Introduction to Virology II

DNA Genomes

The strategy of having DNA as a viral genome appears at first glance to be simplicity itself: the host genetic system is based on DNA, so the viral genome replication and expression could simply emulate the host's system. Many surprises await those who believe that this is all such a strategy entails.

Double-Stranded DNA (dsDNA) (Fig. 1)

Viral genomes may consist of double-stranded or partially double-stranded nucleic acids. There are 22 families of viruses with dsDNA genomes; those that include mammalian viruses are the *Adenoviridae*, *Herpesviridae*, *Papillomaviridae*, *Polyomaviridae*, and *Poxviridae*. These dsDNA genomes may be linear or circular. mRNA is produced by copying of the genome by host or viral DNA-dependent RNA polymerase.
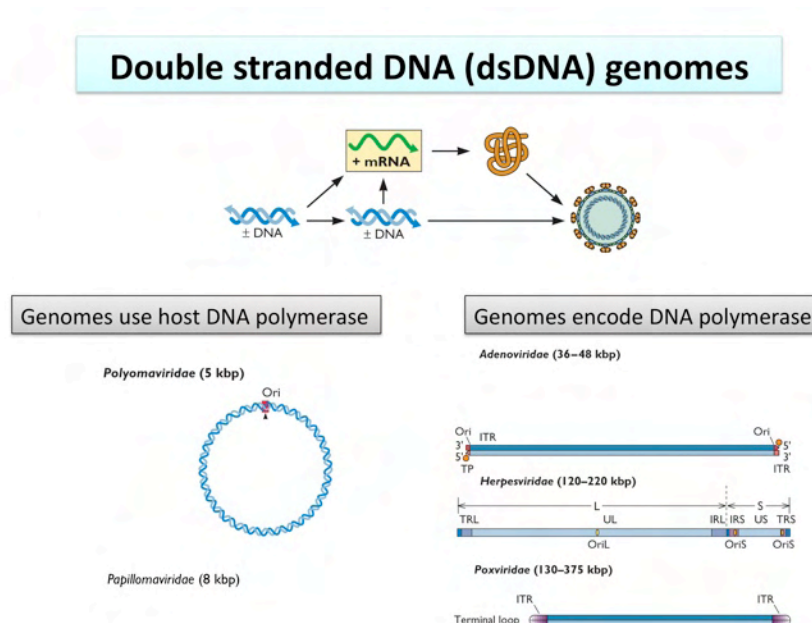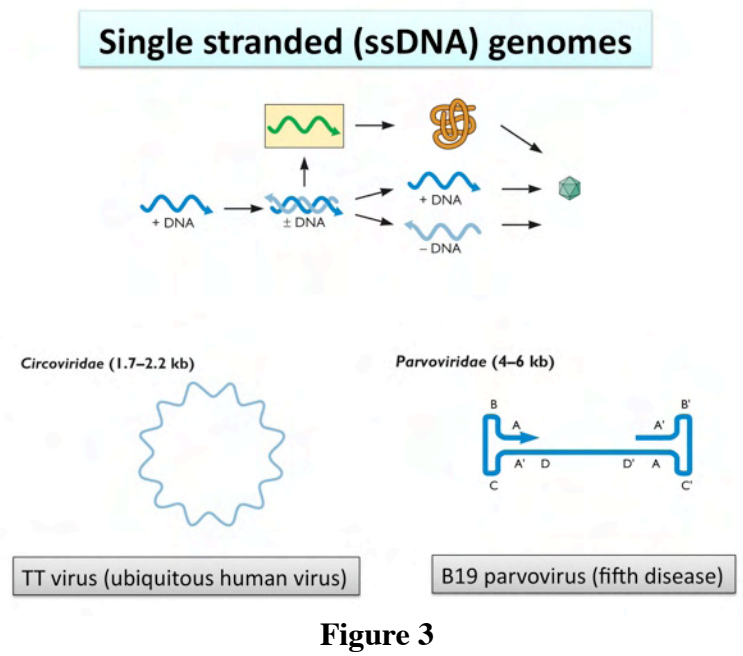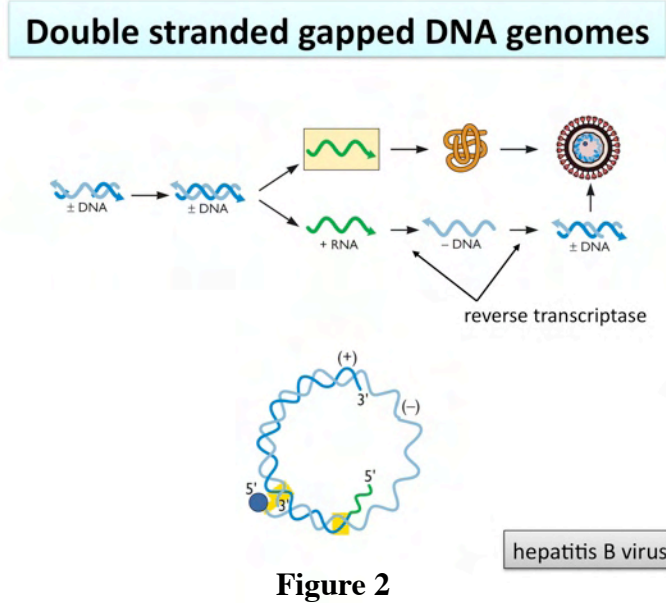

Figure 1

Gapped DNA (Fig. 2)

For partially double-stranded DNA, such as the gapped DNA of *Hepadnaviridae*, the gaps must be filled to produce perfect duplexes. This repair process must precede mRNA synthesis. The unusual gapped DNA genome is produced from an RNA template by a virus-encoded reverse transcriptase homologous to that of the retroviruses.

Single-Stranded DNA (ssDNA) (Fig. 3)

Five families and one genus of viruses containing ssDNA genomes have been recognized; the *Circoviridae* and *Parvoviridae* include viruses that infect mammals. ssDNA must be copied into mRNA before proteins can be produced. However, as RNA can only be made from a double-stranded DNA template, no matter what sense the single-stranded DNA is, DNA synthesis **must**

Double stranded gapped DNA genomes

reverse transcriptase

hepatitis B virus

Figure 2

Single stranded (ssDNA) genomes

Circoviridae (1.7–2.2 kb)

Parvoviridae (4–6 kb)

TT virus (ubiquitous human virus)

B19 parvovirus (fifth disease)

Figure 3

precede mRNA production in the replication cycles of these viruses. The single-stranded genome is produced by cellular DNA polymerases.

RNA Genomes
Cells have no RNA-dependent RNA polymerase that can replicate the genomes of RNA viruses or make mRNA from RNA. The simple solution that has evolved is that RNA virus genomes encode novel RNA-dependent RNA polymerases that produce RNA (both genomes and mRNA) from RNA templates. The mRNA produced is readable by host ribosomes. One notable exception is retrovirus mRNAs, which are produced from DNA templates by host RNA polymerase II.

dsRNA (Fig. 4)

There are seven families of viruses with dsRNA genomes. The number of dsRNA segments ranges from 1 (*Totiviridae* and *Hypoviridae*, viruses of fungi, protozoans, and plants) to 10 to 12 (*Reoviridae*, viruses of mammals, fish, and plants). While dsRNA contains a (+) strand, it cannot be translated as part of a duplex. The (–) strand RNAs are first copied into mRNAs by a viral RNA-dependent RNA polymerase to produce viral proteins. Newly synthesized mRNAs are encapsidated and then copied to produce dsRNAs.

(+) Strand RNA (Fig. 5)

The (+) strand RNA viruses are the most plentiful on this planet; 22 families and 22 genera have been recognized. The *Arteriviridae*, *Astroviridae*, *Caliciviridae*, *Coronaviridae*, *Flaviviridae*, *Picornaviridae*, *Retroviridae*, and *Togaviridae* include viruses that infect mammals. (+) strand RNA genomes usually can be translated directly into protein by host ribosomes. The genome is replicated in two steps. First, the (+) strand genome is copied into a full-length (–) strand. The (–) strand is then copied into full-length (+) strand genomes. A subgenomic mRNA is produced in cells infected with some viruses of this class.
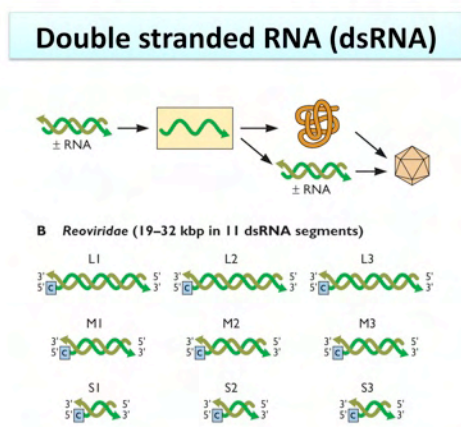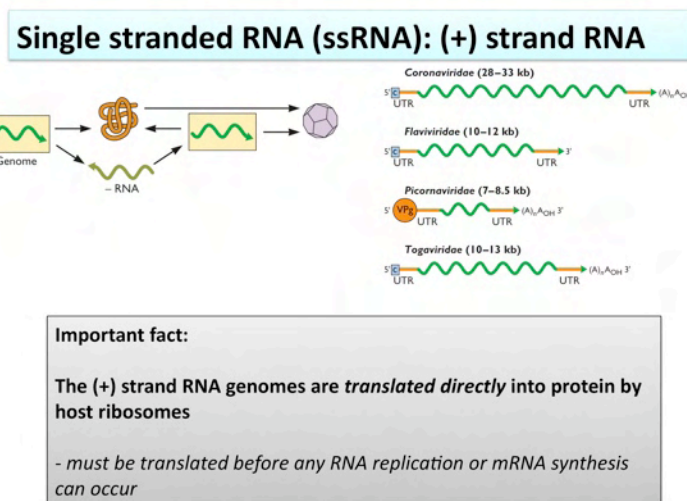

Figure 4

Figure 5

### (+) Strand RNA with DNA Intermediate (Fig. 6)

The retroviruses are unique (+) strand RNA viruses because a dsDNA intermediate is produced from genomic RNA by an unusual viral RNA-dependent DNA polymerase called reverse transcriptase. This DNA then serves as the template for viral mRNA and genome RNA synthesis by cellular enzymes.

(–) Strand RNA (Fig. 7)

Viruses with (–) strand RNA genomes are found in seven families and three genera. Viruses of this class that can infect mammals are found in the *Bornaviridae*, *Filoviridae*, *Orthomyxoviridae*, *Paramyxoviridae*, and *Rhabdoviridae*. By definition, (–) strand RNA genomes cannot be

The retroviral genome strategy is remarkable

RNA is copied into DNA and then back into RNA, some of which is packaged into virions
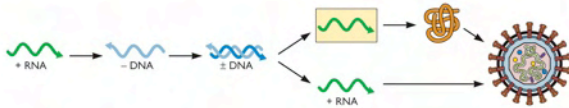
The +ssRNA in the virion is a real mRNA
- however, it is NEVER used as a message!

Upon infection, it is converted to dsDNA by a virion enzyme called **reverse transcriptase.**

This dsDNA intermediate then integrates into the host DNA and becomes a permanent part of the host genome (a "provirus")

Figure 6

Single strand RNA, (-) sense

These genomes cannot be translated directly into protein

- must be FIRST copied to make (+) strand mRNA that can be translated

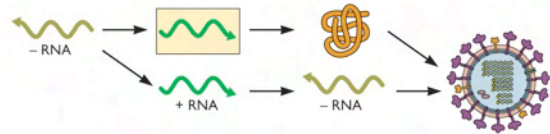- always use a viral encoded, RNA-dependent, RNA polymerase that is found INSIDE the capsid

Figure 7

translated directly into protein but first must be copied to make (+) strand mRNA by a viral RNA-dependent RNA polymerase. There are no enzymes in the cell that can produce mRNAs from the RNA genomes of (–) strand RNA viruses. These virus particles therefore contain virus-encoded, RNA-dependent RNA polymerases that produce mRNAs from the (–) strand genome. This strand is also the template for the synthesis of full-length (+) strands, which in turn are copied to produce (–) strand genomes. (–) strand RNA viral genomes can be either single molecules (nonsegmented) or segmented.

Reassortment

A consequence of a segmented RNA genome is **reassortment**: when cells are coinfected with two different viruses, the progeny includes reassortants that inherit RNA segments from either parental virus (Figure 8). Reassortment among human, avian, and swine influenza viruses gives rise to pandemic strains.
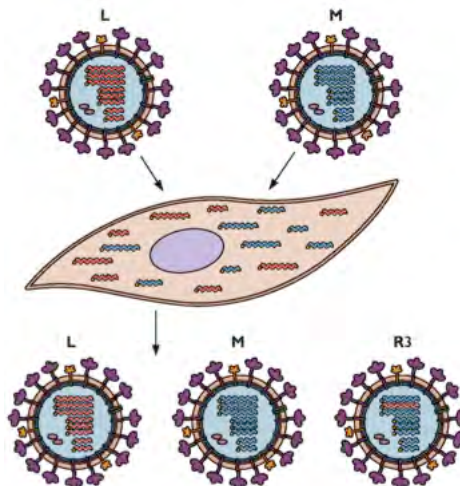

Figure 8

Infectious DNA Clones

Recombinant DNA techniques have made it possible to manipulate the genome of most animal viruses, whether that genome comprises DNA or RNA. In fact, this technology allows the production of infectious viruses from a nucleotide sequence - meaning that no virus can in theory ever be eradicated. The infectious DNA clone is DNA copy of the viral genome that is carried on a bacterial plasmid. Infectious DNA clones, or RNAs derived from them, can be introduced into cultured cells by transfection to recover infectious virus.

The recovery of influenza virus from cloned DNA is achieved by an expression system in which cloned DNA copies of the eight RNA segments of the viral genome are inserted between two promoters (Figure 9). When eight plasmids carrying DNA for each viral RNA segment are introduced into cells, infectious influenza virus is produced. This method was used to recover the 1918 pandemic influenza virus strains from RNA sequences obtained from victims of the disease.
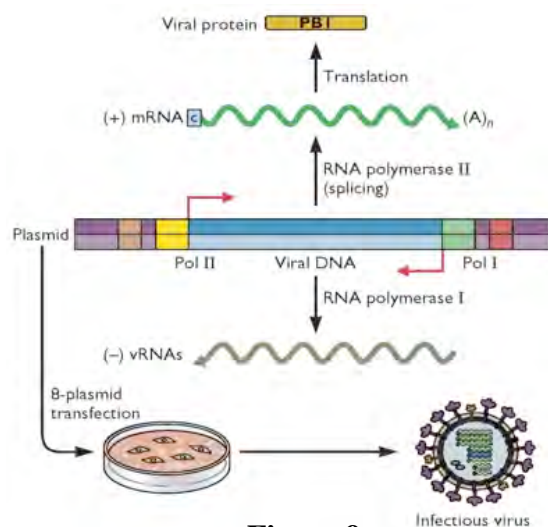

Figure 9

The Cell Surface

In animals, viral infections usually begin at the epithelial surfaces of the body that are exposed to the environment (Fig. 10). Cells cover these surfaces, and the part of these cells exposed to the environment is called the apical surface. Conversely, the basolateral surfaces of such cells are in contact with adjacent or underlying cells or tissues. These cells exhibit a differential (polar) distribution of proteins and lipids in the plasma membranes that creates the two distinct surface domains. As illustrated in Fig. 5, these cell layers differ in thickness and organization. Movement of macromolecules between the cells in the epithelium is prevented by tight junctions, which circumscribe the cells at the apical edges of their lateral membranes. Many viral infections are initiated upon entry of epithelial or endothelial cells at their exposed apical surfaces, often by attaching to cell surface molecules specific for this domain. Viruses that both enter and are released at apical membranes can be transmitted laterally from cell to cell without ever transversing the epithelial or endothelial layers; they generally cause localized infections. In
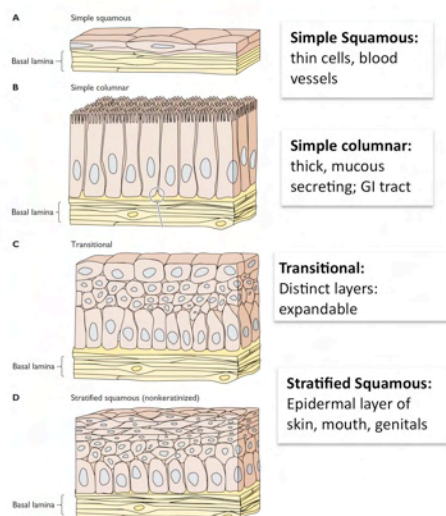
Figure 10

other cases, progeny virions are transported to the basolateral surface and released into the underlying cells and tissues, a process that facilitates viral spread to other sites of replication.

The plasma membrane of every cell type is composed of a similar phospholipid - glycolipid bilayer, but different sets of membrane proteins and lipids allow the cells of different tissues to carry out their specialized functions. The lipid bilayer consists of molecules that possess both hydrophilic and hydrophobic portions; they are known as amphipathic molecules (Fig. 6). They form a sheet-like structure in which polar head groups face the aqueous environment of the cell's cytoplasm (inner surface) or the surrounding environment (outer surface). The polar head groups of the inner and outer leaflets bear side chains with different lipid compositions. The fatty acyl side chains form a continuous hydrophobic interior about 3 nm thick.

Membrane proteins are classified into two broad categories, integral membrane proteins and indirectly anchored proteins (Fig. 11). Integral proteins are embedded in the lipid bilayer, because they contain one or more membrane-spanning domains, as well as portions that protrude out into the exterior and interior of the cell (Fig. 11). Many membrane-spanning domain consist of an α-helix typically 3.7 nm long. Some proteins with multiple membrane-spanning domains form critical components of molecular pores or pumps, which mediate the internalization of required nutrients, or expulsion of undesirable material from the cell, or maintain homeostasis with respect to cell volume, pH, and ion concentration.

The external portions of membrane proteins may be decorated by complex or branched carbohydrate chains linked to the peptide backbone. Such membrane glycoproteins frequently serve as viral receptors.

Some membrane proteins do not span the lipid bilayer, but are anchored in the inner or outer leaflet by covalently attached hydrocarbon chains. Indirectly anchored proteins are bound to the plasma membrane lipid bilayer by interacting either with integral membrane proteins or with the charged sugars of the glycolipids within the membrane lipid.
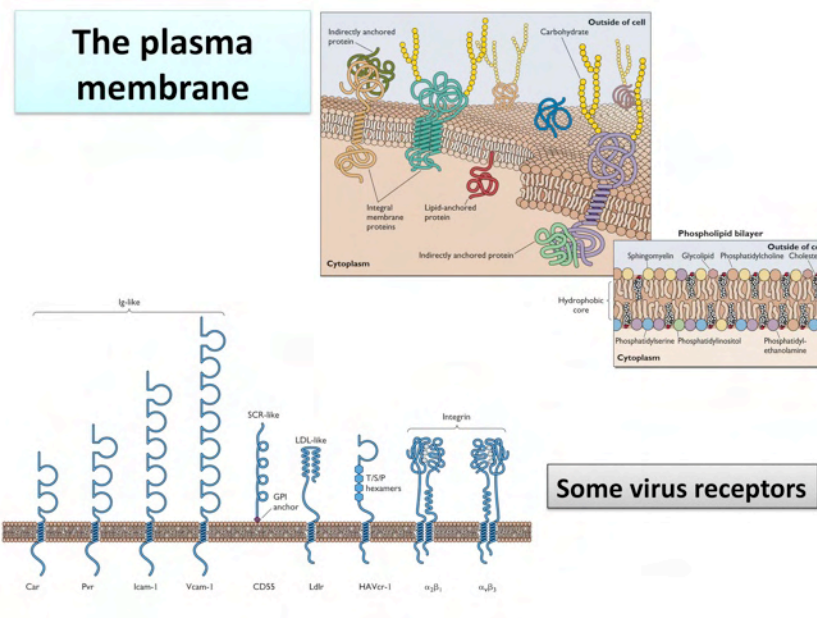
Figure 11

Entering Cells

Viral infection is initiated by a collision between the virus particle and the cell, a process that is governed by chance. Therefore, a higher concentration of virus particles increases the probability of infection. However, a virion may not infect every cell it encounters. It must come in contact with the cells and tissues in which it can replicate. Such cells are normally recognized by means of a specific interaction of a virion with a cell surface **receptor** (Fig. 11). This process can be either promiscuous or highly selective, depending on the virus and the distribution of the cell receptor. The presence of such receptors determines whether the cell will be **susceptible** to the virus. However, whether a cell is **permissive** for the replication of a particular virus depends on other, intracellular components found only in certain cell types. Cells must be both susceptible **and** permissive if an infection is to be successful.

In general, viruses have no means of locomotion, but their small size facilitates diffusion driven by Brownian movement. Propagation of viruses is dependent on essentially random encounters with potential hosts and host cells. Viral propagation is critically dependent on the production of large numbers of progeny virions with surfaces composed of many copies of structures that enable the virions to attach to susceptible cells.

Many icosahedral viruses attach to cell receptors via projections or depressions on the virion surface. An example is the attachment of poliovirus to its cell receptor, CD155 (Fig. 12). CD155 binds to a depression on the virion surface known as the 'canyon'. In contrast, the binding site for low density lipoprotein receptor on the rhinovirus type 2 virion is located on a prominent 'plateau' (Fig. 12).

Enveloped viruses typically attach to cellular receptors via viral membrane glycoproteins (Fig. 13). An example is the attachment of the influenza virus hemagglutinin glycoprotein, HA, to its cellular receptor, sialic acid (Fig. 14).
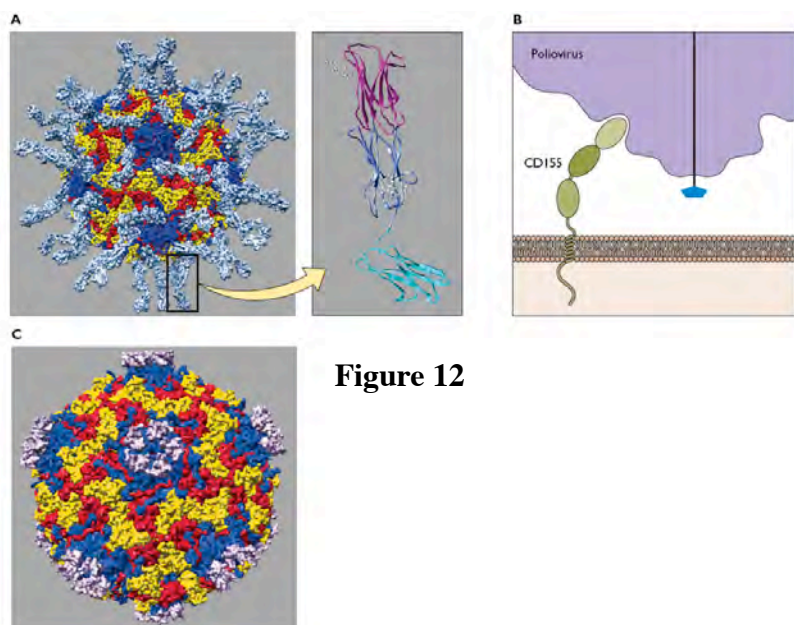
Figure 12

Each HA monomer consists of a long, helical stalk topped by a large HA1 globule, which includes the sialic acid-binding pocket (Fig. 14). Attachment of all influenza A virus strains requires sialic acid, but strains vary in their affinities for different sialic acids. For example, human virus strains are preferentially bound by sialic acids attached to galactose via an $\alpha(2,6)$ linkage, the major sialic acid present on human respiratory epithelium (Fig. 13). Avian virus strains bind preferentially to sialic acids attached to galactose via an $\alpha(2,3)$ linkage, the major sialic acid in the duck gut epithelium. Amino acids in the sialic acidbinding pocket of HA (Fig. 14) determine which sialic acid is preferred and can therefore determine viral host range. An example is the origin of the 1918 influenza virus strain, which may have evolved from an avian virus. It is believed that an amino acid change in the sialic acid-binding pocket of the avian HA allowed it to recognize the $\alpha(2,6)$-linked sialic acids that predominate in human cells.
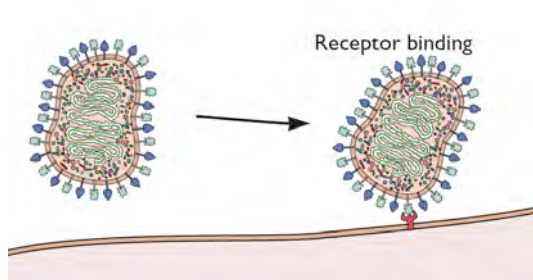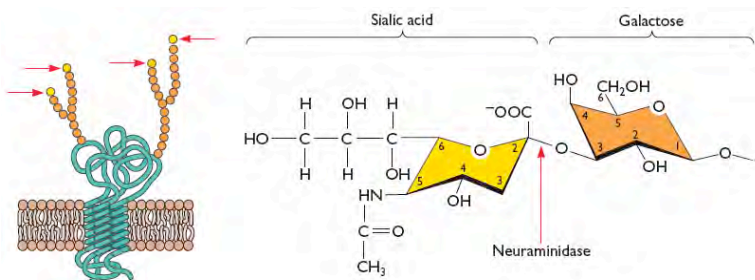

Figure 13

Figure 14

Successful entry of a virus into a host cell requires that the virus cross the plasma membrane, and often the nuclear membrane (Fig. 15). The virus particle must also disassemble to make the viral genome accessible in the cytoplasm, and the nucleic acid must be targeted to the correct cellular compartment. These are not simple processes. Cell membranes are not permeable to virus particles. Virions or critical subassemblies are brought across such barriers by specific transport pathways.

The particles of many enveloped viruses, including members of the *Paramyxoviridae* such as measles virus, fuse directly with the plasma membrane at neutral pH (Fig. 15). Once the viral and cell membranes have been closely juxtaposed by this receptor-ligand interaction, fusion is induced by a second viral glycoprotein know as fusion (F) protein, and the viral nucleocapsid is released into the cell cytoplasm.
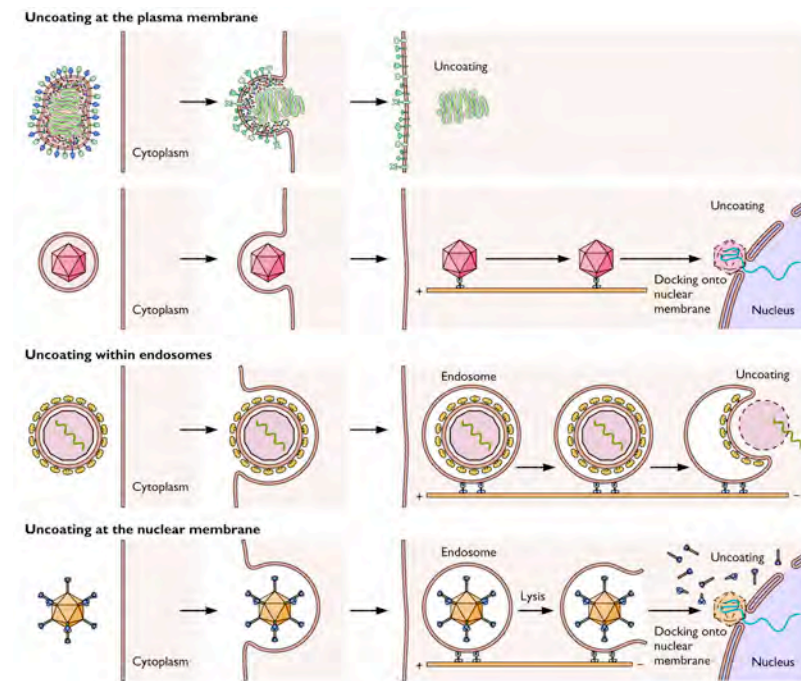

Figure 15

Many other viruses enter cells by receptor-mediated endocytosis, which is also the mechanism of uptake of molecules from the extracellular fluid by. Ligands in the extracellular medium bind to cells via specific plasma membrane receptor proteins. The receptor-ligand complex diffuses along the membrane until it reaches an invagination. Following the accumulation of receptor-ligand complexes, the pit invaginates and then pinches off to form a vesicle containing the ligand-receptor complex. Within a few seconds, the vesicles fuse with small, smooth-walled vesicles located near the cell surface, called early **endosomes**. The lumen of early endosomes is mildly acidic (pH 6.5 to 6.0). The contents of the early endosome are then transported via endosomal carrier vesicles to late endosomes located close to the nucleus. The lumen of late endosomes is more acidic (pH 6.0 to 5.0). Late endosomes in turn fuse with **lysosomes**, which are vesicles containing a variety of enzymes that degrade sugars, proteins, nucleic acids, and lipids. Viruses usually enter the cytoplasm from the early or late endosomes, and a few enter from lysosomes.

Many enveloped viruses undergo fusion within an endosomal compartment. The entry of influenza virus from the endosomal pathway is one of the best-understood viral entry mechanisms. At the cell surface, the virus attaches to sialic acid-containing receptors via the viral HA glycoprotein (Fig. 16). The virus-receptor complex is internalized by endocytosis. When the

endosomal pH reaches approximately 5.0, HA undergoes an acid-catalyzed conformational rearrangement, exposing a fusion peptide. The viral and endosomal membranes then fuse, allowing penetration of the viral RNP (vRNP) into the cytoplasm.

The fusion reaction mediated by the influenza virus HA protein is a remarkable even (Fig. 17). In native HA, the fusion peptide is joined to the three-stranded coiled-coil core by which the HA monomers interact via a 28-amino-acid sequence that forms an extended loop structure buried deep inside the molecule, about 100 Å from the globular head. In contrast, in the low-pH HA structure, this loop region is transformed into a three-stranded coiled coil. In addition, the long α-helices of the coiled coil bend upward and away from the viral membrane. The result is that the fusion peptide has moved a great distance toward the endosomal membrane. Despite these dramatic changes, HA remains trimeric and the globular heads can still bind sialic acid. In this conformation, HA holds the viral and endosome membranes 100 Å apart, too distant for the fusion reaction to occur. To bring the viral and cellular membranes closer, it is thought that the top of the acid-induced coiled coil splays apart, spreading into the lipid bilayer. The stems of the HA tilt, further facilitating close contact of the membranes.
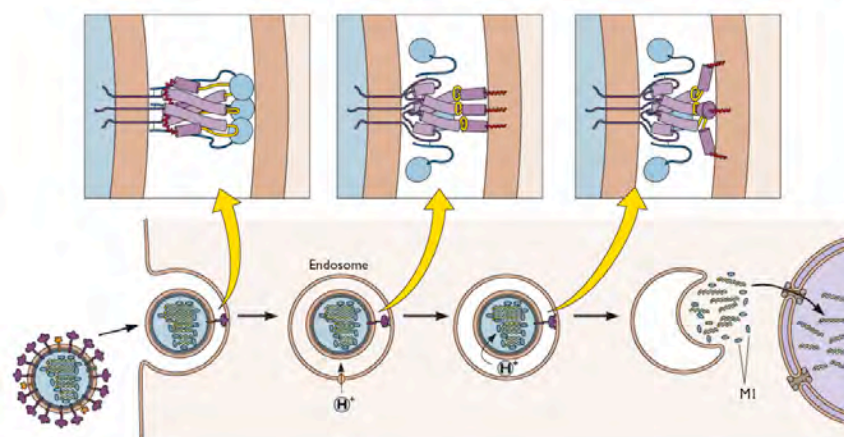

Figure 16

Uncoating: the target of the antiviral amantadines

The three-ringed symmetric amine known as amantadine was the first highly specific, potent antiviral drug effective against any virus. At low concentrations, amantadine specifically inhibits influenza A virus. The target of the drug is the M2 protein, a tetrameric, transmembrane ion channel that transports protons during virus entry (Fig. 18). The drug blocks the M2 channel, preventing acidification of the virion interior. Consequently, the viral RNPs do not dissociate after membrane fusion and cannot enter the nucleus to begin viral RNA synthesis (Fig. 16). Influenza A virus mutants resistant to amantadine arise after therapy. All have amino acid changes in the M2 trans-membrane sequences predicted to form the ion channel. The 2009 pandemic H1N1 influenza virus is resistant to amantadine.

Import of Viral Genomes into the Nucleus

The replication of most DNA viruses, and some RNA viruses including retroviruses and influenza viruses, begins in the cell nucleus. The genomes of these viruses must therefore be
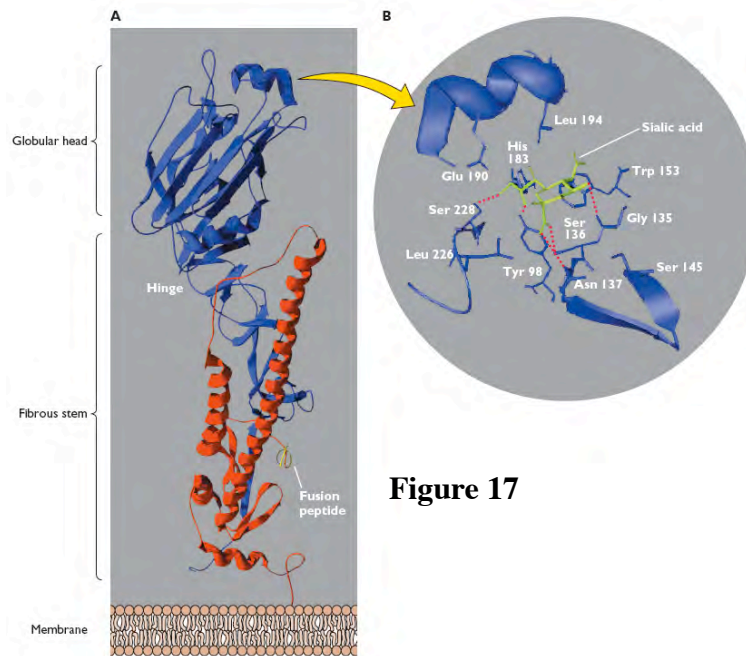
Figure 17

imported from the cytoplasm into the nucleus. One way to accomplish this movement is via the cellular pathway for protein import into the nucleus. An alternative, observed in cells infected by some retroviruses, is to enter the nucleus during cell division. At this time in the cell cycle, cellular chromatin becomes accessible to virus particles. This strategy restricts infection to cells that undergo mitosis.
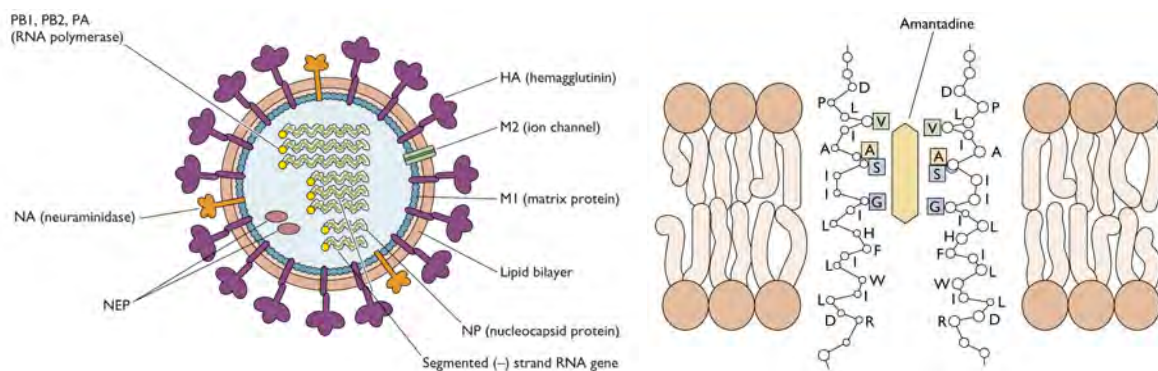

Figure 18

Some subviral particles can pass through the nuclear pore complex, but most are too large to accomplish this passage. There are several strategies to overcome this limitation (Fig. 19). The influenza virus genome, which consists of eight segments that are each small enough to pass through the nuclear pore complex, is uncoated in the cytoplasm. Adenovirus subviral particles dock onto the nuclear pore complex and are disassembled by the import machinery, allowing the viral DNA to pass into the nucleus. Herpes simplex virus capsids also dock onto the nuclear pore but remain largely intact, and the nucleic acid is injected into the nucleus through a portal in the virion.
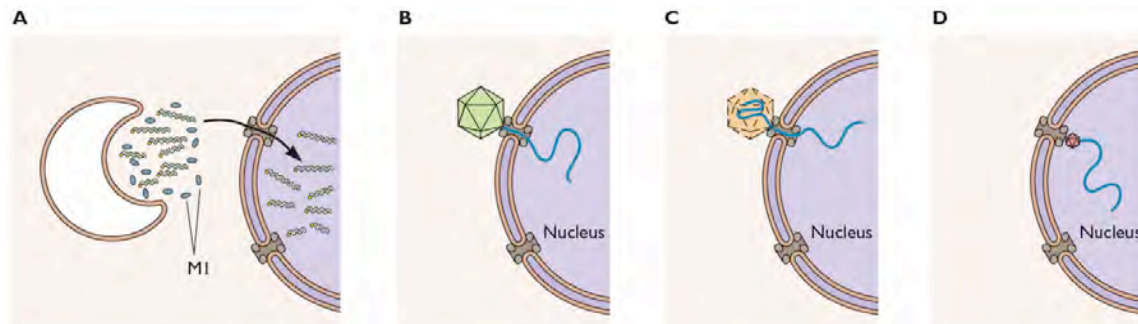
Figure 19

Forming Progeny Virions

Virus particles exhibit considerable diversity in size, composition, and structural sophistication, ranging from those comprising a single nucleic acid molecule and one structural protein to complex structures built from many different proteins and other components. Nevertheless, successful replication of all viruses requires execution of a common set of de novo assembly reactions. These processes include formation of the structural units of the protective protein coat from individual protein molecules, assembly of the coat by interaction among the structural units, incorporation of the nucleic acid genome, and release of newly assembled progeny virions.

The various components of a virion, the nucleic acid genome, capsid protein(s), and in some cases envelope proteins, are often synthesized in different cellular compartments. Their trafficking through and among the cell's compartments and organelles requires that they be equipped with the proper homing signals. Virion components must be assembled at some central location, and the information for assembly must be preprogrammed in the component molecules. The primary sequences of virion proteins contain sufficient information to specify assembly.

In concerted assembly, the structural units of the protective protein shell assemble productively only in association with the genomic nucleic acid. The nucleocapsids of (-) strand RNA viruses form by a concerted mechanism (Fig. 20), as do retrovirus particles. In sequential assembly, the genome is inserted into a preformed protein shell. The formation of herpesviral capsid is an example of this type of assembly (Fig. 21).

Release of Virus Particles

The majority of viruses leave an infected cell by one of two general mechanisms. Some viruses are released into the external environment upon budding from, or lysis of the cell. Others move directly into a new host cell without physical release of the particle, so-called cell-to-cell spread.

Many enveloped viruses assemble at, and bud from, the plasma membrane. Consequently, the final assembly reaction releases the newly formed virus particle into the extracellular milieu. The assembly of enveloped viruses is therefore mechanistically coupled with and coincident with their exit from the host cell. The egress of viruses without envelopes is frequently by destruction (lysis) of the host cell. Large quantities of assembled virions may accumulate in the infected cell, prior to release upon lysis.

Many viruses spread from one host cell to another as extracellular virions released from an infected cell (Fig. 22). Such extracellular dissemination is necessary to infect another host. Some
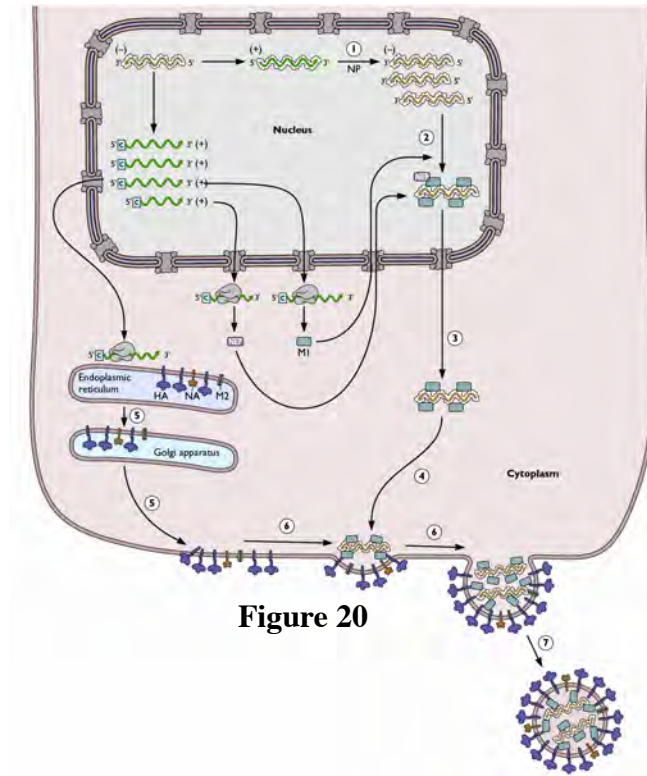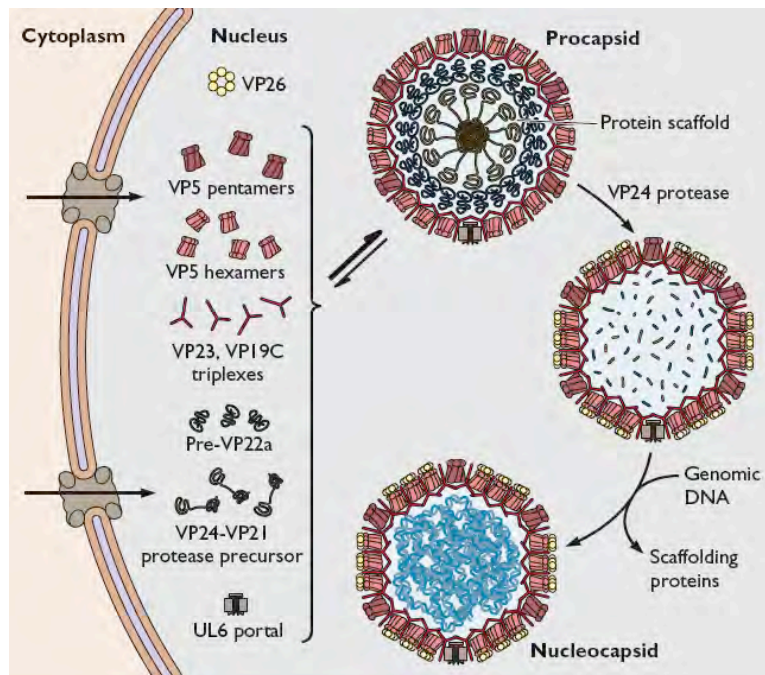
Figure 20

Figure 21

viruses, such as alphaherpesviruses, can also spread from cell to cell without passage through the extracellular environment. Such cell-to-cell spread is dependent on the viral fusion machinery. Virus spread by this mechanism avoids exposure to host defense mechanisms that target extracellular virus.
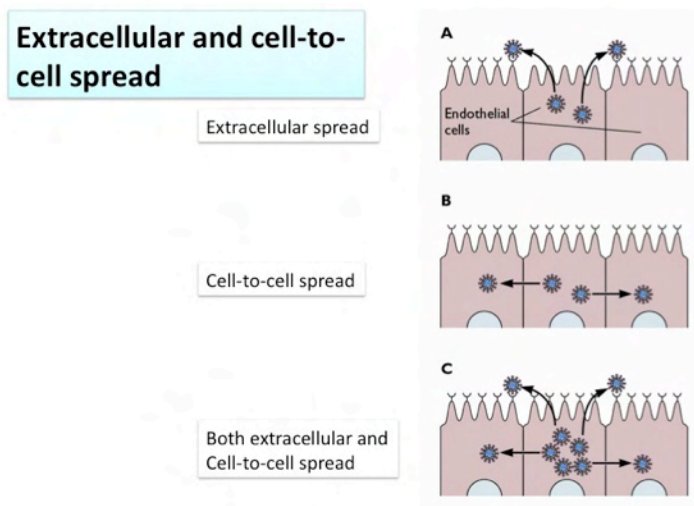
Figure 22.

Function of NA in Release of Virions from Cells

The influenza viral NA glycoprotein plays a role during release of newly budded virions from the cell surface. As virus particles are produced, they immediately bind sialic acid receptors on the cell surface. The viral NA is an enzyme that cleaves sialic acid residues from the cell surface, allowing release of virions (Fig. 23). Sialic acid fits in a pocket on the head of the NA that contains the enzyme active site. If the viral NA activity is blocked, virions remain on the cell surface after budding and the infection cannot spread.
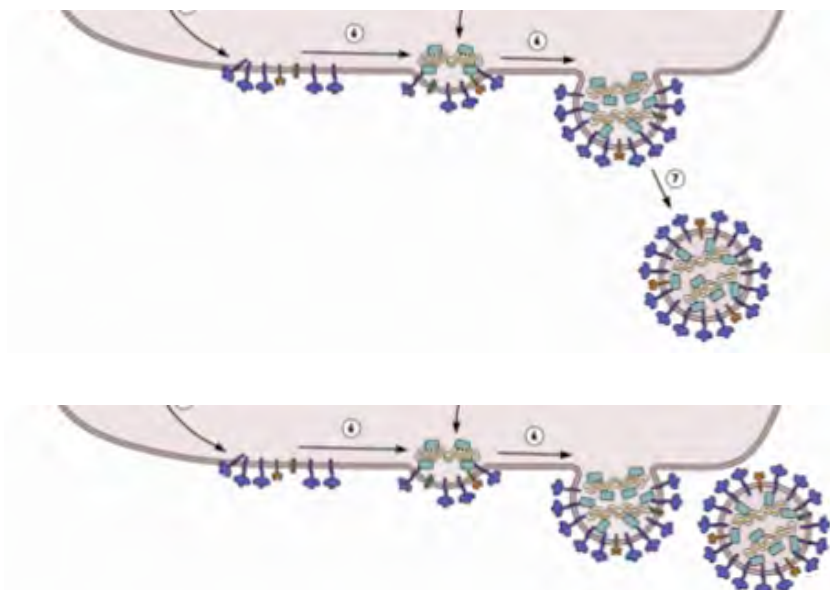

Figure 23

Recently developed NA inhibitors Tamiflu and Relenza function by blocking the activity of the viral NA glycoprotein. Both drugs fit into the sialic-acid binding pocket on the NA, preventing entry of sialic acid (Fig. 24). Consequently sialic acid cannot be cleaved from the cell surface, and newly budded virions cannot spread from cell to cell. When used within 48 h of onset of symptoms, the drugs reduce the median time to alleviation of symptoms by about 1 day compared to placebo. When used within 30 h of disease onset, the drugs reduce the duration of symptoms by about 3 days.

Tamiflu resistant influenza virus mutants have been identified; the NA of these mutants have amino acid changes that prevent binding of the drug (Fig. 25).
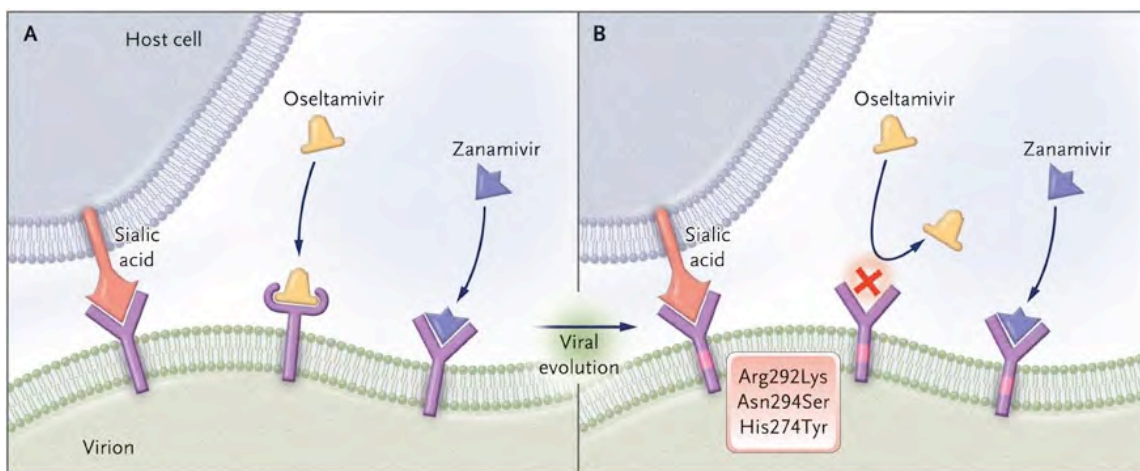

Figure 24