is generally used to ensure infection of all cells.

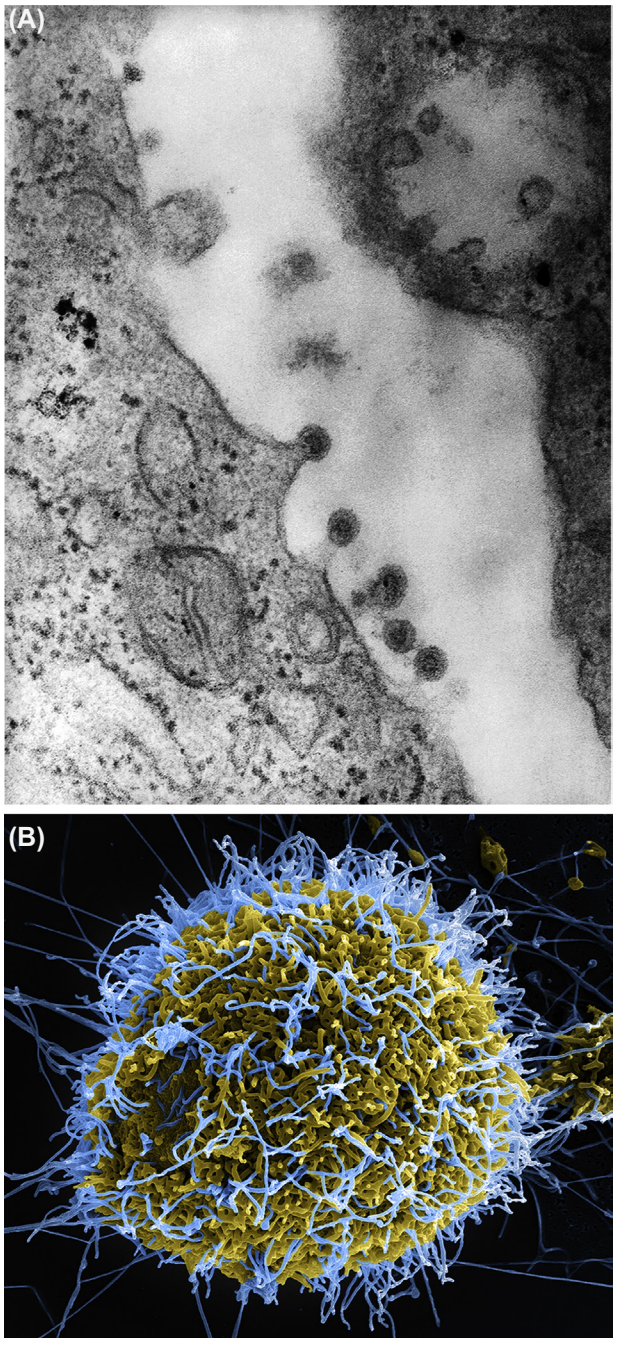


**FIGURE 4.18 Virion budding.** (A) Rubella virus virions are observed budding from the host plasma membrane in this transmission electron micrograph. *(Image courtesy of CDC/Dr. Fred Murphy and Sylvia Whitfield.)* (B) In this digitally enhanced pseudocolored scanning electron micrograph, helical Ebolavirus virions (blue) are budding from an infected cell (yellow). *(Courtesy of the National Institute of Allergy and Infection Diseases (NIAID).)*

When infecting bacteria, the bacteriophages infect the cells immediately. After a very short period of time, generally 1 min, the culture is diluted with media to prevent continued infection, or alternatively, the cells can

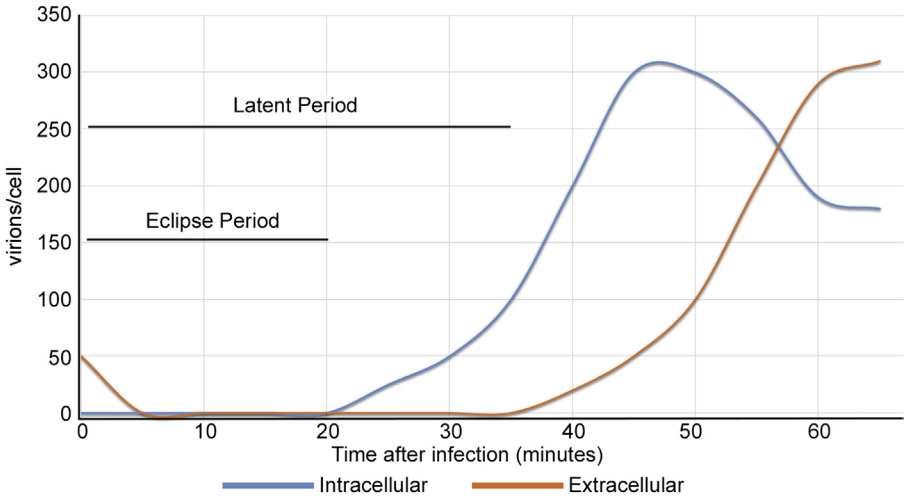


**FIGURE 4.19 The one-step growth curve experiment.** In this experiment, a high MOI is used to synchronize infection of cells with bacteriophages. Once infection occurs, no extracellular virus is observed because all the viruses are replicating internally. No infectious virions are observed intracellularly until assembly occurs within the cell. The amount of time this takes to occur is known as the eclipse period. The infectious virions are released from the cell, at which time they can be observed extracellularly. The amount of time it takes to produce extracellular infectious virions is known as the latent period. A plaque assay, discussed in Chapter 7, "Detection and Diagnosis of Viral Infections," is used to determine the number of infectious virions present.

be centrifuged so the liquid media with any nonadsorbed bacteriophages can be removed and replaced with fresh media. A sample of the culture is taken every minute, and the number of infectious virions is determined using a **plaque assay**, which will be described in detail in Chapter 7, "Detection and Diagnosis of Viral Infections." Initially, there are no additional infectious virions in the culture, because all of the bacteriophages have infected the bacterial cells and are in the process of intracellular replication. This period is known as the **latent period** (Fig. 4.19). Then, the first **burst** occurs as the bacteria are lysed by the bacteriophages, which escape from the cells and are now present in the media. The **burst size** is the number of infectious virions that are released per cell. A typical bacteriophage infection releases 50–200 virions per infected cell, and the latent period is typically 20–30 min.

Because the virions assemble from newly synthesized proteins, scientists can also artificially lyse the bacterial cells at various timepoints to see when the virions are assembled internally and infectious, but not yet released. This period is known as the **eclipse period** and represents the amount of time it takes to form infectious virions within the cell (Fig. 4.19).

One-step growth curves have also been performed with viruses that infect eukaryotes. These viruses take significantly longer to replicate because eukaryotic cells do not replicate as quickly as bacteria; as a result, the enzymes required by some viruses are not necessarily immediately available. Eukaryotic cells are also more complex and involve processes and organelles, such as the rER or Golgi complex, that are not found in bacteria. Consequently, the viral replication cycle takes longer in eukaryotic cells, and the latent period of viruses that infect eukaryotes is significantly longer, about 18–24 h. The eclipse period, where intracellular infectious virions are observed, will vary depending upon whether or not the virus is enveloped. Nonenveloped viruses will be infectious upon assembly and maturation, whereas viruses that obtain their envelope from the plasma membrane will not be infectious until they have budded from the cell.

## SUMMARY OF KEY CONCEPTS

### Section 4.1 Attachment

- Viruses make initial contact with cells at the plasma membrane. The binding of a virion is a specific process that involves the virus attachment protein binding to a cell surface receptor, which will vary depending upon the virus. This determines the tropism of the virus. Some viruses require coreceptors for entry.
- The binding of a virus attachment protein to a cell surface receptor involves electrostatic forces. Virus attachment proteins will be located on the outermost surface of the virion, whether that is the envelope or capsid (for nonenveloped viruses).

### Section 4.2 Penetration

- Penetration is the method by which viruses cross the plasma membrane of the cell.
- Viruses take advantage of normal cellular processes to gain entry into the cytoplasm. One common method involves endocytosis. This can occur through endocytosis of caveolin- or clathrin-coated pits, bulk-phase endocytosis, or phagocytosis.
- Some enveloped viruses use fusion to enter the cell. This process is mediated by viral fusion proteins and merges the viral envelope with the cell membrane.

### Section 4.3 Uncoating

- Uncoating refers to the breakdown of the viral capsid, releasing the genome into the cell. Unlike a cell, which divides into two, nascent virions are assembled de novo by packaging a copied genome into a newly created capsid.
- Many viruses achieve uncoating by escaping from the endosome that they used to enter the cell. Other viruses do not completely uncoat and use the remaining capsid as a home base for replication processes.
- Some viruses uncoat at the nuclear envelope immediately before transporting the genome into the nucleus. A few viruses are small enough to pass through the nuclear pores and uncoat in the nucleus.

### Section 4.4 Replication

- There are seven classes of viruses in the Baltimore classification system, which is based upon the type of nucleic acid and replication strategy of viruses: dsDNA, ssDNA, dsRNA, +ssRNA, −ssRNA, RNA viruses that reverse transcribe, and DNA viruses that reverse transcribe.
- All human dsDNA viruses, with the exception of poxviruses, must gain entry into the nucleus for replication because of the DNA or RNA polymerases that are present there. dsDNA viruses often transcribe their gene products in waves in order to ensure an ordered process that does not put too much overall stress upon the cell. Herpesviruses, poxviruses, and adenoviruses are examples of dsDNA viruses that infect humans.
- The genome of ssDNA viruses is converted into dsDNA by the host cell DNA polymerase after self-priming. At this point, the replication of DNA and transcription of mRNA are carried out in the same way as dsDNA viruses. Anelloviruses and parvoviruses are examples of human ssDNA viruses.
- Human dsRNA viruses, namely reoviruses and picobirnaviruses, are nonenveloped and icosahedral with segmented genomes. To transcribe their genome into vmRNA, they must carry their own RdRp, since the cell does not contain an enzyme that will transcribe mRNA from an RNA template.
- The genomes of +ssRNA viruses are infectious, since positive-sense RNA is able to be directly translated by ribosomes. Consequently, +ssRNA viruses do not carry an RdRp protein but instead encode it so it is immediately translated by ribosomes into the RdRp protein. +ssRNA viruses encode a single polyprotein that is cleaved into several different individual proteins. Seven different families of +ssRNA viruses infect humans and include poliovirus, rhinovirus, Norwalk virus, hepatitis C virus, and rubella virus, among many others.
- −ssRNA viruses are not infectious and must be transcribed into vmRNA before translation can occur. Therefore, −ssRNA viruses must also carry an RdRp into the cell. −ssRNA viruses have helical ribonucleoprotein complexes that are composed of the viral RNA, the protecting capsid protein, and any associated enzymes. As with +ssRNA viruses, an antigenome is transcribed and acts as a template to create the genome. Ambisense viruses are partially positive sense and partially negative sense. Several well-known human viruses have −ssRNA genomes, including influenza viruses, Ebola virus, rabies virus, measles virus, and mumps virus.
- Due to the lack of proofreading ability, RNA polymerases have lower fidelity than DNA polymerases.
- New strains of virus can occur when two different strains infect one cell. Recombination occurs when the genome of an RNA virus is being replicated and the RdRp jumps from the template of one strain to the template of the other strain, creating a hybrid genome. Reassortment occurs when the genome segments of segmented viruses are mixed while being packaged into new capsids.
- Retroviruses are viruses that reverse transcribe an RNA genome into cDNA. Reverse transcriptase is the enzyme that carries this out and has the activity of an RNA-dependent DNA polymerase, DNA-dependent DNA polymerase, and RNase H. HIV, a human retrovirus, becomes a provirus when its IN protein merges the viral cDNA into the host's chromosomal DNA. Transcription and genome replication is carried out by host enzymes.

- Hepatitis B virus is also a retroid virus but instead reverse transcribes in order to create its DNA genome, which is partially single-stranded and partially double-stranded and known as rcDNA. This is repaired to a completely double-stranded episome (cccDNA) in the nucleus of the cell. RNA polymerase II transcribes an RNA pregenome that is reverse transcribed, after being packaged into the capsid, into the rcDNA genome.

Section 4.5 Assembly

- Assembly refers to the packaging of the copied viral genome with newly manufactured viral proteins to create a virion. This can occur in the nucleus, within an organelle like the rER or Golgi complex, or in the cytosol, depending upon the virus.

Section 4.6 Maturation

- Maturation refers to the final changes that must occur within the virion to create an infectious virion rather than an inert particle. This can involve the modification of cell surface receptors, the cleavage of viral polyproteins, or changes to the viral capsid. Maturation is often tightly linked with assembly and/or release.

Section 4.7 Release

- Release refers to the exit of the virion from the cell. This most often occurs through the budding of enveloped viruses or via cell lysis.

Section 4.8 Virus Growth Curves

- One-step growth curves rely on synchronous infection of cells so that the replication process and virion release occur simultaneously in all cells. The eclipse period refers to the amount of time required to assemble infectious virions intracellularly, and the latent period is the amount of time before infectious virions are observed outside the cell. Viruses that infect eukaryotic cells take much longer to replicate than bacteriophages. The eclipse and latent periods will be the same for eukaryotic viruses that obtain their envelope from the plasma membrane, unless maturation occurs after release.

## FLASH CARD VOCABULARY

| | |
|---|---|
| Attachment | Release |
| Penetration | Glycosylation |
| Uncoating | Virus attachment protein |
| Replication | Cell surface receptor |
| Assembly | Coreceptor |
| Maturation | Tropism |
| Host range | Fidelity |
| Phagocytosis | Recombination |
| Fusion | Reassortment |
| Segmented genome | Retroviruses |
| Nonsegmented genome | Reverse transcriptase |
| Self-priming | Opportunistic infections |
| Monocistronic | Long terminal repeats |
| Polycistronic | Integration/Integrase |
| Positive-sense RNA | Proviral DNA |
| Negative-sense RNA | Nascent |
| RNA-dependent RNA polymerase | Retroid virus |
| Polyprotein | Hepatitis |
| Replication complexes | One-step growth curves |
| Antigenomic RNA | Multiplicity of infection |
| Ribonucleoprotein complex | Latent period |
| Ambisense | Eclipse period |
| Proofreading | Burst size |

## CHAPTER REVIEW QUESTIONS

1. List what takes place at each of the seven steps of viral replication.
2. Considering that each virus must bind to a specific cell surface receptor for attachment, explain how you would create a drug that prevents viral attachment.
3. Focusing on the nucleic acids and enzymes involved, draw out the replication strategies of the seven classes of viruses.
4. Regardless of the type of nucleic acid, what are the general requirements for a virus to create functional nascent virions?
5. Make a chart that lists the location of transcription for each of the seven classes of viruses.
6. Explain why +ssRNA viruses do not have to carry their own RdRp within their virions.
7. What is the difference between recombination and reassortment?
8. List the steps involved in the reverse transcription and integration of a retrovirus genome.
9. Describe the steps involved in replicating the genome of HBV.
10. Both HIV and HBV use reverse transcription. Explain how reverse transcription is used differently in the replication of these two viruses.
11. What generally determines whether or not a virus needs to gain entry into the nucleus to replicate?
12. Make a table of the seven classes of viruses and list what the *first* event is that occurs after the virus gains

entry into the cell. Transcription? Reverse transcription? Translation?

13. Which of the cellular processes described in this chapter are limited only to enveloped viruses compared to nonenveloped viruses?
14. Which classes of viruses are more prone to introducing mutations during genome replication?
15. What would be the result of interfering with the maturation of virions?
16. Looking at the one-step growth curves, extracellular virus disappears because the virus enters the cell. Why does the virus initially disappear from the intracellular samples, too?

## FURTHER READING


Baltimore, D., 1970. RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature 226, 1209–1211.

Beilhartz, G.L., Götte, M., 2010. HIV-1 ribonuclease H: structure, catalytic mechanism and inhibitors. Viruses 2, 900–926.

Brenkman, A.B., Breure, E.C., van der Vliet, P.C., 2002. Molecular architecture of adenovirus DNA polymerase and location of the protein primer. J. Virol. 76, 8200–8207.

Clement, C., Tiwari, V., Scanlan, P.M., Valyi-Nagy, T., Yue, B.Y.J.T., Shukla, D., 2006. A novel role for phagocytosis-like uptake in herpes simplex virus entry. J. Cell Biol. 174, 1009–1021.

Ellis, E.L., Delbruck, M., 1939. The growth of bacteriophage. J. Gen. Physiol. 22, 365–384.

Ghigo, E., Kartenbeck, J., Lien, P., et al., 2008. Ameobal pathogen mimivirus infects macrophages through phagocytosis. PLoS Pathog. 4. http://dx.doi.org/10.1371/journal.ppat.1000087.

Goff, S.P., 2013. Chapter 47: Retroviridae. In: Knipe, D.M., Howley, P.M. (Eds.), Fields Virology, sixth ed. Wolters Kluwer Lippincott Williams and Wilkins, pp. 1424–1473.

Haywood, A.M., 2010. Membrane uncoating of intact enveloped viruses. J. Virol. 84, 10946–10955.

Kielian, M., 2014. Mechanisms of virus membrane fusion proteins. Annu. Rev. Virol. 1, 171–189.

Lyman, M.G., Enquist, L.W., 2009. Herpesvirus interactions with the host cytoskeleton. J. Virol. 83, 2058–2066.

Rabe, B., Vlachou, A., Panté, N., Helenius, A., Kann, M., 2003. Nuclear import of hepatitis B virus capsids and release of the viral genome. Proc. Natl. Acad. Sci. U.S.A. 100, 9849–9854.

Ren, R.B., Costantini, F., Gorgacz, E.J., Lee, J.J., Racaniello, V.R., 1990. Transgenic mice expressing a human poliovirus receptor: a new model for poliomyelitis. Cell 63, 353–362.

Risco, C., Fernández de Castro, I., Sanz-Sánchez, L., Narayan, K., Grandinetti, G., Subramaniam, S., 2013. Three-dimensional imaging of viral infections. Annu. Rev. Virol. 1, 140718093033004.

Sanjuán, R., 2012. From molecular genetics to phylodynamics: evolutionary relevance of mutation rates across viruses. PLoS Pathog. 8, 1–5.

Stöh, L.J., Stehle, T., 2013. Glycan engagement by viruses: receptor switches and specificity. Annu. Rev. Virol. 1, 140707224641009.

Verdaguer, N., Fita, I., Reithmayer, M., Moser, R., Blaas, D., 2004. X-ray structure of a minor group human rhinovirus bound to a fragment of its cellular receptor protein. Nat. Struct. Mol. Biol. 11, 429–434.

Temin, H.M., Mizutani, S., 1970. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. Nature 226, 1211–1213.

Tao, Y.J., Ye, Q., 2010. RNA virus replication complexes. PLoS Pathog. 6, 1–3.

Suomalainen, M., Greber, U.F., 2013. Uncoating of non-enveloped viruses. Curr. Opin. Virol. 3, 27–33.

Will, H., Reiser, W., Weimer, T., et al., 1987. Replication strategy of human hepatitis B virus. J. Virol. 61, 904–911.