18 Pavletich, N. P. and Pabo, C. O. (1991) *Science* 252, 809–817
19 Luisi, B. F., Xu, W. C., Otwinowski, Z., Freedman, L. P., Yamamoto, K. R. and Sigler, P. B. (1991) *Nature* 352, 497–505
20 Schwabe, J. W. R., Neuhaus, D. and Rhodes, D. (1990) *Nature* 348, 458–461
21 Landschulz, W. H., Johnson, P. F. and McKnight, S. L. (1988) *Science* 240, 1759–1764
22 O'Shea, E. K., Klemm, J. D., Kim, P. S. and Alber, T. (1991) *Science* 254, 539–544
23 Gilmore, T. D. (1990) *Cell* 62, 841–843

24 Lenardo, M. J. and Baltimore, D. (1989) *Cell* 58, 227–229
25 Murre, C., McCaw, P. S. and Baltimore, D. (1989) *Cell* 56, 777–783
26 Dynlacht, B. D., Hoey, T. and Tjian, R. (1991) *Cell* 66, 1–20
27 Inostroza, J. A., Mermelstein, F. H., Ha, I., Lane, W. S. and Reinberg, D. (1992) *Cell* 70, 477–489
28 Fjermedal, G. (1984) *Magic Bullets*, Macmillan
29 Klevit, R. E., Herriott, J. R. and Horvath, S. J. (1990) *Proteins* 7, 215–226
30 Oas, T. G., McIntosh, L. P., O'Shea, E., Dalquist, F. W. and Kim, P. S. (1990) *Biochemistry* 29, 2891–2994

reviews

Alphaviruses – vectors for the expression of heterologous genes

Sondra Schlesinger

DNA viruses and retroviruses are well established as vectors for the expression of heterologous genes, but there is increasing interest in the possibilities of using RNA viruses, which do not replicate through a DNA intermediate, for this purpose. This article summarizes some of the general properties of RNA viruses and concentrates on one class of RNA viruses – the alphaviruses – and their potential as expression vectors.

Virology is celebrating its centennial in 1992. The existence of a filtrable infectious agent was first recognized 100 years ago, when studies in Russia and The Netherlands on the infectious agent that caused the mosaic disease of tobacco plants led to the concept of infectious agents that were smaller than bacteria[1]. At the beginning of the 20th century, the first human virus was discovered – yellow fever virus – and for much of the intervening time the study of viruses has been involved primarily in the identification and characterization of viral agents that cause disease. However, viruses are now beginning to be associated not only with the cause of disease, but also with its cure. Viruses are being used as vectors for the introduction of heterologous genes into cells, which means that they can be used for the delivery of therapeutic agents into cells, as well as for gene therapy.

RNA viruses as gene expression vectors

RNA viruses infect a wide range of organisms, both prokaryotic and eukaryotic. An important feature of these viruses is that many of them produce high

levels of specific viral proteins, which makes them good candidates as vectors for heterologous gene expression[2,3].

To use the RNA viral genome as a vector necessitates first converting the RNA genome into a genomic complementary DNA (cDNA), which is then placed downstream from a promoter for a bacterial or bacteriophage DNA-dependent RNA polymerase. Inserts of heterologous DNA that encode the proteins of interest may then be introduced into the cDNA downstream from a strong viral promoter by standard restriction-endonuclease digestion and ligation techniques. The whole cDNA is then transcribed *in vitro*, using the RNA polymerase promoter, into the genomic viral RNA, which, in many cases, can be directly transfected into cultured cells, or into plants or animals. Many RNA viruses possess relatively small genomes; cloning of the recombinant cDNA can therefore be accomplished in only a few steps. Many of the RNA viruses that infect higher eukaryotic organisms replicate exclusively in the cytoplasm and hence expression of their genes will be independent of host nuclear functions and will not involve any splicing.

Not all RNA viruses, however, have the same mode of replication and this affects the strategy that must

S. Schlesinger is at the Department of Molecular Microbiology, Washington University School of Medicine, St Louis, MO 63110-1093, USA.

be used in engineering the viral genome to obtain expression of the heterologous genes. A key feature which divides RNA viruses into two categories is the 'polarity' of the RNA genome[4].

● *Positive polarity*

RNA genomes with positive polarity are those in which the genomic RNA functions as a mRNA; the naked RNA genome added to cells is infectious (i.e. it is able to initiate the steps leading to the production of new virus progeny).

● *Negative polarity*

RNA genomes in which the sequence is complementary to the mRNA (i.e. negative polarity) must be transcribed into mRNA before translation to yield protein can occur. Negative-strand RNA viruses contain the enzymes required for transcription of their RNA within the virus particle. Although this makes the engineering of such genomes more complex, the feasibility of using these genomes to express heterologous genes has been established[5,6].

Two other characteristics which vary among the different RNA-virus families and are relevant to their potential use as expression vectors are: (1) whether the genome exists as a single molecule of RNA or as several molecules of different structure (i.e. a 'segmented' genome); and (2) how the genome is translated.

Many RNA viruses have segmented genomes: examples include both the negative-strand RNA myxoviruses, most notably influenza virus; and the positive-strand RNA plant viruses such as the bromoviruses and cucumoviruses. The genome of viruses such as togaviruses, picornaviruses and flaviviruses is a single strand of RNA of positive polarity. The latter two virus families have RNA genomes with a single open reading frame (ORF), from which the individual viral proteins are generated by co- and post-translational proteolytic cleavages. Insertions into these genomes must be in frame, and must permit the correct proteolytic processing. Efforts to insert small heterologous epitopes into the picornavirus genome have met with success[7].

Togaviruses, coronaviruses and the plant viruses, bromoviruses and cucumoviruses, are examples of virus families in which translation of the nonstructural proteins involved in replication, and the translation of the structural proteins essential for packaging of the genome can be independent. During the course of replication, the infected cell produces both genomic- and subgenomic-sized mRNAs. The genomic RNA is translated to produce some of the viral-specific proteins, but the other viral proteins are translated from subgenomic RNAs.

## The replication strategy of alphaviruses

Alphaviruses are the major genus of the Togavirus family[8]. They include Sindbis virus, Semliki Forest virus and the human pathogens eastern and western equine encephalitis viruses.

Sindbis virus and Semliki Forest virus are best known as valuable models for molecular and cell biology, and it is these two viruses that are presently being developed as vectors for the expression of heterologous genes. The alphavirus genome is a single-strand RNA of positive polarity of ~1.2 x 10[4] nucleotides, that is capped at the 5' terminus and polyadenylated at the 3' terminus[8,9]. The genomic RNA is referred to as 49S RNA (on the basis of its sedimentation coefficient), and is encapsidated in an icosahedral protein shell which is composed of a single-protein subunit. This shell is surrounded by a lipid bilayer that contains two transmembranal glycoproteins, organized as trimers of heterodimeric subunits. As many of the steps in virus replication and maturation utilize the host cellular machinery, the virus proteins provide markers for analysing the synthesis and modification of membrane glycoproteins, and the localization of proteins to the plasma membrane.

The 5' portion (approximately two-thirds) of the alphavirus genome contains the genetic information encoding the nonstructural viral proteins that are required for transcription and replication of the viral RNA[8,9]. The 3' portion of the gene (approximately one-third) contains the genes encoding the viral structural proteins – the capsid protein, a hydrophobic 6kDa protein and the two viral envelope glycoproteins.

The replication cycle of the alphaviruses is shown schematically in Fig. 1. Following attachment of the virus to the cell surface, entry is effected by receptor-mediated endocytosis. Fusion of the virus membrane with the endosomal membrane releases the viral nucleocapsid into the cytoplasm, where translation of the genomic RNA occurs. Only the nonstructural viral proteins are translated from the genomic RNA. There is a single ORF and the resulting polyprotein is co- and post-translationally cleaved to form four polypeptides. These function together in a replication complex which is required for the synthesis of the complementary, negative RNA strand. The newly synthesized negative RNA strand serves as a template for the synthesis of more virion RNA, and also for the transcription of a subgenomic (26S) RNA which may then be translated to yield the structural proteins.

The nucleotides spanning the junction between the genes for the nonstructural protein and those for the structural proteins on the negative RNA strand serve as the promoter for transcription of the subgenomic RNA[10,11]. The minimal sequence required to initiate transcription (19 nucleotides upstream and five nucleotides downstream from the start of the subgenomic RNA) is located within a region that is highly conserved among alphaviruses[9–11]. If regions flanking the minimal promoter are also present, the level of transcription of the subgenomic RNA is increased. The viral structural proteins translated from 26S RNA are also synthesized as a polyprotein, with the N-terminal capsid protein functioning as an autoprotease to cleave itself from the nascent polypeptide. The capsid protein can then interact with the genomic
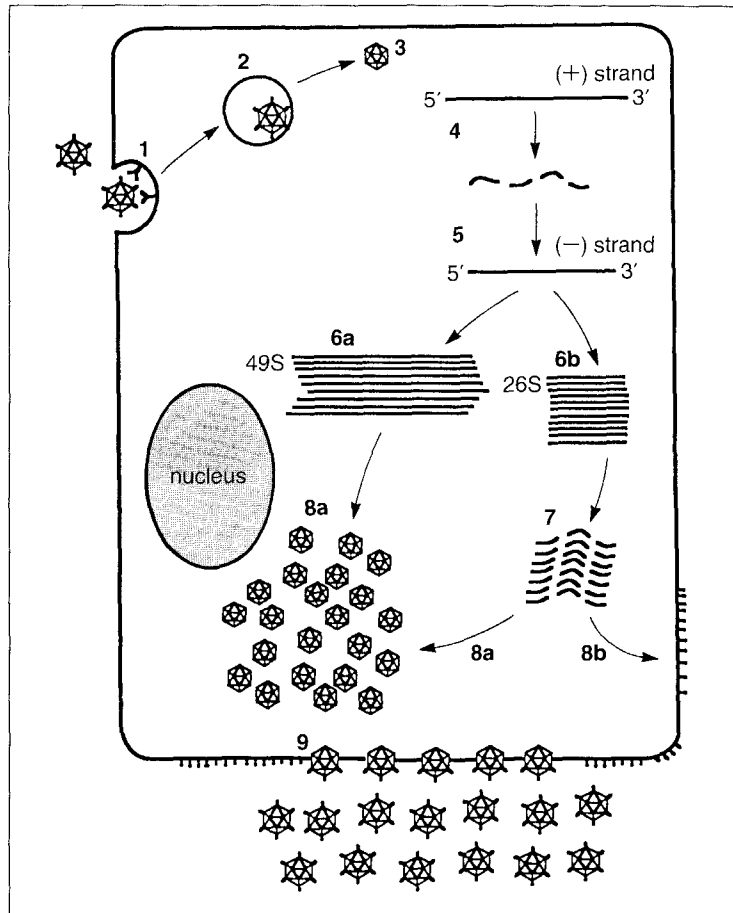
Figure 1

Schematic representation of the life cycle of an alphavirus from the first steps of attachment to the host cell and endocytosis to the final steps of assembly and budding through the plasma membrane (see text for details). (**1**) Receptor-mediated endocytosis. (**2**) Low-pH dependent membrane fusion. (**3**) Virus uncoating and release of nucleocapsid into cytoplasm. (**4**) Translation and processing of the genomic RNA to give a single polyprotein that is cleaved to give four polypeptides, required for the transcription and replication of the viral RNAs. (**5**) Synthesis of negative RNA strand. (**6a**) Synthesis of genomic 49S RNA. (**6b**) Transcription of subgenomic 26S RNA. (**7**) Translation of 26S RNA to yield the virus structural proteins. (**8a**) Nucleocapsid formation. (**8b**) Transport to plasma membrane and processing of viral glycoproteins. (**9**) Virus budding through the plasma membrane. In vertebrate cells, this process takes 7–12 hours *In vitro*. Adapted from an illustration by Chang Hahn.

RNA to form the nucleocapsid. There is a specific recognition signal located near the 5' terminus of the RNA that serves as the nucleation site for encapsidation[12] and thus ensures that the genomic, but not the subgenomic RNA, is packaged. The viral glycoproteins, synthesized on membrane–bound ribosomes, are transported through the host-cell Golgi network to the plasma membrane of the infected cell. Finally, interaction of the nucleocapsid with the cytoplasmic tail of a viral glycoprotein leads to budding of the virus particle and release of infectious virus from the host-cell membrane.

Alphavirus vectors

The basic strategy for using alphaviruses as vectors for the expression of heterologous genes has been to construct cDNAs of the alphavirus genome, in which the heterologous gene is placed downstream from the promoter used to transcribe a subgenomic RNA[13]

(Fig. 2a). The cDNA is transcribed *in vitro* using the SP6 DNA-dependent RNA polymerase and the RNA transcript is transfected into cells either by lipofection[14] or electroporation[15] – methods that enhance the uptake of nucleic acids into cells. The genomic RNA contains the genes for the nonstructural proteins and the *cis*-acting sequences important in the recognition of the RNA for replication and encapsidation[8,9]. The subgenomic RNA synthesized in the transfected cells is translated into the heterologous protein. The first alphavirus vector designed using this strategy replaced the viral structural-protein genes with the gene for the bacterial enzyme chloramphenicol acetyltransferase (CAT)[13]. About 10^8 CAT polypeptides were produced per transfected cell, corresponding to 3% of the total cell protein. More recently, a similar vector was designed using the cDNA of Semliki Forest virus[15].

Infection by a virus is a much more efficient means for introducing its genome into the cell than transfection, and several different strategies have been designed to permit packaging of the RNA vector. Hahn *et al.*[16] engineered the Sindbis virus cDNA so that it contained two promoters for subgenomic RNA synthesis; one upstream of the Sindbis virus genes, the other upstream of the heterologous gene (Fig. 2b). Cells transfected with the recombinant, *in vitro* transcribed RNA produced virus with yields ranging from 10^8 to 10^9 plaque forming units (PFU) per ml. The vector studied in the greatest detail contained the heterologous gene inserted downstream from the structural genes. Infection of cells with the double subgenomic RNA Sindbis virus vector resulted in the production of $>10^6$ CAT molecules seven hours after transfection.

The size of the heterologous insert can affect the stability of the recombinant RNA. Large inserts (i.e. >3 kb) are much less stable than smaller ones (i.e. <2 kb). The level of expression of the heterologous gene is higher when it is located downstream from the structural genes, although the recombinant RNA appears to be more stable when placed upstream of the structural gene.

The alphavirus vector containing a single subgenomic RNA promoter can be packaged by complementation with a defective helper-virus RNA (Fig. 3). Two types of packaging helper RNAs have been used:

● One, first described by Geigenmüller-Gnirke *et al.*[17], contains the structural genes downstream from the promoter for the subgenomic RNA and also contains the *cis*-acting sequences required for replication and encapsidation (Fig. 3). Most of the released particles contain both the helper and vector RNAs in a single particle; such particles are able to form plaques and appear to be stable over several passages (i.e. repeated subculture). This type of packaged vector should prove to be of value in animal experiments in which both production of the heterologous gene and spread of the virus is desired.

● The second type of defective helper-virus RNA lacks the region required for encapsidation of the RNA. In studies of the packaging of the Semliki Forest virus vector, Liljeström and Garoff[15] found that co-transfection of the latter type of helper RNA with vector RNA led to efficient packaging of the vector RNA, but the helper RNA was not packaged. Electroporation of both helper and vector RNAs produced yields of up to 10^9 particles from 5×10^6 cells (as observed by immunofluorescence analysis).

Packaging of the Sindbis virus vector using Sindbis defective helper RNAs has proved to be more complex (P. Bredenbeck, C. Rice, I. Frolov and S. Schlesinger, unpublished). Deletion of the packaging signal leads to a considerable reduction in the amount of helper RNA in the progeny; however, its presence can still be detected by analysis of the RNA synthesized in vector-infected cells, and by a low level of PFU in the media, presumably due to co-packaging of the RNAs. This may be due to the ability of the encapsidation process to include other RNAs in the capsid in conjunction with the viral RNA encoding the specific packaging signal.

A second potential problem is recombination between the packaging helper virus RNA and vector RNAs. The two Sindbis RNAs can undergo recombination to produce a single molecule of RNA containing the genes that encode both the nonstructural and structural proteins[18]. Although the growth and replication of such recombinants are suppressed under conditions in which most of the cells have been transfected with both helper and vector RNAs, they can be detected after passaging of the infectious particles.

A completely different approach to using alphaviruses as vectors has been to insert heterologous sequences into the viral structural proteins in such a way that the virus is still viable, but the heterologous epitope is expressed and can function as an immunogen[19]. This concept was originally tested with picornaviruses, where the site of insertion was chosen on the basis of the crystallographic structure of the virus[7]. However, the three-dimensional structures of the Sindbis virus glycoproteins are not known, and therefore it has not been possible to predict the regions of the glycoprotein which form surface loops and into which insertion of additional sequence would not adversely affect glycoprotein folding. Initial studies, using an epitope from Rift Valley fever virus, involved making random insertions into the structural genes, followed by selection of viable viruses containing the insert. The chimeric Sindbis virus that could tolerate insertions and that displayed immunogenic Rift Valley fever virus epitopes was used as a vaccine to make mice resistant to infection by Rift Valley fever virus.

Expression of heterologous genes using alphavirus vectors

There are now several reports in the literature in which the Sindbis virus vectors have provided an
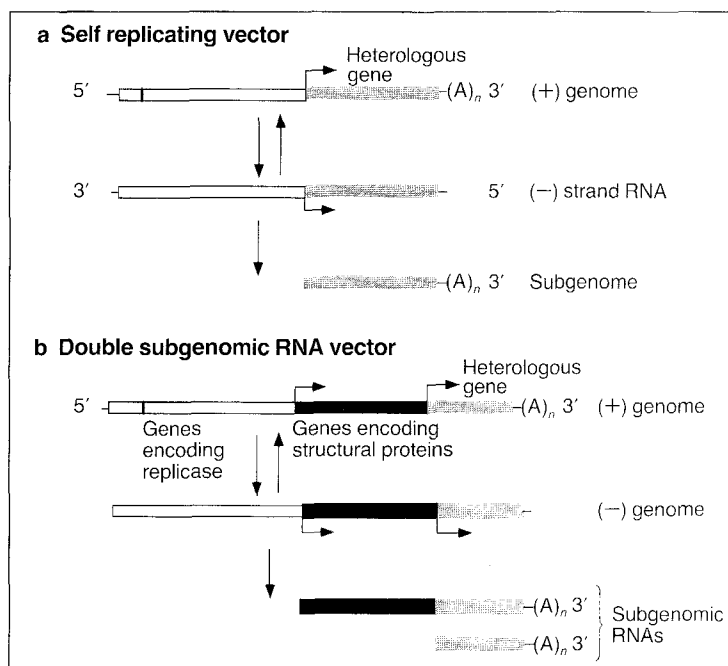
Figure 2

Two alphavirus vectors that express a heterologous gene. (**a**) Self-replicating vector[13,15]. This vector contains the viral genes and *cis*-acting sequences required for replication and transcription of the RNA, but lacks the structural protein genes. This vector is transfected into cells and does not produce any progeny virus. (**b**) Double subgenomic RNA vector[16]. This vector contains two promoters for the transcription of subgenomic RNAs. In this example the larger subgenomic RNA is translated to produce the viral structural proteins; the smaller is translated into the heterologous protein. The arrows on the (+) and (–) strands of RNA indicate the location of the promoters. The promoter on the (–) strand is recognized for transcription of the subgenomic RNAs. The vertical black line near the 5' terminus of the (+) genomes represents the sequence required for encapsidation of the RNA (the packaging signal)[12].
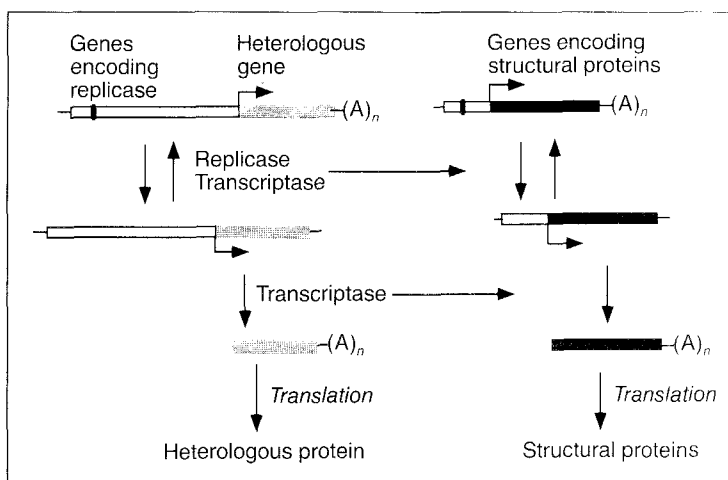
Figure 3

Complementation between alphavirus vector and helper RNAs. The vector RNA provides the genes coding for the nonstructural proteins essential for replication and transcription of the RNAs. Both vector and helper RNAs are replicated and transcribed. The helper subgenomic RNA codes for the structural proteins. In the illustration the helper RNA contains the packaging signal (represented by the vertical black line) and both vector and helper RNAs are packaged efficiently and are also co-packaged[17]. If the helper RNA lacks the packaging signal the vector RNA is packaged but the helper RNA is not (Semliki Forest virus vector) or is packaged much less efficiently (Sindbis virus vector).

important tool for studying biological questions. One of the first examples was the production of complete, infectious hepatitis B virus from the pregenomic RNA expressed from the Sindbis RNA vector[20]. This study demonstrated that this RNA contains all the information necessary for production of infectious hepatitis B DNA.

The double subgenomic RNA Sindbis viruses have been used in two studies. Hahn *et al.*[16] inserted either a truncated form of the influenza hemagglutinin (HA) protein, or minigenes encoding immunodominant cytotoxic T-cell epitopes into the Sindbis virus genome. These vectors were used to sensitize target cells for lysis by appropriate major histocompatibility complex-restricted, HA-specific, cytotoxic T cells. They were also used to infect mice to prime a specific T-cell response. In addition, the double subgenomic RNA vector was used to express the insulin-regulated glucose-transporter protein in cultured cells[21]. This protein, when expressed by the Sindbis vector, was found in adipocytes and muscle cells – its normal intracellular location.

The Semliki Forest virus vector has been used to express a variety of proteins, including dihydrofolate reductase (DHFR), the transferrin receptor, and lysozyme, proving that the vector can be used to express cytoplasmic, membrane and secreted proteins, respectively[15]. Studies on the expression of β-galactosidase indicated that about 400 μg of the heterologous protein could be produced from 5×10^6 cells, corresponding to approximately 25% of total cellular protein[15].

Conclusions

The initial studies with Sindbis and Semliki Forest virus suggest that both viruses are promising as vectors for heterologous gene expression. The potential of using alphaviruses to stimulate production of specific subsets of T cells by infection of animals is only just beginning to be explored. The use of alphaviruses to express proteins for production is ready to be exploited.

One particular aspect of using alphaviruses as vectors needs further investigation. Infection by these viruses is cytopathic for most vertebrate cells that are grown in culture. There are some variants that seem to be much less cytopathic and, for most of the Sindbis virus variants, infected mosquito cells in culture do not show cytopathic effects[22]. The mechanisms of cell killing are not known. Infection of cultured cells with the Semliki Forest virus vectors shuts off the synthesis of host-cell proteins, demonstrating that the viral structural proteins are not required for this effect[15]. The inhibition of host-cell protein synthesis may be due to the expression of a particular viral gene product, or it could be a consequence of the very high level of expression of the heterologous gene. It is possible to modulate the expression of the subgenomic RNA by the use of temperature-sensitive mutants or by the down-regulation of the subgenomic RNA promoter[13].

A second issue of importance is the possible biohazard of working with alphaviruses. The vectors derived from Semliki Forest virus that are unable to provide any viral structural proteins (termed 'suicide' vectors) should be safe for laboratory procedures, although the possibility of recombination and the production of segmented genome particles has not been fully resolved. There is one report of a fatal infection by Semliki Forest virus[23], so caution in using this virus must be maintained. Most strains of Sindbis virus are not pathogenic and so they may be preferred for many purposes, particularly for those investigators who are less familiar with working with infectious agents. There are many ways that these viruses can be attenuated and made safe for routine work, so that in the future their use should become essentially risk-free.

The initial studies on the development of alphaviruses as vectors uncovered the phenomenon of recombination. It also led to some important discoveries about these viruses (such as the identification of the packaging signal and a definition of the subgenomic RNA promoter). In the future, continued work on these vectors should not only provide us with improved expression systems, but also reveal other interesting facets of the biology of these viruses.

References

1 Levine, A. J. (1992) *Viruses*, W. H. Freeman & Co.
2 Bredenbeek, P. J. and Rice, C. M. (1992) *Semin. Virology* 3, 297–310
3 Huang, H. V., Rice, C. M., Xiong, C. and Schlesinger, S. (1989) *Virus Genes* 3, 85–91
4 Baltimore, D. (1971) *Bacteriol. Rev.* 35, 235–241
5 Luytjes, W., Krystal, M., Enami, M., Parvin, J. D. and Palese, P. (1989) *Cell* 59, 1107–1113
6 Collins, P. L., Mink, M. A. and Stec, D. S. (1991) *Proc. Natl Acad. Sci. USA* 88, 9663–9667
7 Almond, J. W. and Burke, K. L. (1990) *Semin. Virology* 1, 11–20
8 Schlesinger, S. and Schlesinger, M. J. (1990) in *Virology* (Fields, B. N. and Knipe, D. M., eds), pp. 697–711, Raven Press
9 Strauss, E. G. and Strauss, J. H. (1986) in *The Togaviridae and Flaviviridae* (Schlesinger, S. and Schlesinger, M. J., eds), pp. 35–90, Plenum Press
10 Levis, R., Schlesinger, S. and Huang, H. V. (1990) *J. Virol.* 64, 1726–1733
11 Raju, R. and Huang, H. V. (1991) *J. Virol.* 65, 2501–2510
12 Weiss, B., Nitschko, H., Ghattas, I., Wright, R. and Schlesinger, S. (1989) *J. Virol.* 63, 5310–5318
13 Xiong, C., Levis, R., Shen, P., Schlesinger, S., Rice, C. M. and Huang, H. V. (1989) *Science* 243, 1188–1191
14 Felgner, P. L. and Ringold, G. N. (1989) *Nature* 337, 387
15 Liljeström, P. and Garoff, H. (1991) *Bio/Technology* 9, 1356–1361
16 Hahn, C. S., Hahn, Y. S., Braciale, T. J. and Rice, C. M. (1992) *Proc. Natl Acad. Sci. USA* 89, 2679–2683
17 Geigenmüller-Gnirke, U., Weiss, B., Wright, R. and Schlesinger, S. (1991) *Proc. Natl Acad. Sci. USA* 88, 3253–3257
18 Weiss, B. and Schlesinger, S. (1991) *J. Virol.* 65, 4017–4025
19 London, S. D., Schmaljohn, A. L., Dalrymple, J. M. and Rice, C. M. (1991) *Proc. Natl Acad. Sci. USA* 89, 207–211
20 Huang, M. and Summers, J. (1991) *J. Virol.* 65, 5435–5439
21 Piper, R. C., Tai, C., Slot, J. W., Hahn, C. S., Rice, C. M., Huang, H. and James, D. E. (1992) *J. Cell Biol.* 117, 729–743
22 Brown, D. T. and Condreay, L. D. (1986) in *The Togaviridae and Flaviviridae* (Schlesinger, S. and Schlesinger, M. J., eds), pp. 171–207, Plenum Press
23 Willems, W. R., Kaluza, G., Boschek, C. B. and Bauer, H. (1979) *Science* 203, 1127–1129