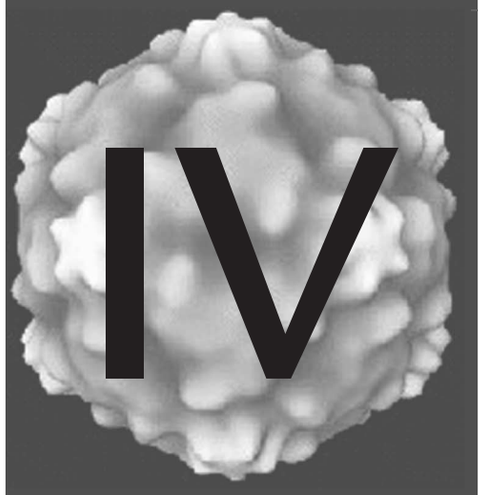
Replication Patterns of Specific Viruses



PART IV

- REPLICATION OF POSITIVE-SENSE RNA VIRUSES
- REPLICATION OF POSITIVE-SENSE RNA VIRUSES WHOSE GENOMES ARE TRANSLATED AS THE FIRST STEP IN GENE EXPRESSION
- POSITIVE-SENSE RNA VIRUSES ENCODING A SINGLE LARGE OPEN READING FRAME
- POSITIVE-SENSE RNA VIRUSES ENCODING MORE THAN ONE TRANSLATIONAL READING FRAME
- REPLICATION OF PLANT VIRUSES WITH RNA GENOMES
- REPLICATION OF BACTERIOPHAGE WITH RNA GENOMES
- REPLICATION STRATEGIES OF RNA VIRUSES REQUIRING RNA-DIRECTED mRNA TRANSCRIPTION AS THE FIRST STEP IN VIRAL GENE EXPRESSION
- REPLICATION OF NEGATIVE-SENSE RNA VIRUSES WITH A MONOPARTITE GENOME
- INFLUENZA VIRUSES – NEGATIVE-SENSE RNA VIRUSES WITH A MULTIPARTITE GENOME
- OTHER NEGATIVE-SENSE RNA VIRUSES WITH MULTIPARTITE GENOMES
- VIRUSES WITH DOUBLE-STRANDED RNA GENOMES
- SUBVIRAL PATHOGENS
- REPLICATION STRATEGIES OF SMALL AND MEDIUM-SIZED DNA VIRUSES
- PAPOVAVIRUS REPLICATION

Replication of Positive-sense RNA Viruses

CHAPTER

RNA viruses – general considerations

By definition, RNA viruses use RNA as genetic material and thus, must use some relatively subtle strategies to replicate in a cell since the cell uses DNA. Ultimately, to express its genetic information, any virus must be able to present genetic information to the cell as translatable mRNA, but the way this happens with RNA viruses will depend on the type of virus and the nature of the encapsidated RNA.

According to Watson-Crick base-pairing rules, once the sequence of one strand of either RNA or DNA is known, the sequence of its complementary strand can be inferred. The complementary strand serves as a template for synthesis of the strand of RNA or DNA in question. While the sequence of a strand of RNA is in a sense equivalent to its complement, the actual "sense" of the information encoded in the virion RNA is important for understanding how the virus replicates. As noted in Chapter 1, viral mRNA is the obligate first step in the generation of viral protein; therefore, an RNA virus must be able to generate something that looks to the cell like mRNA before its genome can be replicated.

The ways that viruses, especially RNA viruses, express their genomes as mRNA, of necessity, are limited and form an important basis of classification. The use of this criteria in the Baltimore classification of viruses was outlined in Chapter 5. The fundamental basis of this classification for RNA viruses is whether the viral genome can be directly utilized as mRNA or whether it must first be transcribed into mRNA. This classification breaks RNA viruses that do not utilize a DNA intermediate (an important exception) into two basic groups: the viruses containing mRNA as their genomes and those that do not. This second group, which comprises the viruses encapsidating an RNA genome that is complementary (antisense) to mRNA and the viruses that encapsidate a double-stranded (ds) RNA genome, requires the action of a specific viral-encoded transcriptase. Such viral transcriptases are contained in the virion as a structural protein, and utilize the virion genomic RNA as a template for transcription.

The basic strategy for the initiation of infection by these two groups of viruses, members of which are described in some detail in this and the following chapter, is outlined in Fig. 15.1a.

This classification ignores a very significant complication: It makes no accommodation for the fact that a very important group of viruses with genomes that can serve as mRNA use DNA as the intermediate in their replication. These are the retroviruses. These viruses and their relatives use a very complex pattern of viral-encoded and cellular functions in their replication, and are described only after a full survey of the "simpler" RNA and DNA viruses is presented.

A general picture of RNA-directed RNA replication

With the exception of retroviruses and some unusual viruses related to viroids, single-stranded (ss) RNA virus genome replication requires two stages; these are shown in Fig. 15.1b. First, the input strand must be transcribed (using Watson-Crick base-pairing rules) into a strand of complementary sequence and **opposite polarity**. Replication occurs as a "fuzzy," multibranched structure. This complex, dynamic structure contains molecules of viral transcriptase (replicase), a number of partially synthesized product RNA strands ("nascent" strands), and the genome-sense template strand. The whole **ribonucleoprotein (RNP)** complex is termed the type 1 **replicative intermediate** or **RI-1**. The single-stranded products generated from RI-1 are antisense to the genomic RNA.

This complementary strand RNA serves as a template for the formation of more genomic-sense RNA strands. This second replicative intermediate (**RI-2**) is essentially the same in structure as RI-1 except that the template strand is of opposite sense to genomic RNA and the nascent product RNA molecules are of genome sense.

Remember:

Virion RNA is the template in RI-1.

RI-1 produces template RNA of opposite sense to virion RNA.

RNA that is complementary to virion RNA is the template in RI-2.

RI-2 is the intermediate for expression of RNA of the same sense as the virion.

One further general feature of the replication of RNA viruses is worth noting. The **error frequency** (i.e., the frequency of incorporating an incorrect base) of RNA-directed RNA replication is quite high compared to that for dsDNA replication. Thus, typically DNA-directed DNA repli-

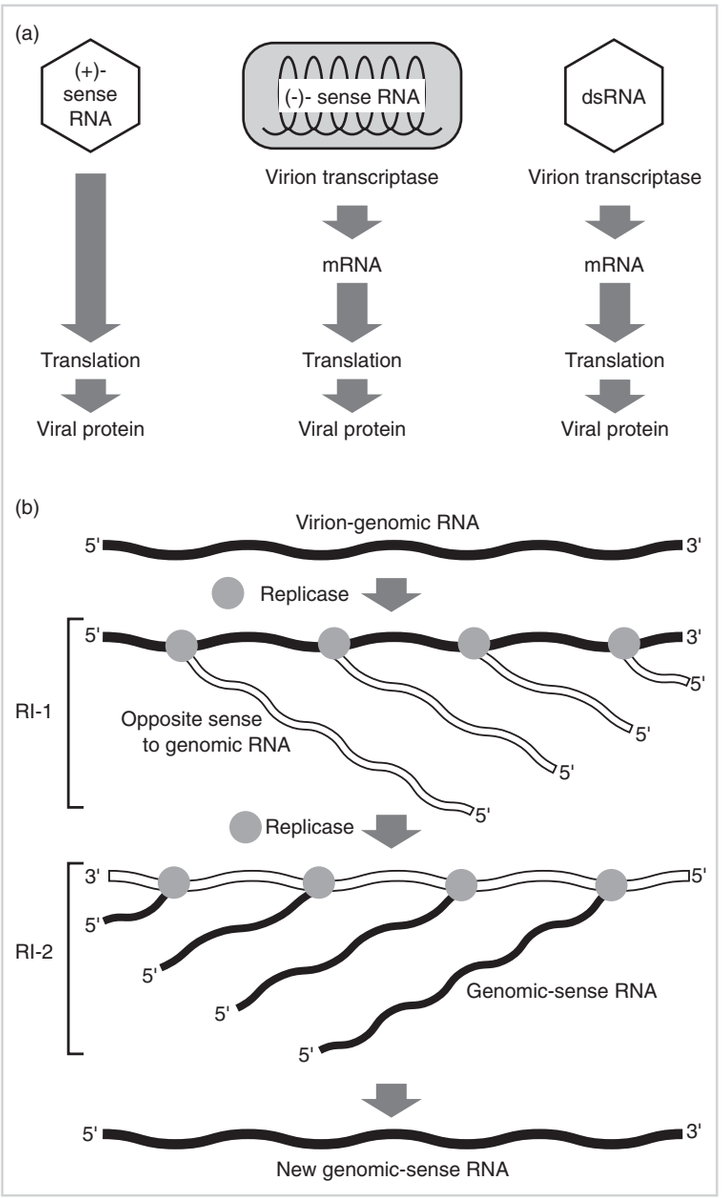


Fig. 15.1 Some general features of viruses containing RNA genomes that use RNA-directed RNA transcription in their replication. *a*. The general relationship between viruses containing a genome that can be translated as the first step in the expression of viral genes versus those viruses that first must carry out transcription of their genome into mRNA utilizing a virion-associated transcriptase. *b*. The basic rules for RNA-directed RNA replication. As with DNA-directed RNA and DNA synthesis, the new (nascent) strand is synthesized 5′ to 3′ antiparallel to the template, and the Watson-Crick base-pairing rules are the same, with U substituting for T. However, the very high thermal stability of dsRNA leads to complications. The major complication is that newly synthesized RNA must be denatured and removed from the template strand to avoid its "collapsing" into a double-stranded form. Formation of such dsRNA is an effective inducer of interferon (see Chapter 8), and it appears to be refractory to serving as a template when free in the cytoplasm. A second complication is that in order to generate an ssRNA molecule of the same coding sense as the virion genome, *two* replicative intermediates (RI's) must be generated. These intermediates are dynamic structures of ribonucleoprotein containing a full-length template strand, and a number of newly synthesized product RNA molecules growing from virion-encoded replicase that is traversing the template strand. RI-1 generates RNA complementary to the virion genomic RNA. This serves as a template for new virion genome RNA in RI-2.

cation leads to incorporation of one mismatched base per 10^7 to 10^9 base pairs, while RNA-directed RNA synthesis typically results in one error per 10^5 bases. Indeed, the error rate in the replication of some RNA genomes can be as high as one error per 10^4 nucleotides.

Part of the reason for this error rate for RNA is that there is no truly double-stranded intermediate; therefore, there is no template for error correction or "proofreading" of the newly synthesized strand as there is in DNA replication. A second reason is that RNA polymerases using RNA templates seem to have an inherently higher error frequency than those utilizing DNA as a template.

For these reasons, infection of cells with many RNA viruses is characterized by the generation of a large number of progeny virions bearing a few or a large number of genetic differences from their parents. This high rate of mutation can have a significant role in viral pathogenesis and evolution;

further, it provides the mechanistic basis for the generation of defective virus particles described in Chapter 14. Indeed, many RNA viruses are so genetically plastic that the term **quasi-species swarm** is applied to virus stocks generated from a single infectious event, as any particular isolate will be, potentially at least, genetically significantly different from the parental virus.

REPLICATION OF POSITIVE-SENSE RNA VIRUSES WHOSE GENOMES ARE TRANSLATED AS THE FIRST STEP IN GENE EXPRESSION

The first step in the infectious cycle of this group of positive-sense RNA viruses (also called **positive (+) strand viruses**) leading to expression of viral proteins is *translation of viral protein*. If the virion (genomic) RNA is incubated with ribosomes, transfer RNA (tRNA), amino acids, ATP, GTP, and the other components of an in vitro protein synthesis system, protein will be synthesized.

Further, if virion RNA is transfected into the cell in the absence of any other viral protein, infection will proceed and new virus will be produced. This can occur in the laboratory provided there are proper precautions to protect the viral RNA, which is chemically labile.

Positive-sense RNA viruses (other than retroviruses) do not require a transcription step prior to expression of viral protein. This means that the nucleus of a eukaryotic cell is either somewhat or completely superfluous to the infection process. All the replication steps can take place more or less efficiently in a cell from which the nucleus is removed.

For instance, removal of the nucleus can be accomplished in poliovirus infections by use of a drug, **cytochalasin B**, which breaks down the actin-fiber cytoskeleton that anchors the nucleus inside the cell. Cells treated with this drug can be subjected to mild centrifugal force, causing the nucleus to "pop" out of the cell. Such enucleated cells can be infected with poliovirus and new virus synthesized at levels equivalent to those produced in normal nucleated cells.

A very large number of positive-sense RNA viruses can infect bacteria, animals, and especially plants, and the patterns of their replication bear strong similarities. The replication patterns of the positive-sense RNA important to human health can be outlined by consideration of just a few, if the replication of retroviruses is considered separately.

A basic distinction between groups of positive-sense RNA viruses involves whether the viral genome contains a single open translational reading frame (**ORF**) as defined in Chapter 13, or multiple ones. This difference correlates with the complexity of mRNA species expressed during infection.

POSITIVE-SENSE RNA VIRUSES ENCODING A SINGLE LARGE OPEN READING FRAME

Picornavirus replication

Picornaviruses are genetically simple and have been the subject of extensive experimental investigation owing to the number of diseases they cause. Their name is based on a pseudoclassical use of Latin mixed with modern terminology: *pico* ("small")-RNA-virus.

The replication of poliovirus (the best-characterized picornavirus, and perhaps, best-characterized animal virus) provides a basic model for RNA virus replication. Studies on poliovirus were initiated because of the drive to develop a useful vaccine against paralytic poliomyelitis. These studies successfully culminated in the late 1950s and early 1960s. Protocols developed for replicating the virus in cultured cells formed the basis for successful vaccine development and production.

At the same time, the relative ease of maintaining the virus and replicating it in culture led to its early exploitation for molecular biological studies. It is still a favored model.

Other closely related picornaviruses include rhinoviruses and hepatitis A virus. These replicate in a generally similar manner, as do a number of positive-sense RNA–containing bacterial and plant viruses. Indeed, close genetic relationships among many of these viruses are well established.

The poliovirus genetic map and expression of poliovirus proteins

A schematic of the icosahedral poliovirus virion is shown in Fig. 15.2. In accordance with its classification as a positive-sense RNA virus, the poliovirus genomic RNA isolated from purified virions is mRNA sense and acts as a viral mRNA upon infection. Full characterization and sequence analysis established that the genome is 7741 bases long with a very long (743-base) leader sequence between the 5′ end of the mRNA and the (ninth!) AUG, which initiates the beginning of an ORF extending to a translation termination signal near the 3′ end. There is a short untranslated trailer following the 7000 base ORF, and this is followed by a polyA tract. The polyA tail of the poliovirus mRNA is actually part of the viral genome; therefore, it is not added posttranscriptionally as with cellular mRNA (see Chapter 13). A simple genetic map of the viral genome is shown in Fig. 15.2.

While poliovirus RNA *is* mRNA and can be translated into protein in an in vitro translation system, it has two properties quite different from cellular mRNA. First, poliovirus virion RNA has a protein VPg at its 5′ end instead of the methylated cap structure found in cellular mRNA. The VPg protein is encoded by the virus. The viral mRNA also has a very long leader that can assume a complex structure by virtue of intramolecular base pairing in solution. The structure of this leader sequence, especially near the beginning of the translational reading frame (the **internal ribosome entry site [IRES]**), mediates association of the viral genome with ribosomes. With poliovirus RNA, the normal Kozak rules for the selection of the AUG codon to initiate translation in an mRNA (see Chapter 13) do not apply. Indeed, the AUG triplet that begins the large poliovirus ORF is preceded by eight other AUG triplets within the leader that are not utilized to initiate translation.

Upon successful initiation of infection, viral genomic mRNA is translated into a single large protein that is the precursor to all viral proteins. This precursor protein is also shown in Fig. 15.2; it contains all the poliovirus proteins that are expressed during infection. Thus, all the viral proteins such as those shown in Fig. 12.1 are derived from it.

The smaller proteins are cleaved from the precursor polyprotein by means of two proteases (2A and 3C) that comprise part of this large viral protein. As briefly outlined in Chapter 6, many viruses utilize proteolytic cleavage of large precursor proteins via virus-encoded proteases during the replication process, and such proteases are important potential targets for antiviral chemotherapy (see Chapter 8). Indeed, the development of protease inhibitors is having a very encouraging effect on attempts to treat AIDS.

The steps in processing are complex, and have yet to be fully worked out in complete detail. Both viral proteases utilize a cysteine residue as part of their active sites; thus, they are termed **C-proteases**. They exhibit a very high specificity, and although both cleave the precursor peptide at sites between specific amino acids (Tyr-Gly for protease 2A and Gln-Gly for protease 3C), neither cleaves all available sites and protease 2A does not cleave nonviral peptides with any efficiency at all. Clearly, secondary structure and other features of the substrate protein are important in determining cleavage sites.

The first two cleavages take place intramolecularly, that is, within the protein in which the proteases are covalently linked. These cleavages result in the formation of three large precursor proteins, P1, P2, and P3. Protein P1 contains the capsid proteins, VP1, VP3, and VP0, as well as a short leader protein (L) of unknown function. The P2 and P3 proteins are precursors for a number of

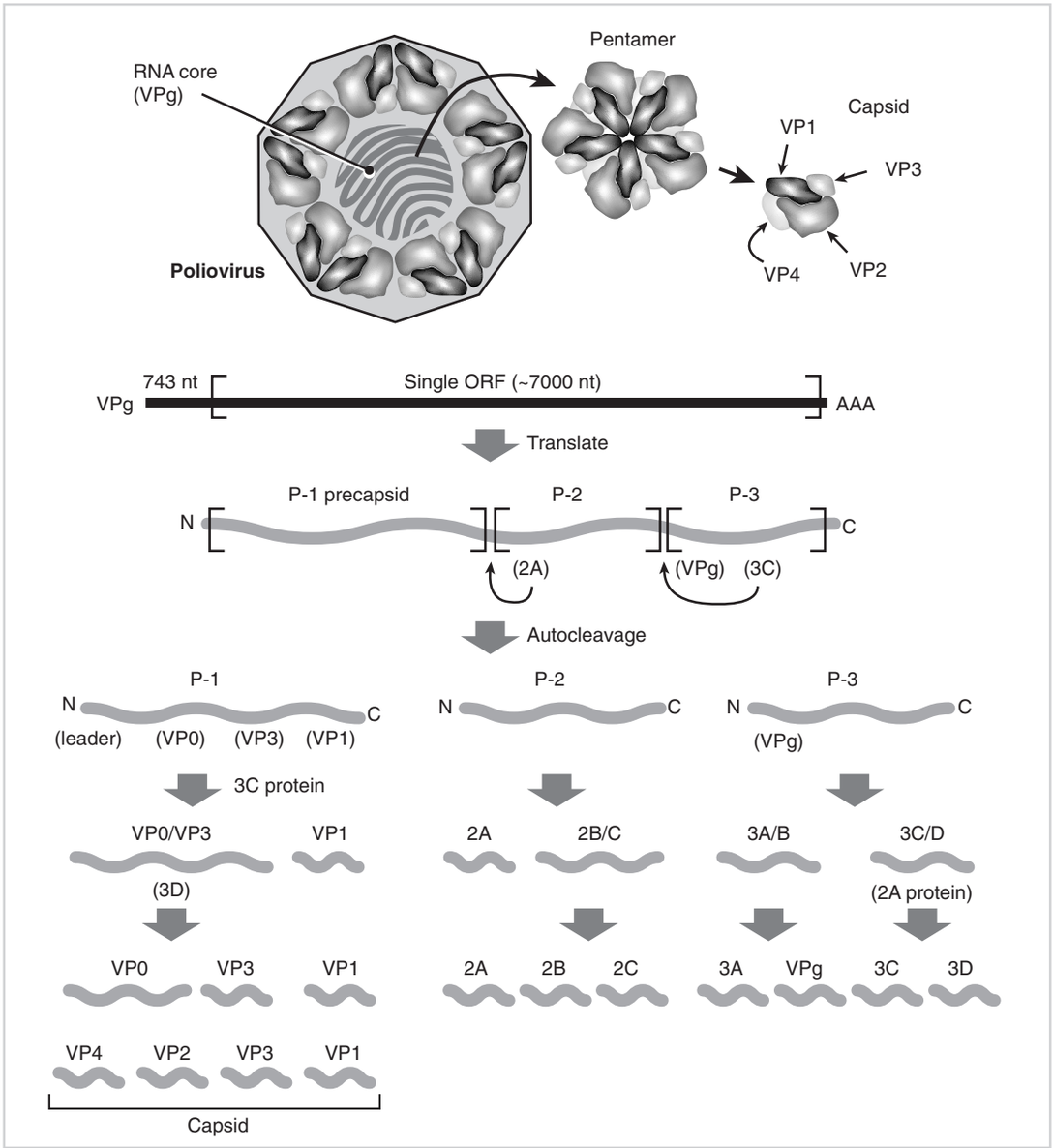


Fig. 15.2 Poliovirus, a typical picornavirus. The 30 nm diameter icosahedral capsid comprises 60 identical subunits—each a pentamer of subunits (often called protomers) containing a single copy of VP1, VP2, VP3, and VP4. The map of the approximately 7700 nucleotide (nt) single-stranded RNA genome that serves as mRNA in the initial stages of replication is also shown. Unlike cellular mRNA, poliovirus genomic RNA has a viral protein (VPg) at its 5′ end instead of a methylated nucleotide cap structure. The RNA has a ca. 740 nt sequence at the 5′ end that encodes no protein, but assumes a complex secondary structure to aid ribosome entry and initiation of the single translational reading frame. The single precursor protein synthesized from the virion RNA is cleaved by internal proteases (2A and 3C) initially into three precursor proteins, P1, P2, and P3. Protein P1 is then proteolytically cleaved in a number of steps into the proteins that assemble into the precapsid, VP0, VP1, and VP3. Proteins P2 and P3 are processed into replicase, VPg, and a number of proteins that modify the host cell, ultimately leading to cell lysis. With three exceptions, all proteolytic steps are accomplished by protease 3C, either by itself or in association with protein 3D. Protease 2A carries out the first cleavage of the precursor protein into P1 and P2 as an intramolecular event. It also mediates cleavage of the protease 3CD precursor into protease 3C and protein 3D. It is not known how the third cleavage that does not utilize protease 3C occurs. This is the maturation of the capsomers by the cleavage of VP0 into VP2 and VP4.

nonstructural proteins, including the viral replicase enzyme and proteins and enzymes that alter structure of the infected cell. Protein P3 also contains the VPg protein. The general steps in derivation of mature viral proteins from the precursor protein are shown in the genetic map of Fig. 15.2.

The later stages in processing of the precursor proteins involve mainly protease 3C, although protease 2A cleaves the 3CD precursor of protease and replicase into variants then termed 3C′ and 3D′. It is not known whether these have any role in replication, and they are not seen in infections with all strains of the virus. While protein 3D is not a protease (it is the replicase protein), it aids in cleavage of the VP0-VP3 precursor into VP0 and VP3. The 3CD precursor itself, however can also act as a protease, and may have a specific role in some of the early cleavage events.

Since the poliovirus ORF is translated as a single, very large protein, poliovirus technically has only one "gene." This is not strictly true, however, since different portions of the ORF contain information for different types of protein or enzyme activities. Further, different steps in processing of the precursor proteins are favored at different times in the replication cycle; therefore, the pattern of poliovirus proteins seen varies with time following infection, as shown earlier in Fig. 12.1.

The demonstration of precursor–product relationships between viral proteins can be tricky and experimentally difficult, but the procedure's theory is simple and based on analysis of proteins encoded by the virus, consideration of the virus's genetic capacity to encode proteins, and a general understanding of the translation process itself. The separation and enumeration of viral proteins based on their migration rates in denaturing gels, which is a function of protein size, are outlined in Chapter 12, and estimates of protein coding capacity based on genome size are described in Chapter 14.

For poliovirus, many years of analysis can be summarized as follows: The total molecular size of the proteins encoded by the virus cannot exceed approximately 2300 amino acids (7000/3). Despite this, the total size of viral proteins estimated by adding radioactive amino acids to an infected cell and then performing size fractionation on the resulting radiolabeled material is significantly greater. Further, it is known that poliovirus efficiently inhibits cellular protein synthesis, so most proteins detected by the addition of radioactive precursor amino acids to infected cells (also termed a *pulse* of radioactive material) are, indeed, viral.

This conundrum can be resolved by using a technique called a **pulse-chase experiment**, and by using *amino acid analogues*, which inhibit protease processing of the precursor proteins. In pulse-chase experiments, radioactive amino acids are added for a short time. This is the "pulse." Then a large excess of nonradioactive amino acids is added to dilute the label. This is the "chase."

Only the largest viral proteins isolated from a poliovirus-infected cell exposed only to the radioactive pulse for short periods (followed by isolation of the infected cell) had radioactivity. This finding suggests that these proteins are the first viral products synthesized. If the pulse period is followed by chase periods of various lengths, radioactivity is eventually seen in the smaller viral proteins. Such a result is fully consistent with a kinetic precursor-product relationship between large (precursor) proteins and smaller mature (product) viral proteins.

The relationship between precursor and product was confirmed by adding translation inhibitors at specific times following a pulse of radioactive amino acids. This step resulted in the loss of label incorporated into large proteins, but did not affect the appearance of label in the smaller proteins derived from the precursor proteins already labeled during the pulse. Finally, addition of amino acid analogues that inhibited proteolysis of the precursor protein contributed a further confirmation of the process.

The poliovirus replication cycle

As shown in Fig. 15.3, everything tends to "happen at once" during the poliovirus replication cycle. Viral entry involves attachment of the virions by association with the cellular receptor. As described in Chapter 6 (see Fig. 6.2), this leads to the formation of a coated pit into which the capsid is

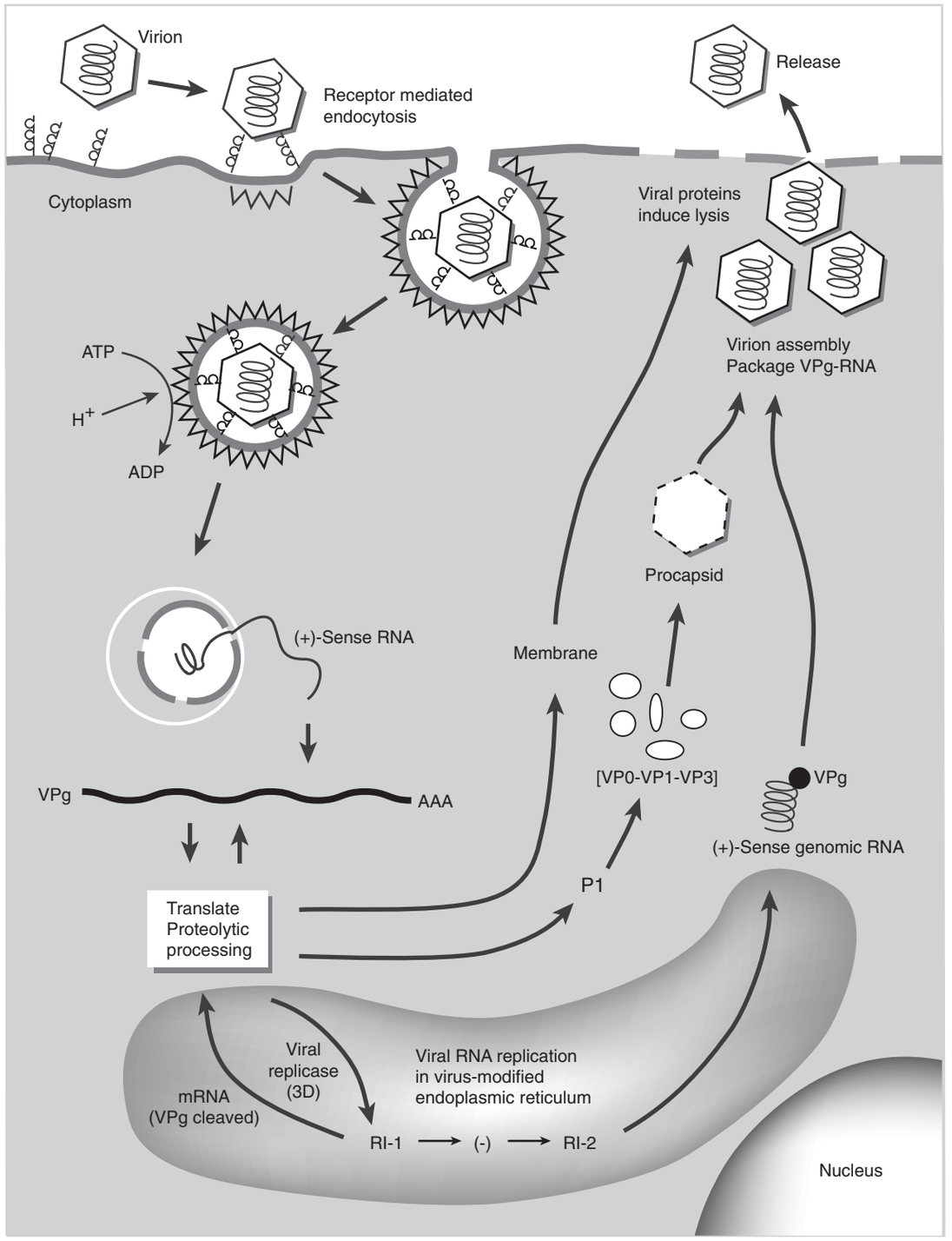


Fig. 15.3 The poliovirus replication cycle. The schematic representation is broken into discrete steps. Viral entry is by receptor-mediated endocytosis during which the virion proteins are sequentially removed, releasing virion-associated positive-sense RNA. This RNA is translated into a large polyprotein. Viral replicase released from the precursor protein then mediates generation of RI-1 and RI-2 to generate more mRNA that, unlike the original genomic RNA, has the VPg protein cleaved off. As infection proceeds, the replication complexes become associated with cellular membrane structures into replication compartments. Newly synthesized positive-sense RNA is also translated and the process repeats many times until sufficient capsid protein precursors are formed to allow assembly of the procapsid. Procapsids associate with newly synthesized positive-sense RNA still containing VPg at its 5′ end, and entry of viral genomes results in capsid maturation. As the process continues, virions accumulate in the cytoplasm until viral proteins induce cell lysis and virus release occurs. The entire process can take place in the absence of a nucleus.

engulfed and transported into the infected cell's cytoplasm. The acidic environment of the endocytotic vesicle leads to a change in conformation of capsid proteins, leading to loss of VP4 and insertion of the RNA genome into the cytoplasm through the vesicle's membrane. Viral RNA is translated into protein, portions of which are involved in replication of the viral genome by generation of RI-1 and RI-2. The protein VPg is a primer for this replication. Poliovirus replicase, protein $3D^{pol}$, catalyzes the generation of both negative- and positive-sense products. It has recently been demonstrated that cis-acting sequence elements that control replication are present in the poliovirus genome. Secondary structure features at the 5′ end as well as within coding regions appear to be required for efficient RNA replication. Other poliovirus proteins are also involved, as well as one or more host proteins, since much of the viral genome's replication takes place in membrane-associated compartments generated by these proteins within the infected cell's cytoplasm. Generation of new mRNA sense (positive) strands of poliovirus RNA leads to further translation, further replication, and finally, capsid assembly and cell lysis.

Details of the poliovirus capsid's morphogenesis were worked out several decades ago. While there is still some controversy concerning the timing of certain steps in the assembly process (especially the timing of the association of virion RNA with the procapsids), poliovirus assembly serves as a model for such processes in all icosahedral RNA viruses (see Chapter 6). Proteolytic cleavage of precursor proteins plays an important role in the final steps of maturation of the capsid. This cleavage does not involve the action of either protease 2A or 3C. Rather, it appears to be an intramolecular event mediated by the capsid proteins themselves as they assemble and assume their mature conformation. The molecular sizes of the poliovirus capsid proteins are given in Table 11.1.

The most generally accepted scheme is shown in Fig. 15.4. In viral morphogenesis, P1 protein is cleaved from the precursor protein by protease 2A segment. Five copies of this protein aggregate and the protein are further cleaved by protease 3C into VP0, VP1, and VP3, which forms one of the 60 capsid *protomers*. Five of these protomers assemble to form the 14s pentamer. Finally, twelve of these 14s pentamers assemble to form an empty capsid (**procapsid**).

This procapsid is less dense than the mature virion, so its proteins can be separated readily by centrifugation. Analysis of the procapsid proteins demonstrates equimolar quantities of VP0, VP1, and VP3. Following formation of the procapsid, viral RNA associates with the particle and a final cleavage of VP0 into VP2 and VP4 occurs to generate the mature virion. After virions are assembled, the cell lyses and virus is released.

Picornavirus cytopathology and disease

The most obvious cytopathology of poliovirus replication is cell lysis. But prior to this, the virus specifically inhibits host cell protein synthesis. Inhibition of host cell protein synthesis involves proteolytic digestion of the translation initiation factor eIF-4G so that ribosomes can no longer recognize capped mRNA (see Chapter 13). Such modification leads to the translation of only uncapped poliovirus mRNA because its IRES allows it to assemble the translation complex with the virus-modified ribosomes. Note that this rather elegant method of shutoff will not work with most types of viruses because they express and utilize capped mRNA!

There are three related types, or serotypes, of poliovirus. They differ in the particular antigenic properties of viral structural proteins. Most poliovirus infections in unprotected human populations result in no or only mild symptoms, but one serotype (type 3) is strongly associated with the disease's paralytic form. Infection with this serotype does not invariably lead to a paralytic episode, but the probability of such an episode is much higher than with the others. All serotypes are distributed throughout the regions where poliovirus is endemic in a population, although some predominate in some locations.

Poliovirus is spread by fecal contamination of food or water supplies. Receptors for the virus are found in the intestine's epithelium, and infection results in local destruction of some tissue in the

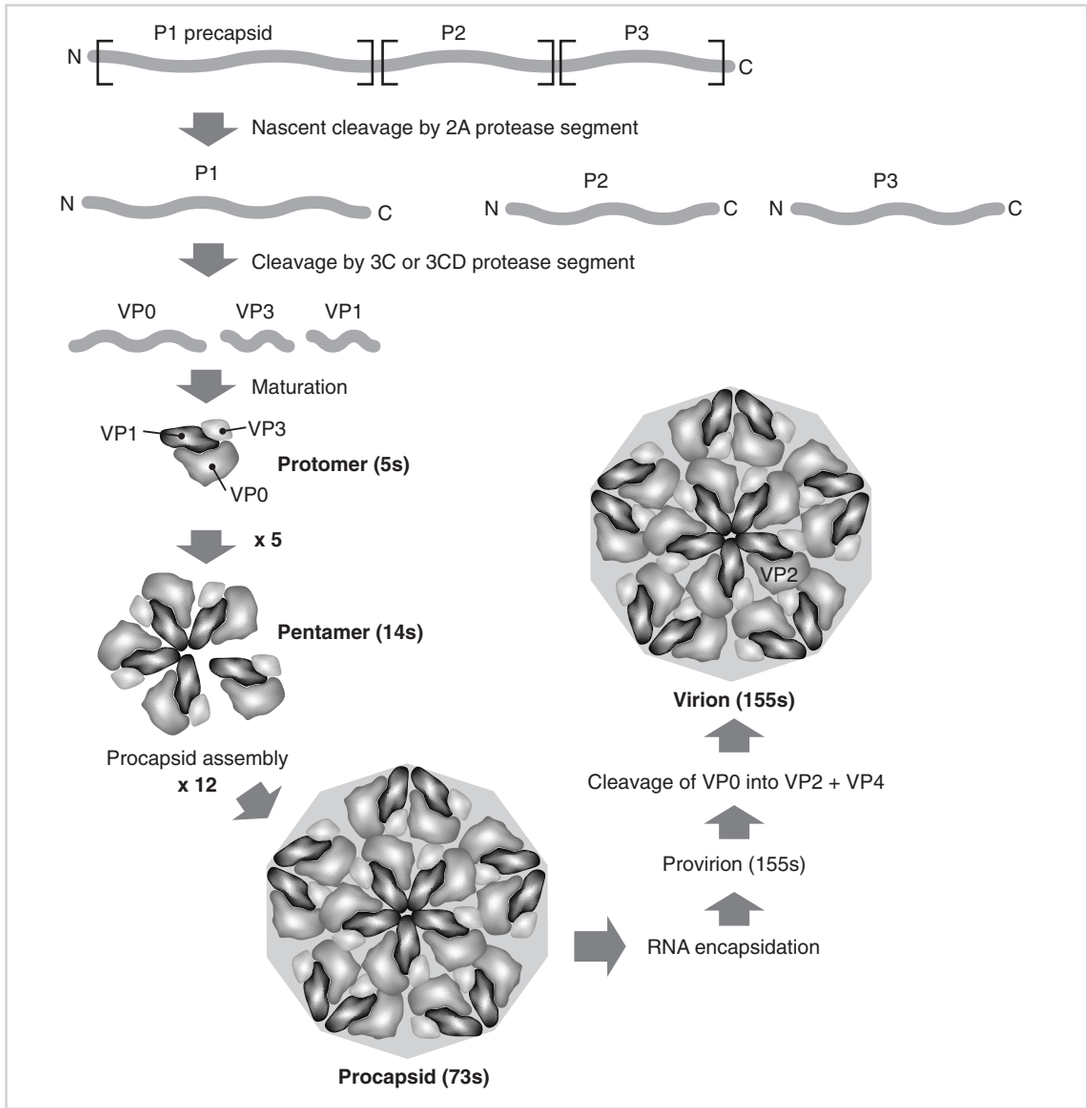


Fig. 15.4 The steps in the assembly of the poliovirus virion. Precursor proteins associate to form 5s protomers, which then assemble to form pentamers. Twelve of these assemble to form the procapsid into which virion RNA is incorporated. Final cleavage of VP0 into VP2 and VP4 takes place to form the mature capsid that has a diameter of 28 to 30 nm.

intestine, which can result in diarrhea. Unfortunately, motor neurons also have receptors for poliovirus, and if the virus gets into the bloodstream, it can replicate in and destroy such neurons, leading to paralysis. This result is of no value to the virus since the virus initiating neuronal infection cannot be spread to other individuals and is eventually cleared; thus, the paralytic phase of the disease is a "dead end" for the virus. The virus stimulates an immune response and the individual recovers and is resistant or immune to later infection.

Vaccination against poliovirus infections is accomplished effectively with both inactivated and attenuated live-virus vaccines, as described in Chapter 8. Since the only reservoir of poliovirus is humans, immunity through vaccination against the virus is an effective way of preventing disease.

Currently, a major effort is underway to completely eradicate the disease from the environment (see Chapter 22).

A number of other picornaviruses cause disease; many are spread by fecal contamination and include hepatitis A virus, echoviruses, and coxsackievirus. Like poliovirus, these viruses occasionally invade nervous tissue. Coxsackievirus generally causes asymptomatic infections or mild lesions in oral and intestinal mucosa, but can cause encephalitis. Echoviruses are associated with enteric infections also, but certain echovirus serotypes cause infant nonbacterial meningitis, and some epidemic outbreaks with high mortality rates in infants have been reported.

Another widespread group of picornaviruses are the rhinoviruses, one of the two major groups of viruses causing common head colds. Unlike the other picornaviruses detailed here, rhinoviruses are transmitted as aerosols. Because of the large number (~ 100) of distinct serotypes of rhinovirus it is improbable that an infection will generate immunity that prevents subsequent colds. There are no known neurological complications arising from rhinovirus infections.

Flavivirus replication

The success and widespread distribution of picornaviruses and their relatives demonstrate that the replication strategy found in translation of a single large ORF is a very effective one. If more evidence were needed on this score, the plethora of mosquito-borne flaviviruses should settle the matter completely!

Flaviviruses are enveloped, icosahedral, positive-sense RNA viruses. They appear to be related to picornaviruses, but clearly have distinct features, notably an envelope. Because mosquitoes and most other arthropods are sensitive to weather extremes, it is not surprising that arboviral diseases occur throughout the year in the tropics and subtropics, but occur only sporadically, and in the summer, in temperate zones.

Many flaviviruses demonstrate tropism for neural tissue, and flaviviruses are the causative agents of yellow fever, dengue fever, and many types of encephalitis. In the United States, mosquito-borne St. Louis encephalitis virus leads to periodic epidemics in the summer, especially during summers marked by heavy rains and flooding, such as the summer of 1997 in northeastern states. An emerging problem is the establishment of West Nile virus in regions of the eastern United States. In the late summer of 1999, confirmed human cases of encephalitis were documented in New York. Although it is not known how this virus arrived, the strains that are present seem to be related to virus found in the Middle East. During the summer of 2002, the virus spread was documented all the way to California. As of this writing, West Nile virus is now known to be present in virtually all of the contiguous 48 states.

An abbreviated outline of the yellow fever virus replication cycle can be inferred from the genetic and structural map shown in Fig. 15.5. The yellow fever genome is over 10,000 bases long, and unlike poliovirus, it is (i) capped at the 5′ end and (ii) not polyadenylated at the 3′ end. Like poliovirus, the large ORF is translated into a single precursor protein that is cleaved by integral proteases into individual proteins. Some of these cleavage steps are shown in Fig. 15.5. The structural protein precursor includes an integral membrane protein (M) and an envelope glycoprotein. These membrane-associated proteins are translated by membrane-bound polyribosomes, and the process of insertion into the cell's membrane follows the basic outline described for togaviruses later. The M protein contains a "signal" sequence at its N-terminal that facilitates the insertion of the nascent peptide chain into the endoplasmic reticulum. This signal is cleaved from the PreM protein within the lumen of the endoplasmic reticulum—probably by the action of cellular enzymes. The NS (nonstructural) proteins encode the replicase enzymes and do not form part of the virion. Despite this, it is interesting that antibodies directed against the precursor, NS1, protect animals against infection.

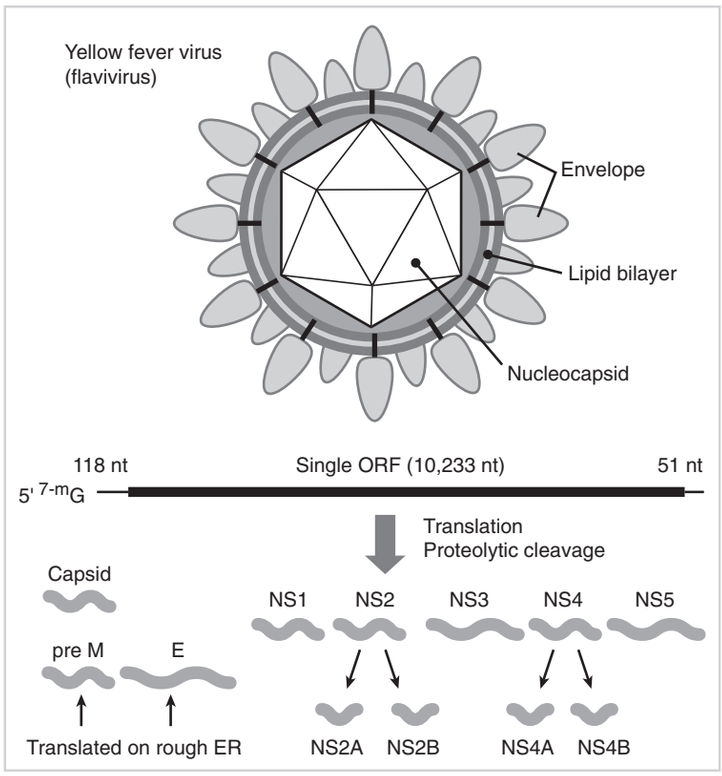


Fig. 15.5 The yellow fever virus (a flavivirus) and its genome. This flavivirus has a replication cycle very similar in broad outline to that detailed for poliovirus. Unlike poliovirus, flaviviruses encode a single envelope glycoprotein, and its approximately 10,000 nucleotide (nt) genome is capped, although not polyadenylated. Also in contrast to poliovirus, the yellow fever virus precursor polyprotein is cleaved into a large number of products as it is being translated, so the very large precursor proteins of poliovirus replication are not seen. The enveloped capsid is larger than that of poliovirus, with a diameter of 40 to 50 nm. (ER, endoplasmic reticulum.)

POSITIVE-SENSE RNA VIRUSES ENCODING MORE THAN ONE TRANSLATIONAL READING FRAME

A positive-sense RNA virus that must regulate gene expression while infecting a eukaryotic host faces a fundamental problem: The eukaryotic ribosome cannot initiate translation of an ORF following translation of one upstream of it. While a positive-sense RNA virus genome could (and some do) contain more than one ORF, these ORFs cannot be independently translated at different rates during infection without some means to overcome this fundamental mechanistic limitation.

One way to overcome the problem is for a virus to encapsidate more than one mRNA (in other words, for the virus to contain a segmented genome). This approach is utilized by a number of positive-sense RNA viruses infecting plants, but has not been described for animal viruses. This finding is somewhat surprising since there are numerous negative-sense RNA viruses with segmented genomes that are successful animal and human pathogens. The list contains influenza viruses, hantaviruses, and arenaviruses.

Despite the disinclination of positive-sense RNA viruses that infect animal cells to encapsidate segmented genomes, another strategy for regulating mRNA expression is utilized successfully. This strategy involves the encoding of a cryptic (hidden) ORF in the genomic RNA, which can be translated from a viral mRNA generated by a transcription step during the replication cycle. With this strategy, viral gene expression from the full-length positive-sense mRNA contained in the virion results in translation of a 5′ ORF, and this protein (an enzyme) is involved in generation of a second, smaller mRNA by transcription.

The second mRNA (which is not found in the virion), in turn, is translated into a distinct viral protein. Such a scheme allows the nonstructural proteins encoded by the virus—the enzymes re-

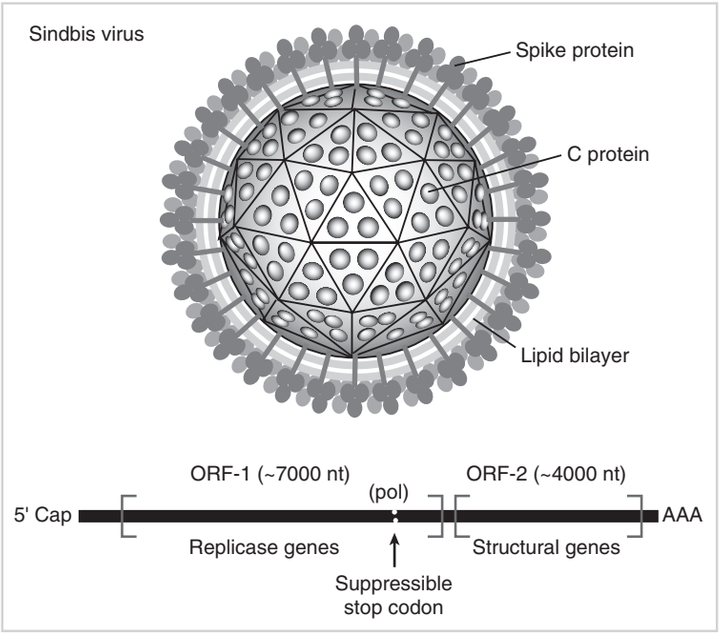


Fig. 15.6 Sindbis virus—a typical togavirus. The virion (60–70 nm in diameter) and genetic map are shown. The Sindbis genome contains two translational reading frames; only the upstream (5′) one can be translated from the approximately 11,000 nucleotide (nt) capped and polyadenylated 49s (positive) virion-associated genomic RNA. This upstream translational frame encodes nonstructural proteins via expression of two precursor proteins. The larger, which contains the polymerase precursor, is translated by suppression of an internal stop codon in the reading frame.

quired for replication—to be expressed in lesser amounts or at different times in the infection cycle than the proteins ending up in the mature virion. Clearly, this approach is effective as witnessed by the number of important pathogens that utilize it.

Two viral mRNAs are produced in different amounts during togavirus infections

Togaviruses are enveloped RNA viruses that display a complex pattern of gene expression during replication. Sindbis virus is a well-studied example. This arthropod-borne virus causes only very mild diseases in (rare) humans, but its size and relative ease of manipulation make it a useful laboratory model for the group as a whole.

Sindbis virus has a capsid structure similar to picornaviruses and flaviviruses, and like flaviviruses, the capsid is enveloped. The viral genome contains two translational ORFs. Initially, only the first frame is translated into viral replication enzymes. These enzymes both replicate the virion RNA *and* generate a second mRNA that encodes viral structural proteins.

The viral genome

Sindbis virus and its 11,700 base genome is shown in Fig. 15.6. The virion genomic RNA (termed 49s RNA for its sedimentation rate in rate zonal centrifugation—see Chapter 11) has a capped 5′ end and a polyadenylated 3′ end. Both capping and polyadenylation appear to be carried out by viral replication enzymes, possibly in a manner somewhat analogous to that seen for the negative-sense vesicular stomatitis virus (VSV), which is discussed in Chapter 16.

The Sindbis virus genome contains two ORFs. The 5′ ORF encodes a replication protein precursor that is processed by proteases to generate four different replicase polypeptides. The 3′ ORF encodes capsid protein and envelope glycoproteins.

The virus replication cycle

Virus entry Viral entry is via receptor-mediated endocytosis as shown in Fig. 15.7a. The entire virion, including envelope, is taken up in the endocytotic vesicle. Acidification of this vesicle leads to modification of the viral membrane glycoprotein. This allows the viral membrane to fuse with the vesicle, and causes the capsid to disrupt so that viral genomic mRNA is released into the cytoplasm.

Early gene expression As shown in Fig. 15.7b, only the 5′ ORF can be translated from intact viral mRNA, because the eukaryotic ribosome falls off the viral mRNA when it encounters the first translation stop signal (either UAA, UAG, or UGA—see Chapter 13). With Sindbis virus, this situation is complicated by the fact that this first ORF in the genomic RNA contains a stop signal about three-fourths of the way downstream of the initiation codon. This termination codon can be recognized to generate a shorter precursor to the nonstructural proteins, but it can also be *suppressed*. (In genetics, the term **suppression** refers to the cell periodically ignoring a translation stop signal either because of an altered tRNA or a ribosomal response to secondary structure in the mRNA encoding it.) With Sindbis virus infection, the suppression is ribosomal, and results in about 25% of the nonstructural precursor protein containing the remaining information shown in ORF-1 in the genetic map. As discussed in Chapter 20, suppression of an internal stop codon also has a role in the generation of retrovirus protein.

In Sindbis virus infection, translation of infectious viral RNA generates replication enzymes that are derived by autoproteolytic cleavage (i.e., self-cleavage) of the replicase precursor protein. This can be considered an "early" phase of gene expression; however, things happen fast in the infected cell and this may only last for a few minutes.

Viral genome replication and generation of 26s mRNA The replication enzymes expressed from genomic 49s positive-sense mRNA associated with genomic RNA to generate 49s negative-sense RNA through RI-1 is shown in Fig. 15.8a. The next step in the process is critical to regulated expression of the two virus-encoded precursor proteins. With Sindbis, the negative-sense RNA complementary to genomic positive-sense RNA is the template for *two* different positive-sense mRNAs. Both are capped and polyadenylated. The first is more 49s positive-sense virion RNA. The second is 26s positive-sense RNA. The shorter 26s mRNA is generated by replicase beginning transcription of negative-sense RNA in the middle and generating a "truncated" or **subgenomic mRNA**. The region on the negative-sense strand where the transcriptase binds is roughly analogous to a promoter, but its sequence does not exhibit the features of promoters found in DNA genomes.

Generation of structural proteins The short 26s mRNA contains only the second ORF contained in the full-length genomic RNA. This ORF was hidden or inaccessible to translation of the full-length virion mRNA. With the 26s mRNA, however, cellular ribosomes can translate the ORF into precursors of capsid and envelope proteins. Expression of structural proteins, thus, requires at least partial genome replication and is generally termed *late* gene expression, although it occurs very soon after infection. Translation of the 5′ region of late 26s mRNA generates capsid protein that is cleaved from the growing peptide chain by proteolytic cleavage. This cleavage generates a new N-terminal region of the peptide. The new N-terminal region of the peptide contains a stretch of aliphatic amino acids, and the hydrophobic nature of this "*signal*" sequence results in the growing peptide chain inserting itself into the endoplasmic reticulum in a manner analogous to synthesis of any cellular membrane protein. This process is shown in Fig. 15.8b.

Following initial insertion of the membrane proteins' precursor, the various mature proteins are formed by cleavage of the growing chain within the lumen of the endoplasmic reticulum. This maturational cleavage is carried out by cellular proteins.

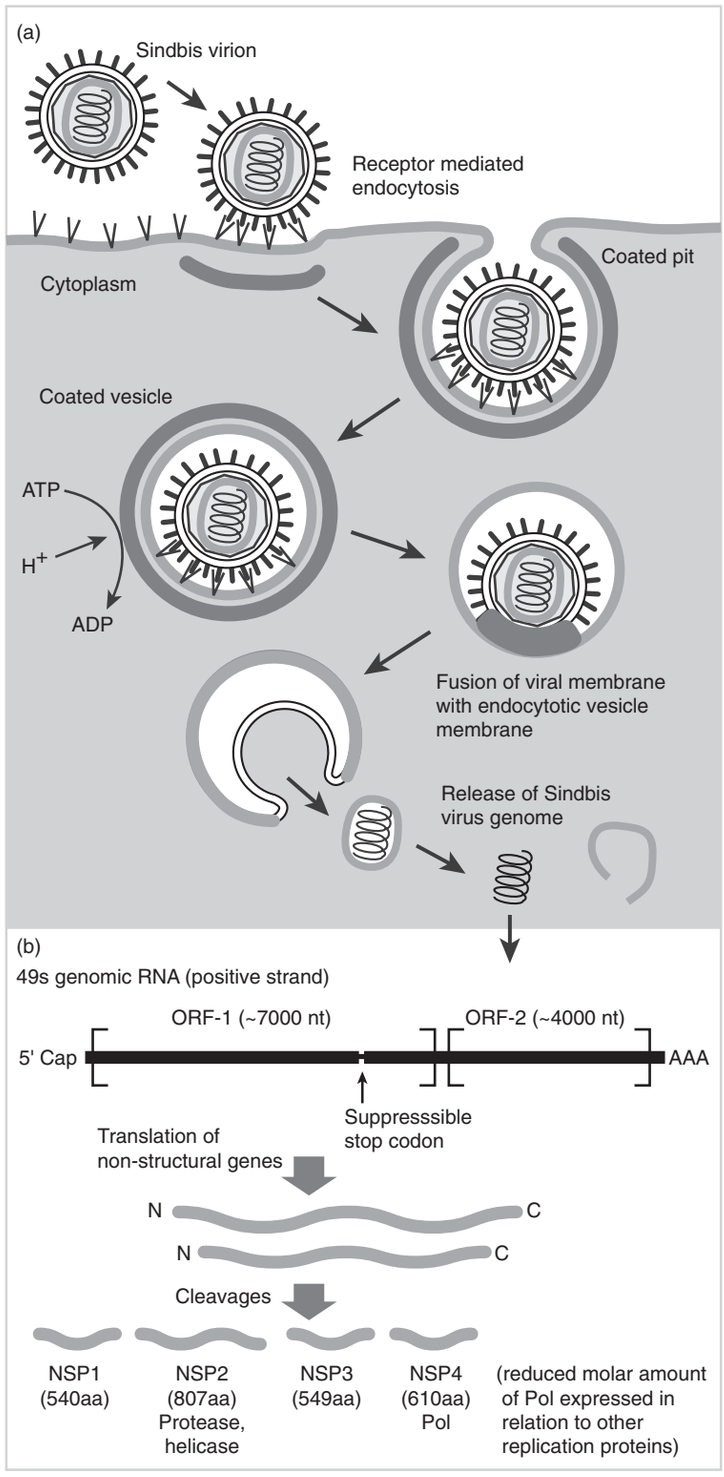


Fig. 15.7 The early stages of Sindbis virus infection. *a*. The first step is receptor-mediated endocytosis, leading to fusion of the viral membrane with that of the endocytotic vesicle, which leads to release of the Sindbis virus genome (mRNA) into the infected cell's cytoplasm. As outlined in Chapter 6, internalization of the enveloped virion within an endocytotic vesicle is followed by acidification and covalent changes in membrane proteins. This results in fusion of the viral membrane with that of the endocytotic vesicle and release of the viral genome. *b*. Translation of the virion RNA results in expression of the precursors to the nonstructural replicase and other viral proteins encoded in the 5′ translational reading frame. These proteins mediate replicase, capping, and protease functions.

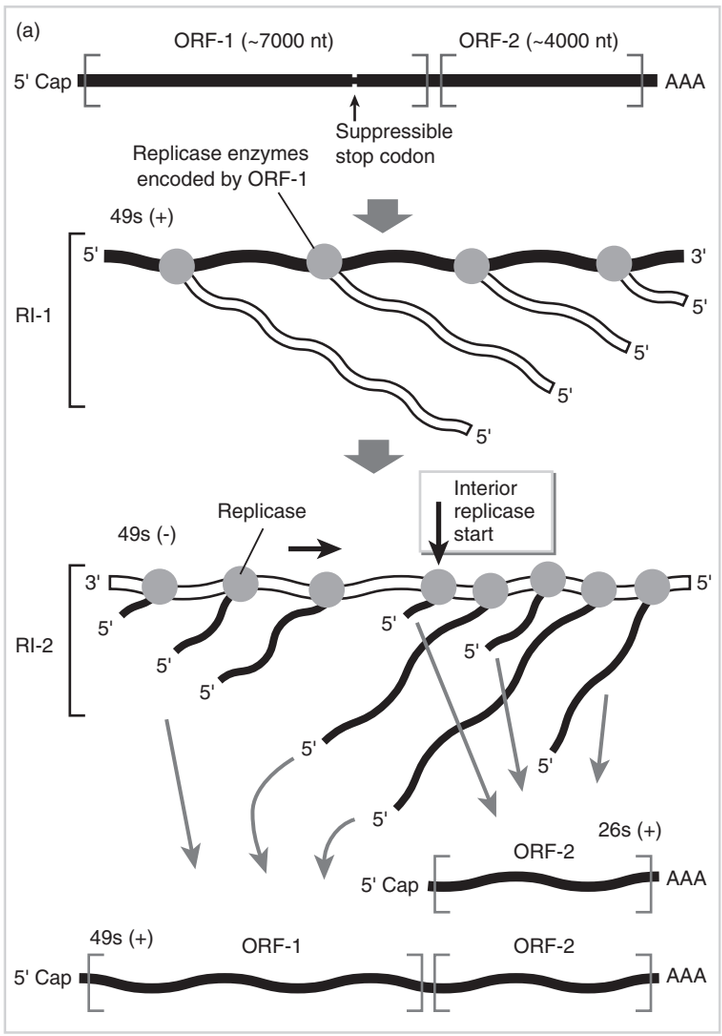


Fig. 15.8 *a.* The replication of Sindbis virus genome, and generation of the subgenomic 26s mRNA. This mRNA is expressed by an internal start site for viral replicase, and is translated into structural proteins since it encodes only the open reading frame (ORF) that was cryptic in the 49s positive-sense virion RNA. *b.* The synthesis of Sindbis virus structural proteins. Structural proteins are translated as a single precursor. When the N-terminal capsid protein is cleaved from the precursor, a signal sequence consisting of a stretch of aliphatic amino acids associates with the endoplasmic reticulum. This association allows the membrane protein portion of the precursor to insert into the lumen of the endoplasmic reticulum. As the protein continues to be inserted into the lumen, it is cleaved into smaller product proteins by cellular enzymes. Cellular enzymes also carry out glycosylation.

Posttranslational processing, such as glycosylation of membrane-associated components of the late structural protein, takes place in the Golgi apparatus, and viral envelope protein migrates to the cell surface. Meanwhile, capsid formation takes place in the cytoplasm, genomes are added, and the virion is formed by budding through the cell surface, as described in Chapter 6.

Togavirus cytopathology and disease

The replication process of togaviruses is a step more complex than that seen with picornaviruses, and the cell needs to maintain its structure to allow continual budding of new virus. Accordingly, there is less profound shutoff of host cell function until a long time after infection.

A major cytopathic change is alteration of the cell surface. This can lead to fusion with neighboring cells so that virus can spread without ever leaving the first infected cells. This alteration to the cell surface also involves antigenic alteration of the cell. Such types of cytopathology are found with many enveloped RNA viruses, whether they are positive or negative sense.

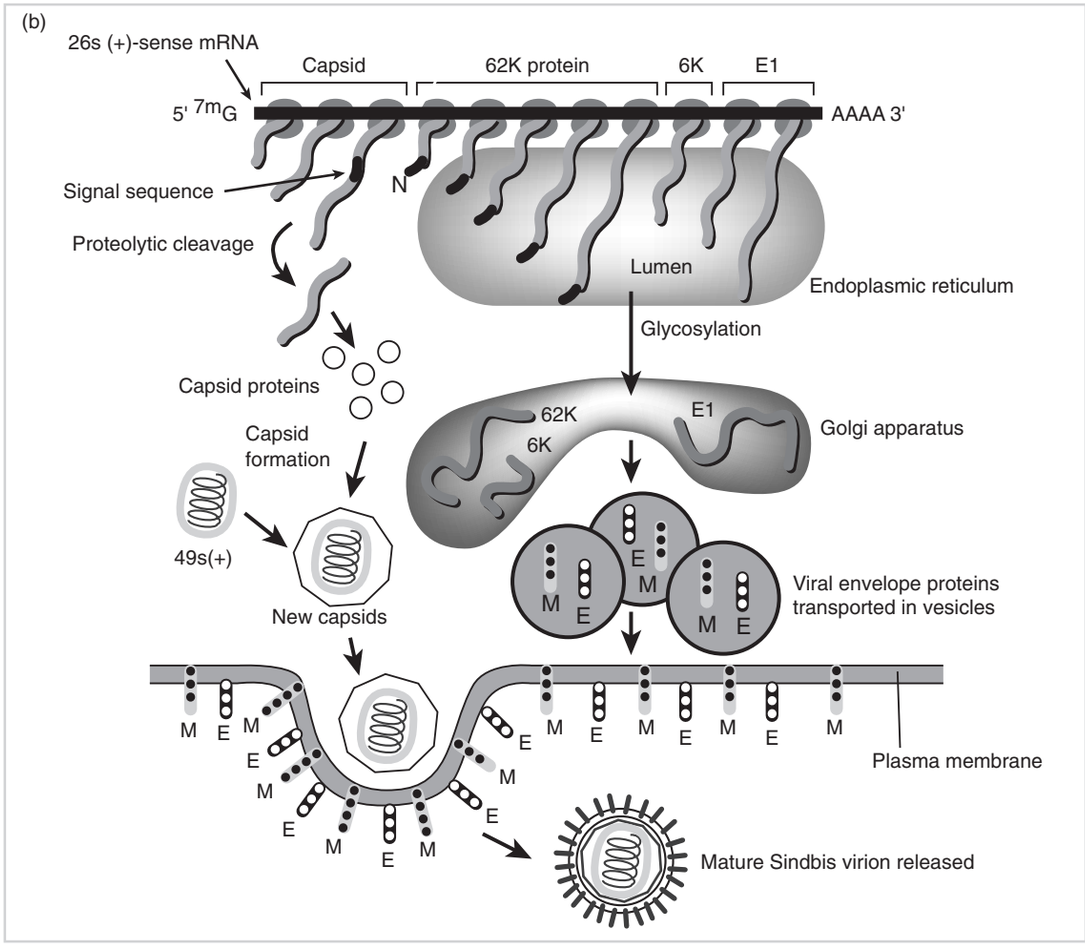


Fig. 15.8 *Continued*

The togaviruses are an extremely successful group of viruses, and like the flaviviruses, many are transmitted by arthropods. As noted in Chapter 5, it is for this reason that these two groups of positive-sense RNA viruses are termed arboviruses (arthropod-borne viruses). While this terminology is convenient for some purposes, it does not recognize significant differences in the replication strategies of these two groups of viruses. Further, numerous other types of viruses are spread by arthropod vectors, and some togaviruses and flaviviruses are *not* transmitted by such vectors. A striking example is rubella (German measles) virus.

Many togaviruses cause sporadic outbreaks of mosquito-borne encephalitis because they have a propensity for replication in cells making up the brain's protective lining. Although such disease can be severe, many forms have a favorable prognosis with proper medical care, as neurons are not the primary targets of infection.

The only known host for rubella virus is humans. The virus causes generally mild and often asymptomatic diseases in children and adults, although a mild rash may be evident. Despite the generally benign course of infection, it is remarkable that rubella is associated with a diverse group of clinical diseases, including rubella arthritis and neurological complications.

Periodic local epidemics are characteristic of rubella virus infections, and although the virus induces an effective immune response, the endemic nature of the virus ensures that once a large-

enough pool of susceptible individuals arises, sporadic regional epidemics occur. The major problem with these periodic occurrences is the very fact that the disease is often so mild as to be asymptomatic in adults of childbearing age. While the symptoms are very mild for adults and children, this is not the case for fetal infections. Infection of the mother in the first trimester of pregnancy often leads to miscarriage, and a fetus who survives is almost inevitably severely developmentally impaired. Infection of the mother later in pregnancy has a more benign outcome.

The tragedy of rubella infections is that although there are effective vaccines, the disease is often so mild that an individual can be infected and can spread the virus without knowing it. For this reason, women of childbearing age who are in contact with young children or other adults at risk of infection should be vaccinated.

A somewhat more complex scenario of multiple translational reading frames and subgenomic mRNA expression: coronavirus replication

Even more complex scenarios exist for expression and regulation of gene function in infections by positive-sense RNA viruses. The replication strategy of the coronaviruses is a good example of such complexity. The structure of coronaviruses is shown in Fig. 15.9, and is unusual for a positive-sense RNA virus.

The nucleocapsid is helical within a roughly spherical membrane envelope, and the envelope glycoproteins project as distinct "spikes" from this envelope. These glycoprotein spikes from the lipid bilayer appear as a distinctive crown-like structure in the electron microscope, hence, the name *corona* (crown)-viruses.

The 30 kb coronavirus genome encodes at least five separate translational reading frames, and is the template for the synthesis of at least six subgenomic mRNAs. Each subgenomic mRNA contains a short, identical leader segment at the 5′ end that is encoded within the 5′ end of the genomic RNA. All subgenomic mRNAs have the same 3′ end, and thus are a nested set of transcripts. Only the 5′ translational reading frame is recognized in each, and the others are cryptic. These features are also shown in Fig. 15.9.

Coronavirus replication

Coronavirus replication involves the generation and translation of genomic and subgenomic viral mRNAs as shown in Fig. 15.10. Virus entry is by receptor-mediated fusion of the virion with the plasma membrane followed by release of genomic RNA. This RNA (one of the largest mRNAs characterized) is translated into a replication protein that, interestingly, is encoded in an ORF encompassing 70% of the virus's coding capacity. The reason why coronavirus replication proteins are encoded by such a large gene is not yet known.

The mature replication proteins derived from the first translation product are used to produce all subsequent mRNA species. There are two competing models that have been presented for coronavirus transcription (Fig. 15.10): leader-primed transcription and discontinuous transcription during negative-strand synthesis.

Leader-primed transcription proposes that the replication proteins first produce a full-length negative strand copy of the genome, using a standard RI-1 structure. From this template is then transcribed multiple copies of the extreme 3′ end, called the leader region. These leader transcripts then function to prime synthesis of subgenomic mRNAs, initiated at homologous regions in between each of the genes (intergenic sequences).

Discontinuous transcription during negative-strand synthesis proposes that the replication proteins transcribe negative-strand copies of the genome, using RI-1 structures. Some of these products are subgenomic. These subgenomic species are produced when the replicase complex in the RI-1 pauses at the intergenic regions and then jumps to the end of the genome, copying the

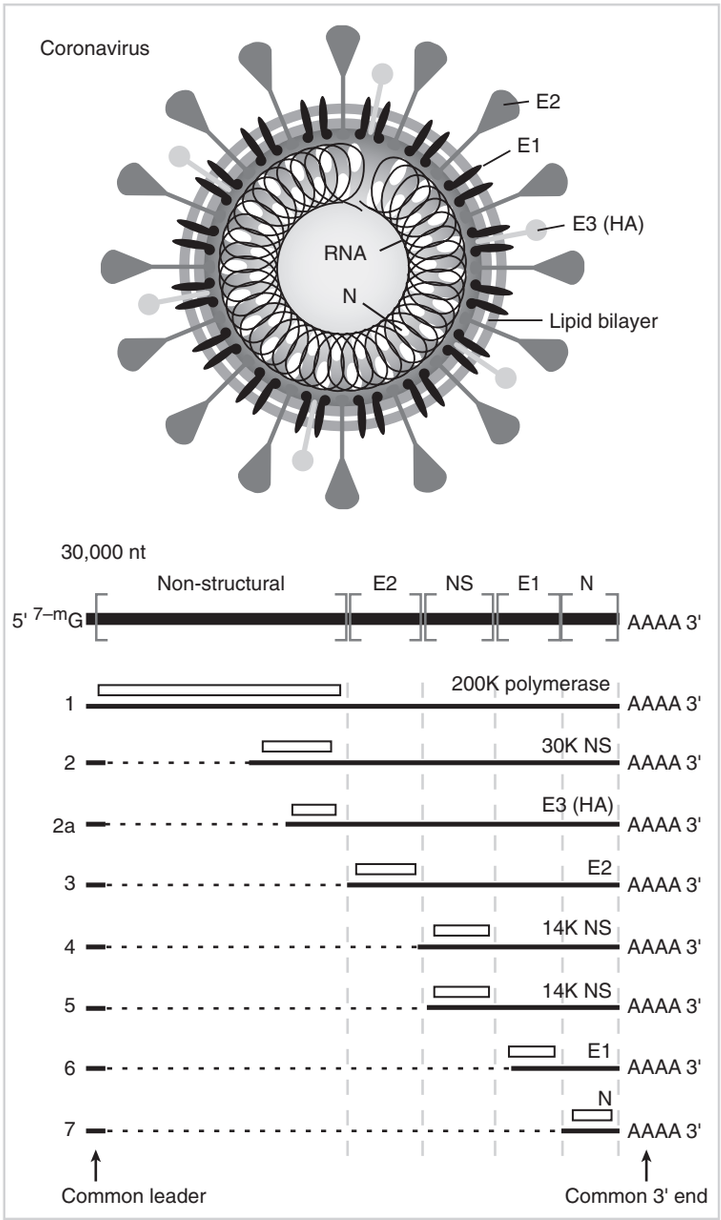


Fig. 15.9 A schematic representation of the coronavirus virion. This is the only known group of positive-sense RNA viruses with a helical nucleocapsid. The name of the virus is derived from appearance of the glycoproteins projecting from the envelope, which gives the virus a crown-like shape. The diameter of the spherical enveloped virion ranges between 80 and 120 nm depending on experimental conditions in visualization. The 30,000 nucleotide (nt) capped and polyadenylated positive-sense genome encodes five translational reading frames that are expressed through translation of the genomic RNA and six subgenomic positive-sense mRNAs. These capped and polyadenylated subgenomic mRNAs each have the same short 5′ leader and share nested 3′ sequences. They are derived by the viral polymerase starting each transcript at the 3′ end of the negative-sense template. The polymerase then, apparently, can translocate or "skip" to various sites on the template where transcription resumes. The result is an mRNA that looks as if it has been spliced, but it all takes place at the level of transcription.

leader sequence. The result of this step is a subgenomic negative-strand RNA that is the complement of the mRNA. Subsequent transcription of this template produces the mRNA itself, using RI-2 structures that are also subgenomic.

Evidence can be obtained in support of both of these models, and no conclusion can currently be drawn as to which of them is in operation. Both models result in mRNAs that have common 5′ sequences (the leader) and common 3′ regions. This nested set of mRNAs is observed during coronavirus infection. Both full-length and subgenomic replicative intermediates can be found in cells at various times after infection. At this writing, it would seem prudent to consider the possibility that transcription may involve features of each of the proposed models.

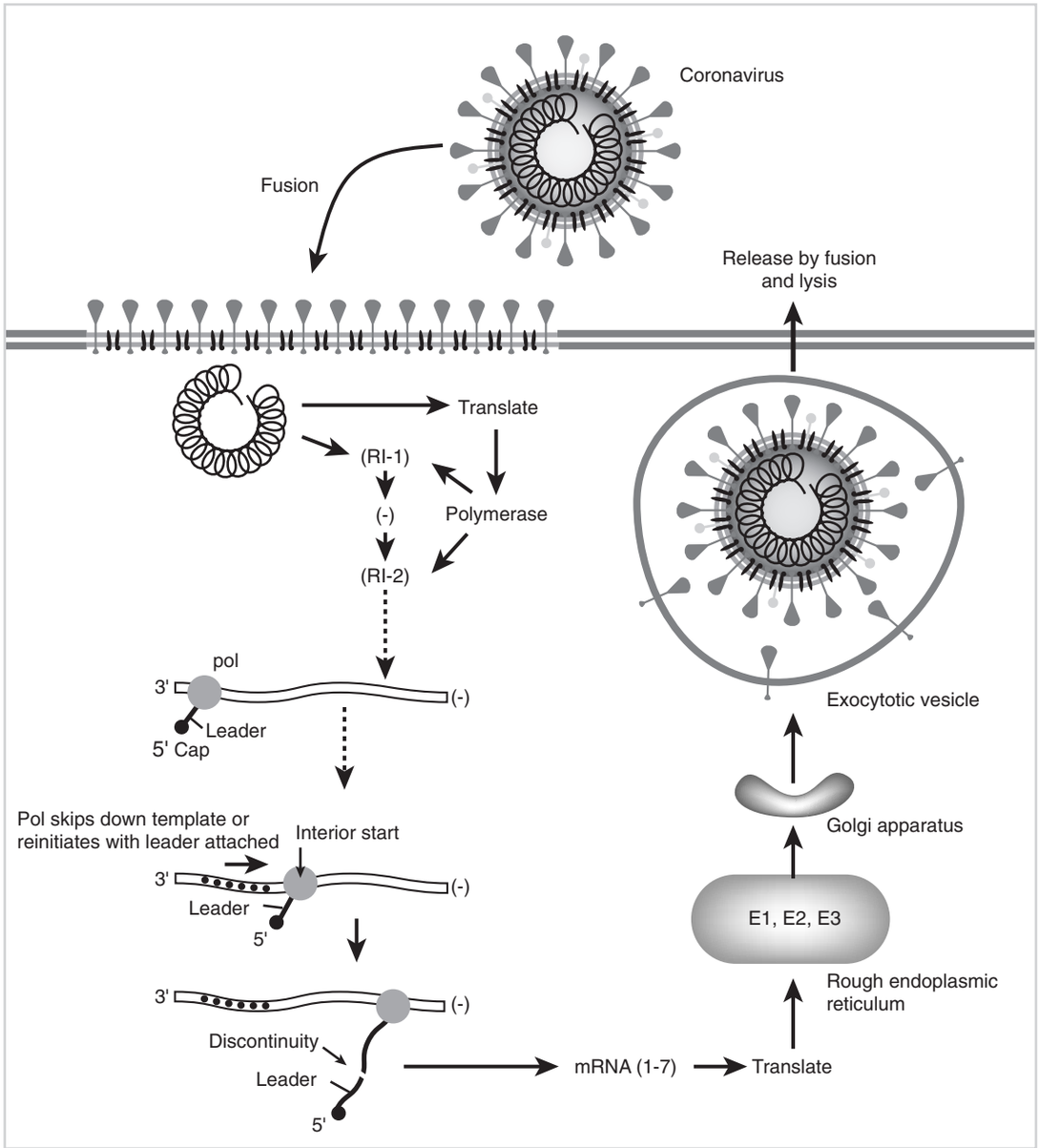


Fig. 15.10 The replication cycle of a coronavirus. Replication is entirely cytoplasmic. Infection is initiated by receptor-mediated membrane fusion to release the genomic mRNA. This RNA is translated into the very large (>200 kd) polymerase/capping enzyme. The interaction between full-length virion positive-sense RNA and replicase generates the templates for the mRNAs. Two models are proposed for the synthesis of subgenomic mRNA: leader-primed synthesis and discontinuous negative-strand synthesis. The two models are shown in the figure and are described in the text. The result of both models is the synthesis of a nested set of mRNAs that contain the same 5′ leader sequence and overlapping 3′ ends. Translation of the various subgenomic mRNAs leads to synthesis of the various structural and nonstructural proteins encoded by interior translational reading frames. The mature virions assemble and become enveloped by budding into intracytoplasmic vesicles; these exocytotic vesicles then migrate to the cell surface where virus is released. At later times, cell lysis occurs.

The specific mechanism of the transcriptase jumping in each model is not known, but it may function to ensure that the virus only has one sequence of RNA needing to be capped (the 5′ leader sequence). The addition of the polyA tracts onto the individual mRNAs also only requires the recognition of one sequence on the positive-sense template by viral replicase since all mRNAs have the same 3′ end.

Cytopathology and disease caused by coronaviruses

The coronaviruses, along with the rhinoviruses, cause mild and localized respiratory tract infections (head colds). The mildness of colds results from a number of both viral and cellular factors.

First, the viruses causing the common cold have a very defined tissue tropism for nasopharynx epithelium. Spread of the virus is limited by ill-defined localized immune factors of the host. The ability of a cold virus infection to remain localized at the site of initial infection is a great advantage to the virus. Local irritation leads to sneezing, coughing, and runny nose, all important for viral spread. Mildness and localization of the infection tend to limit the immune response, which is another distinct advantage. A mild infection results in short-lived immunity, and this, along with the fact that a large number of serotypes exist as a result of the high error frequency of the genome replication process, mean that colds are a common and constant affliction.

In the late winter and spring of 2003 a new illness broke out, focused in China and Singapore. Severe acute respiratory syndrome (SARS) proved to be more than the common cold, having a case fatality rate of 10–20%. As of this writing, the etiologic agent of SARS has been identified as a coronavirus, named SARS-CoV.

REPLICATION OF PLANT VIRUSES WITH RNA GENOMES

A large number of plant viruses contain RNA genomes, and many of the early discoveries in virology were accomplished with plant viruses. The discovery of viruses as specific infectious particles at the end of the nineteenth century focused on work to elucidate the cause of tobacco mosaic disease, culminating in the first description of the tobacco mosaic virus (TMV). This virus took center stage for a number of important early events in biochemical virology, including the first crystallization of a virus particle by W. M. Stanley at University of California, Berkeley; demonstration of the infectious nature of a positive-sense RNA genome by Gierer and Schramm; and in vitro assembly from isolated protein and RNA of an infectious particle by F. Fraenkel-Conrat.

The majority of plant RNA viruses are nonenveloped and have single-stranded genomes. The exceptions are two groups of plant viruses with negative-sense genomes (the plant rhabdoviruses and the *Tospovirus* genus of the bunyavirus family) and one group with dsRNA genomes (e.g., wound tumor virus).

All of the positive-sense plant RNA viruses have genomes that can be translated entirely or in part immediately after infection. Structure of the genome RNA is varied (Table 15.1). The 5′ end may be capped or may have a covalently linked genome protein similar to picornavirus VPg. The 3′

Table 15.1 Genomic structure of some positive-sense RNA viruses infecting eukaryotes.

Virus	No. of genome segments	5′ end	3′ end
Poliovirus	1	VPg	PolyA (genome encoded)
Yellow fever virus	1	Methylated cap	NonpolyA
Sindbis virus	1 (expresses subgenomic mRNA)	Methylated cap	PolyA (A)
Coronavirus	1 (expresses nested subgenomic mRNA)	Common leader with methylated cap	PolyA (A)
Tobacco mosaic virus	1	Methylated cap	tRNAhis
Potato virus Y	1	VPg	PolyA
Tomato bushy stunt virus	1	Methylated cap	NonpolyA
Barley yellow dwarf virus	1	VPg	NonpolyA
Tobacco rattle virus	2	Methylated cap	NonpolyA
Cowpea mosaic virus	2	VPg	PolyA
Brome mosaic virus	3	Methylated cap	tRNAtyr

end may be polyadenylated or not, or may be folded into a tRNA-like structure that can actually be charged with a specific amino acid. There appears to be no role in virus translation for this tRNA, but the fact that the cytoplasm of eukaryotic cells has an enzyme that functions to regenerate the CCA at the 3′ end of tRNA molecules suggests that the tRNA structure may provide the viral genome with a means of avoiding exonucleolytic degradation from the 3′ end.

While expression of the positive-sense RNA genomes of plant viruses follows the same general rules outlined for replication of corresponding animal viruses, there is an added complication. A number of plant virus RNA genomes are segmented. This segmentation means that individual mRNA-sized genomic fragments can be (theoretically, at least) independently replicated and translated. Independent replication and translation allow the virus to maintain a replication cycle in which individual viral genes can be expressed at significantly different levels.

Use of this strategy in virus replication adds the complication that the packaging process is potentially very inefficient. This is certainly true for the packaging of influenza virus described in the next chapter. Alternatively, the packaging process might be controlled in some way to ensure that each viral particle gets its requisite number of genomic fragments. Despite this complication, segmented genomes are a viable strategy for RNA virus replication, and it is not clear why it is not used in the replication of any known positive-sense animal viruses.

With viruses of vascular plants, the limitations in the size of objects that can pass through the cell wall led to another adaptation. The plant viruses with segmented positive-sense RNA genomes package each segment *separately*. Although this separate packaging means that each cell must be infected with multiple virions, plant viruses seem to thrive using this approach, probably for the following reason: Plant viruses are often transmitted mechanically and then spread from cell to cell via the plant's circulation without involvement of a specific immune defense; therefore, high concentrations of virus at the surface of the cell can be maintained.

Viruses with one genome segment

TMV has a helical capsid that encloses a single RNA genome segment of 6.4 kb. Primary translation of the genome produces the replicase complex consisting of the 126 kd and 183 kd replication proteins. Two subgenomic mRNAs are transcribed from negative-sense RNA generated from RI-1. The translation of these two species yields the 17.5 kd coat protein and a 30 kd protein involved in movement of the virus within the infected plant.

Tomato bushy stunt virus has a single RNA genome of 4.8 kb packaged into an icosahedral capsid. Translation of the capped genome results in production of the 125 kd viral replicase. Two subgenomic mRNAs are transcribed from the full-length negative-sense strand generated from RI-1. Translation of these two species leads to synthesis of the 41 kd coat protein and two other proteins thought to be required for cell-to-cell movement of the virus.

Viruses with two genome segments

The genome of cowpea mosaic virus consists of two separate strands of RNA packaged into *separate* icosahedral particles. Since both strands are required for infection, a cell must be infected together by each of the two particles. The larger of the two RNAs (5.9 kb) is translated into a polyprotein that is cleaved into a 24 kd protease, the 4 kd VPg, a 110 kd replicase, and a 32 kd processing protein. The smaller (3.5 kb) RNA encodes a polyprotein that is cleaved into the 42 kd and 24 kd coat proteins and a set of proteins required for cell-to-cell movement of the virus.

Viruses with three genome segments

Brome grass mosaic virus has three separate RNA genome strands (3.2 kb, 2.8 kb, and 2.1 kb) con-

tained in *three separate* icosahedral particles. Again, since all three genome segments are required for infection, cells must receive one of each of the particles. Each of the capped genome segments is translated into a protein. These products include the 94 kd viral replicase, a 109 kd capping enzyme, and a 32 kd cell-to-cell movement protein. In addition, one of the RNAs is transcribed into a subgenomic mRNA that encodes the 20 kd viral coat protein.

REPLICATION OF BACTERIOPHAGES WITH RNA GENOMES

The great majority of well-characterized RNA bacteriophages have linear, single-stranded, positive-sense genomes enclosed within small, icosahedral capsids. These phages (grouped together as the **Leviviridae**) include the male bacteria-specific phage Qβ, MS2, and R17, which attach to the bacteria's F pili.

In broad outline, the replication process of these RNA-containing bacteriophages follows that described for eukaryotic viruses. Infection begins with a translation step, and replication of the viral genome occurs through production of the RI-1 and RI-2 intermediates described in the preceding section.

Regulated translation of bacteriophage mRNA

There is a major difference in the way protein synthesis occurs on bacterial ribosomes as compared to eukaryotic ribosomes, and this leads to a significant difference in the way expression of viral-encoded protein is controlled. As discussed in Chapter 13, bacterial ribosomes can initiate translation at start sites in the interior of bacterial mRNA. This means that a bacterial mRNA molecule with several ORFs can be translated independently into one or all of the proteins. In an RNA bacteriophage infection, protein synthesis programmed by the incoming genome is characterized by synthesis of viral RNA replicase only. Later in infection, after genome replication begins, transition to synthesis of capsid and other proteins begins.

This temporal regulation is governed by the secondary structure of the genome, and initiation of protein synthesis encoded by interior ORFs by ribosomal mechanisms. This can be seen in the phage Qβ, which is diagrammed in Fig. 15.11. This virus encodes three distinct translational reading frames encoding genes for the A (maturational) protein, the coat protein, and replicase. The

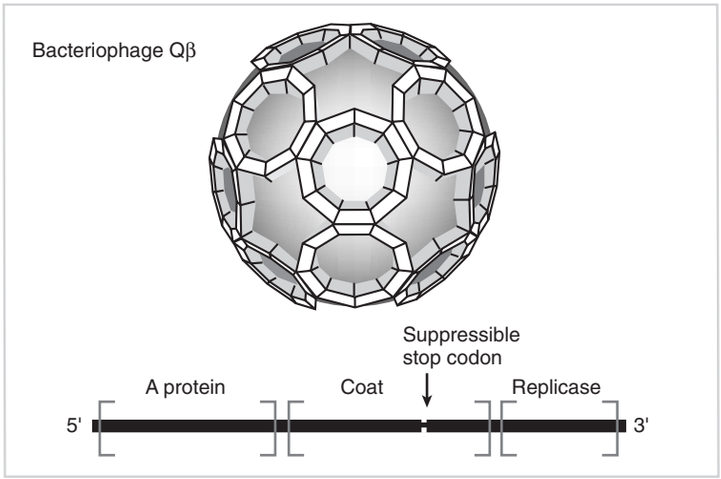


Fig. 15.11 The approximately 25 nm diameter icosahedral capsid of positive-sense RNA bacteriophage Qβ. The positive-sense RNA genome contains three separate open reading frames (ORFs). These ORFs can be independently translated from the full-length virion RNA because unlike the situation in eukaryotic viruses, bacterial ribosomes can initiate translation at interior start signals provided that the ribosome can interact with them. With this bacteriophage, ribosome attachment and translation require active transcription to allow the nascent positive-sense RNA to be unfolded so that the translation start is accessible.

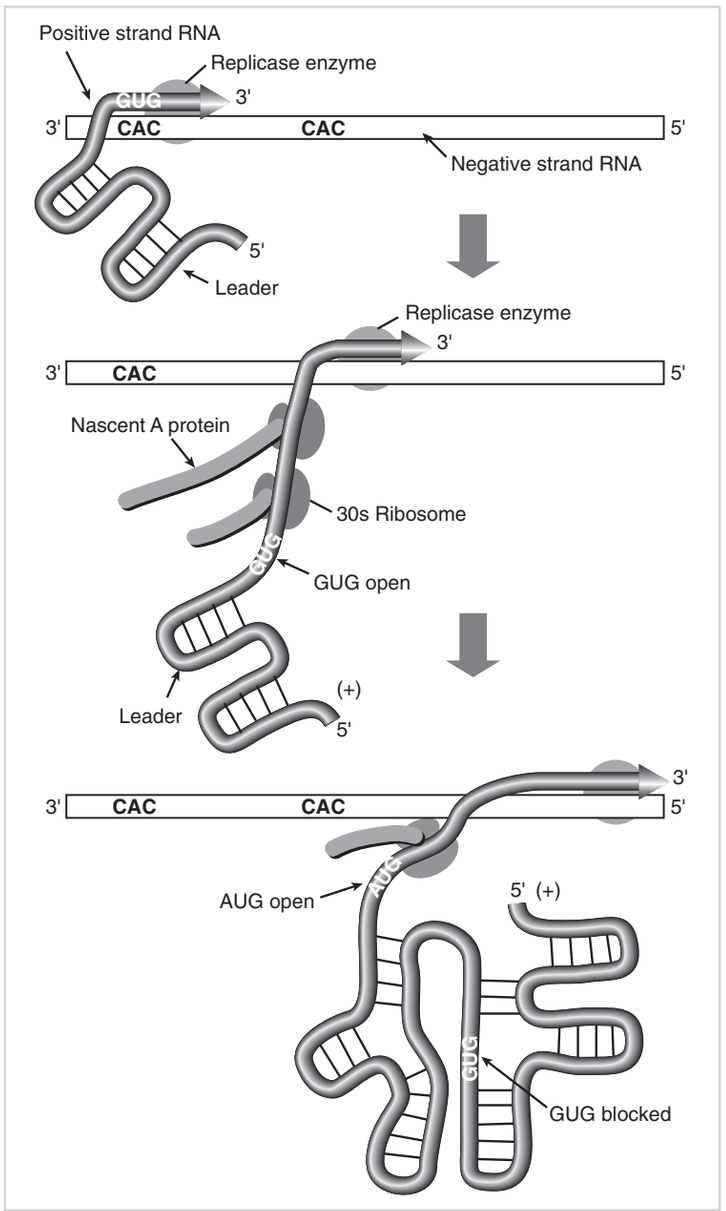


Fig. 15.12 Coupled transcription-translation of bacteriophage Qβ RNA results in opening the blocked translational start site for the A (maturational) and coat proteins. As the replicase enzyme passes the region containing the translation start site on the negative-sense template (which is a GUG for the A protein) the nascent positive-sense mRNA can interact with a ribosome before it has a chance to fold into a structure in which this initiator codon is sterically blocked. Multiple ribosome entry results in translation of a large number of copies of the maturational and coat proteins being synthesized. High levels of coat protein specifically inhibit translation of replicase from full-length genomic RNA so that replicase is only synthesized at early times in the replication cycle. For this reason, it is often termed an "early" protein or gene product.

coat protein translational reading frame has a translation terminator that is misread (suppressed) as a tryptophan residue about 1% of the time, and when this happens, a larger capsid protein with additional amino acids is generated. Suppression of the termination is absolutely required for phage replication.

A portion of the replication cycle of Qβ is shown in Fig. 15.12. Ribosomes can associate with the genomic RNA, but this positive-sense genome is folded in such a way that the only start codon available for interaction with a ribosome is the one that begins translation of phage RNA replicase. All other start codons are involved in base-pairing interactions as a part of the secondary structure. For this reason, replicase is the only phage protein expressed at the start of infection.

Synthesis of new positive-sense genomes takes place through formation of RI-1 and RI-2. As

new positive-sense genomic RNA disassociates from the negative-sense template near the replicase, secondary structure has not yet formed. This results in the start codon for the A and coat proteins being available to begin translation. The A protein uses a GUG instead of an AUG initiation codon. Similarly, newly replicated positive-sense strands immediately interact with ribosomes to yield the capsid proteins necessary for the formation of new virus particles.

This simple mechanism ensures that the earliest protein expressed will be replicase. Further, since a relatively large amount of RI-2 will need to be present, synthesis of A and capsid proteins will only occur when there are a large number of genomes waiting to be encapsidated. Multiple entry of ribosomes onto the nascent viral mRNA ensures that a large amount of structural protein will be available when necessary.

Finally, the phage controls the amount of replicase synthesized in infection so that progeny positive-sense strand does not end up recycling too long. Such control is accomplished by the capsid protein actually inhibiting synthesis of replicase from mature positive-sense RNA. Therefore, after about 20 minutes, increasing levels of capsid proteins shut off replicase synthesis.

QUESTIONS FOR CHAPTER 15

1 What are the steps in the attachment and entry of poliovirus in a susceptible host cell?

2 The Picornaviridae (e.g., poliovirus) have, as their genome, one molecule of single-stranded RNA. This genomic RNA functions in the cell as a monocistronic mRNA. However, picornavirus-infected cells contain 10 or more viral proteins.

a What mechanism have these viruses evolved such that this monocistronic mRNA produces this large number of translation products?

b The poliovirus mRNA does not have a 5′ methylated cap that is present on host cell mRNA. How do host cell ribosomes begin translation of this message?

3 Foot-and-mouth disease virus (FMDV) is a member of the family Picornaviridae. Based on your knowledge of the properties of members of this family, complete the following table with respect to FMDV and each of the characteristics listed. State whether the characteristic is present or absent.

Characteristic	Present or absent for FMDV
5′ methylated cap	
Subgenomic RNAs	
3′ polyadenylation	
Single-stranded, positive-sense genome	
Expression of genome as a polyprotein	

4 The poliovirus genome is a single-stranded RNA of about 7500 nucleotides, with a covalently linked terminal protein, VPg, at the 5′ end and a polyA sequence at the 3′ end. The polyA tail is not added after replication but is derived from the template during replication. VPg is important for replication of this viral RNA, along with poliovirus polymerase and certain host enzymes.

There are two models for the action of VPg:

Model 1. VPg may act as a primer for RNA synthesis, being used as VPg-p$U_{OH.}$

Continued

Model 2. VPg may act as an endonuclease, attaching itself to the 5′ end of a new RNA chain. In this model, RNA synthesis is primed after addition of U residues to the 3′ A at the end of the genome by a host enzyme, followed by a loop-back and self-priming mechanism.

Given these two models, imagine that you have an in vitro system to test the properties of poliovirus genome replication. Your system contains viral genomic RNA as a template and all of the necessary proteins, except as indicated below.

a Assume that model 1 is true. What would you expect to see as the product of the reaction if VPg was left out of the mixture?

b Assume that model 2 is true. What would you expect to see as a product of the reaction if endonucleolytic activity of VPg was inhibited?

5 Draw the structure of the poliovirus RI-1 and RI-2. What are the similarities and differences of these two structures?

6 Which of the following statements is (are) true in regards to the poliovirus genome?

a It lacks posttranscriptional addition of repeating adenines.

b It is approximately 1400 bases long.

c It contains a VPg protein that is cleaved prior to packaging.

d It has a single precursor protein that is cleaved by cellular cytoplasmic nucleases.

7 How are the structural proteins of Sindbis virus generated during the infectious cycle?