Molecular anatomy of viral RNA-directed RNA polymerases

Brief Review

A. Ishihama[1] and **P. Barbier**[2]

[1] National Institute of Genetics, Department of Molecular Genetics, Mishima,
Shizuoka, Japan,
[2] John Innes Institute, Department of Virus Research, Norwich, U.K.

Introduction: the RNA genomes

All cellular organisms contain DNA as the genome, which is transcribed into RNA for gene expression, whereas the RNA genome remains only in the world of viruses. The enzyme, RNA-dependent RNA polymerase, is absolutely required for replication of the RNA genome and it is henceforth called "RNA replicase". However, it also functions as "RNA transcriptase" to produce mRNA for gene expression. Biochemical characterization of the RNA-dependent RNA polymerase was initiated more than 30 years ago soon after the discovery of RNA phages. In spite of numerous efforts, the purification of RNA polymerase has been successful only for a limited number of bacterial, animal and plant viruses. Difficulty in molecular analysis arose from the low level of viral RNA polymerase production, their low extractability, the instability of enzyme activity, and the complexity of enzyme structure and functions.

Since experimental techniques for gene cloning and nucleic acid sequencing were developed, a new trend in virus research has come of age, relying on interpretation of the nucleotide sequences of the RNA genome. The availability of many RNA polymerase sequences led to recognition of several conserved motifs among the RNA polymerases from different groups of RNA viruses. For instance, a unique three-amino-acid sequence, GDD, is now recognized as a hallmark of viral RNA polymerases. Further, viral RNA polymerases can now be over-produced from cDNA clones, for in vitro analysis of their structure and functions. In addition, the establishment of transfection systems of cDNA and cDNA-directed transcripts provided a new tool for testing in vivo functions. The role of each conserved motif in RNA polymerase molecules can now be examined by a combination of localized mutagenesis of cDNA, in vitro assays using over-produced proteins, and in vivo transfection assays.

This article aims at reviewing of our up-to-date knowledge of the structure-function relationships in RNA-dependent RNA polymerases from bacterial, animal and plant RNA viruses. Since the replication strategies of the RNA viruses are reflected in the structure and function of the RNA-dependent RNA polymerase, our effort is focussed on the comparison of the RNA polymerases between viruses sharing the same genetic systems, but including bacterial, animal and plant viruses. The viruses largely discussed in this paper are listed in Table 1. Historical background related to these topics has been described in a previous review [37], which gives the original references not cited in this review.

RNA polymerases of plus-strand RNA viruses

The genomic RNA of plus-strand RNA viruses can be directly translated into viral proteins including viral RNA polymerase, which then participates in both of the two-step reactions: viral RNA (vRNA)-directed synthesis of complementary RNA (cRNA) of negative polarity; and cRNA-directed vRNA synthesis. The molecular anatomy of the RNA polymerase from this group viruses has been investigated in three different ways: i) biochemical isolation and characterization of the RNA polymerase from virus-infected cell extracts; ii) cloning and sequence comparison of the RNA polymerase genes; and iii) in vitro mutagenesis of cDNA for the polymerase genes and in vivo transfection assays of cDNA transcripts.

RNA polymerases of single-strand RNA phages

The majority of RNA phages have about 3.5 kb (kilobase)-long plus-strand RNA as the genome. The catalytic subunit, β, of the enzyme for RNA replication is encoded by the phage genome, but it associates with three bacterial proteins, ribosomal protein S1 (subunit α), and translation elongation factors Tu (γ) and Ts (δ) to form the RNA polymerase holoenzyme [8]. Ribosomal protein S1 (α subunit) carries RNA helicase activity [16], which may be involved in strand-separation of RNA duplexes formed during RNA replication. In addition a host factor (HF-I) of 11 kDa (kilodalton) is required, but only for the first step of RNA replication, i.e., vRNA-directed cRNA synthesis. The hfq gene coding for HF-I was recently cloned from the E. coli chromosome [41]. The function of this protein in E. coli is not yet known.

There are four groups of RNA phages infecting E. coli, each serologically distinct from the others (Table 1). Among the amino acid sequences of the RNA polymerase from MS2 (group I), GA (group II), Q_β (group III) and SP (group IV), there is a great deal of conservation throughout the central region, which suggests the terminal regions are more tolerant to small changes in amino acid sequences than are the internal regions. The consensus YGDD sequence, common to all these RNA phages, also exists in most other viral RNA-dependent RNA polymerases [42]. The pioneering study by Inokuchi and Hirashima [35]

Table 1. Viruses described in this review

Genome type Host	Virus family	Virus species
Plus-strand RNA		
Bacteria	Group-I	Phage MS2 (MS2)
	Group-II	Phage GA (GA)
	Group-III	Phage Q_β (Q_β)
	Group-IV	Phage SP (SP)
Plants	Ilarvirus	Alfalfa mosaic virus (AlMV)
	Bromovirus	Bromo mosaic virus (BMV)
		Cowpea chlorotic mosaic virus (CCMV)
	Cucumovirus	Cucumber mosaic virus (CMV)
	Tymovirus	Turnip yellow mosaic virus (TYMV)
	Tobamovirus	Tobacco mosaic virus (TMV)
Animals	*Picornaviridae*	Polio virus (POL)
	Flaviviridae	Yellow fever virus (YFV)
	Alphaviridae	Sindbis virus (SND)
		Semliki Forest virus (SFV)
	Coronaviridae	Infectious bronchitis virus (IBV)
		Murine coronavirus (MCV)
	Toroviridae	Berne virus (BEV)
Double-strand RNA		
Bacteria	*Reoviridae*	Phage $\phi 6$ ($\phi 6$)
Animals	*Reoviridae*	Human reovirus (HRV)
		Bovine rotavirus (BRV)
		Infectious bursal disease virus (IBDV)
Plants	Phytoreovirus	Rice ragged stunt virus (RRSV)
	Cryptovirus	White clover crypic virus (WCCV)
Minus-strand RNA		
Animals	*Rhabdoviridae*	Vesicular stomatitis virus (VSV)
	Paramyxoviridae	Sendai virus (HVJ)
	Myxoviridae	Influenza virus A (INF)
Plants	*Rhabdoviridae*	Wheat rosette stunt virus (WRSV)
Ambisense RNA		
Animals	*Arenaviridae*	Pichinde virus (PIV)
	Bunyaviridae	Punta Toro virus (PTV)
		La Crosse virus (LCV)
Plants	Tenuivirus	Rice stripe virus (RSV)
	Bunaryvirus	Tomato spotted wilt virus (TSWV)

showed that amino acid substitutions within this sequence specifically impaired the functions of Q_β polymerase. They also proposed that this motif is involved in the binding of Zn^{2+}, an essential cofactor of all the enzymes involved in nucleic acid synthesis.

Recombination of phage RNA takes place by a copy choice mechanism during RNA replication [61]. It is a source of self-replicative variant RNA

species and of new vRNA-associated sequences such as parts of tRNA [7, 57]. Such variant or unrelated RNA sequences may arise as a result of template switching by the viral RNA polymerase or by the RNA polymerase jumping between templates [37].

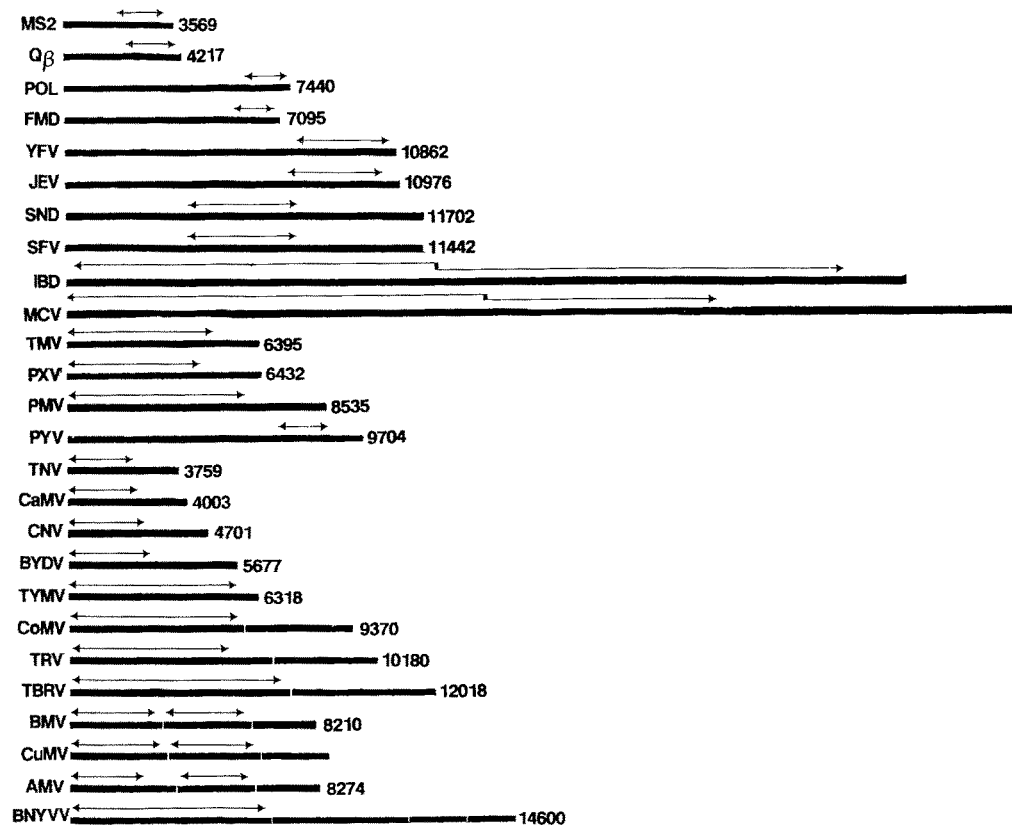## RNA polymerases of plus-strand animal RNA viruses

Picornavirus RNA polymerase family

Poliovirus, the prototype of the family *Picornaviridae*, is a plus-strand RNA virus with a genome of about 7.5 kb, which is translated in virus-infected cells into a single large polyprotein precursor. The RNA polymerase core enzyme (the 3D gene product), primer protein VPg, and polyprotein processing protease (3C) are all located in the C-terminal proximal region (region 3) of this precursor polyprotein (Fig. 1). Processing of the precursor polyprotein also takes place in the membrane fraction, concomitantly with the formation of complexes with transcription and replication activities. The structure of the transcription/replication apparatus is not yet known, except that the 3D protein forms its core.
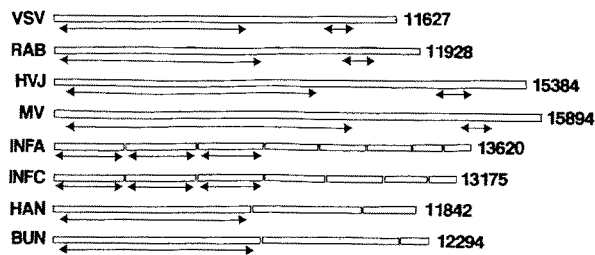
By treatment of the transcription/replication complex with high concentrations of salt, the 3D polymerase can be solubilized, and exhibits activities of primer- and template-dependent RNA synthesis, such as oligo(U)-dependent poly(A)-directed poly(U) synthesis and oligo(U)-dependent poliovirus RNA-directed minus-strand RNA synthesis (reviewed in [37]). A viral RNA polymerase with similar functional specicities has also been isolated from the cytoplasm of polio virus-infected cells. Product RNAs in these reactions are covalently linked to the primers added. The primer requirement can be replaced by a host factor with RNA uridylation activity (or terminal uridyltransferase; TUTase). In this case, newly synthesized oligo(U) is covalently linked to the vRNA-poly(A) tail

----

Fig. 1. The genomes of RNA viruses. The relative sizes of the genome RNAs from the prototype viruses of each virus family or group are shown from 5′ to 3′. The numbers shown at the 3′ termini represent the total length (in nucleotides). The gene loci encoding the RNA-dependent RNA polymerase are indicated above (in the case of plus-strand RNA) or below (in the case of minus-strand RNA) genome RNAs. *MS2* Phage MS2; Q_β phage Q_β; *POL* poliovirus; *FMD* foot-and-mouth disease virus; *YFV* yellow fever virus; *JEV* Japanese encephalitis virus; *SND* Sindbis virus; *SFV* Semliki forest virus; *IBD* infectious bursal disease virus; *MCV* mouse corona virus; *TMV* tobacco mosaic virus; *PXV* potato virus X; *PMV* potato virus M; *PYV* potato virus Y; *TNV* tobacco necrosis virus; *CaMV* carnation mottle virus; *CNV* cucumber necrosis virus; *BYDV* barley yellow dwarf virus; *TYMV* turnip yellow mosaic virus; *CoMV* cowpea mosaic virus; *TRV* tobacco rattle virus; *TBRV* tomato black ring virus; *BMV* brome mosaic virus; *CuMV* cucumber mosaic virus; *AMV* alfalfa mosaic virus; *BNYVV* beat necrotic yellow vein virus; *VSV* vesicular stomatitis virus; *RAB* rabies virus; *HVJ* hemagglutinin virus of Japan (Sendai virus); *MV* measles virus; *INFA* influenza virus A; *INFC* influenza virus C; *HAN* Hantaan virus; *BUN* Bunyamwaera virus; *UUK* Uukuniemi virus; *TSWV* tomato spotted wilt virus; *LCM* lymphatic choriomeningitis virus; *TAC* Tacaribe virus; *PIC* Pichinde virus; *HRV* human rotavirus
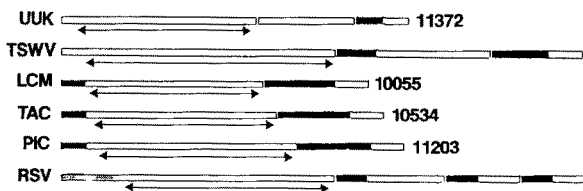
Plus-Strand RNA Viruses

MS2 ▬▬▬ 3569
Qβ ▬▬▬ 4217
POL ▬▬▬▬ 7440
FMD ▬▬▬▬ 7095
YFV ▬▬▬▬▬▬ 10862
JEV ▬▬▬▬▬▬ 10976
SND ▬▬▬▬▬▬ 11702
SFV ▬▬▬▬▬▬ 11442
IBD ▬▬▬▬▬▬▬▬▬▬▬
MCV ▬▬▬▬▬▬▬▬▬▬▬▬
TMV ▬▬▬▬ 6395
PXV ▬▬▬▬ 6432
PMV ▬▬▬▬▬ 8535
PYV ▬▬▬▬▬▬ 9704
TNV ▬▬ 3759
CaMV ▬▬ 4003
CNV ▬▬ 4701
BYDV ▬▬▬ 5677
TYMV ▬▬▬ 6318
CoMV ▬▬▬▬▬ 9370
TRV ▬▬▬▬▬ 10180
TBRV ▬▬▬▬▬▬ 12018
BMV ▬▬▬▬ 8210
CuMV ▬▬▬▬
AMV ▬▬▬▬ 8274
BNYVV ▬▬▬▬▬▬▬ 14600

Minus-Strand RNA Viruses

VSV ▭▭▭ 11627
RAB ▭▭▭ 11928
HVJ ▭▭▭▭ 15384
MV ▭▭▭▭ 15894
INFA ▭▭▭ 13620
INFC ▭▭▭ 13175
HAN ▭▭▭ 11842
BUN ▭▭▭ 12294

Ambi-Sense RNA Viruses

UUK ▭▭▭ 11372
TSWV ▭▭▭▭▭
LCM ▬▭▭ 10055
TAC ▬▭▭ 10534
PIC ▬▭▭ 11203
RSV ▭▭▭▭▭

Double-Strand RNA Viruses

HRV ▬▬▬▬▬▬▬▬ 18247

and hybridizes with the poly(A) sequence through a snap-back mechanism, allowing the oligo(U) to serve as a primer for elongation of cRNA chains. As a result the RNA products are twice the size of the template vRNA.

The cap structures commonly found in eukaryotic mRNA are absent from the 5' terminus of poliovirus RNA, but instead a small virus-coded protein, VPg, 22 amino acid residues in length, is covalently attached to the vRNA 5' terminus. After infection, VPg is removed from vRNA prior to translation. VPg is attached not only to the vRNA but also to the cRNA. Thus, the VPg may have a signalling function for initiation of RNA synthesis.

The 3D protein expressed in *E. coli* was found to be active in primer- and template-dependent elongation of RNA chains [55, 63, 73]. This new experimental system became available for detailed systematic analysis of the effect of amino acid substitution in the RNA polymerase (for example see [14]). Systematic mutagenesis affecting the conserved YGDD sequence of poliovirus RNA polymerase indicated that substitutions of the Tyr and Gly residues gave profound effects on the in vitro enzymatic activity [38, 39]. Most alteration of these conserved amino acids result in non-functional RNA polymerases. The capability of the mutant 3D polymerases to function in poliovirus replication in vivo was also examined by cDNA transfection assay. Both in vivo and in vitro activities are retained only in a few cases (only Phe can replace Tyr; and only Ala or Ser can replace Gly).

The genomes of flaviviruses resemble those of picornaviruses in the gene organization: the genes encoding seven non-structural proteins are located in the 3' terminal-proximal region (Fig. 1). Viral non-structural protein NS5 carries the structural motifs for viral RNA polymerase (Fig. 2). Although the RNA polymerase has not yet been purified, NS5 is always associated with membrane fractions with the RNA polymerase activity. In spite of the similarity in gene organization between picornaviruses and flaviviruses, the RNA polymerase sequences of flaviviruses are more closely related to those of alphaviruses than to those of picornaviruses. However, the flavivirus RNA polymerase is more closely related to that of plant luteoviruses [13, 45].

Alphavirus RNA polymerase family

Most of our current knowledge on the transcription and replication of alphaviruses came from studies of two closely related viruses, Semliki Forest virus (SFV) and Sindbis virus (SND). SND, the prototype virus of the genus alphavirus, contains an 11.7 kb plus-strand RNA genome (Fig. 1), which is capped and polyadenylated. In sharp contrast to the gene organization found in picornaviruses and flaviviruses, the genes for non-structural proteins are located at the 5'-terminal proximal region. Four non-structural proteins (NS1, NS2, NS3 and NS4) are generated after proteolytic processing of two polyprotein precursors, NS123 or NS1234, formed by translation of full-size vRNA. The conserved motifs for viral RNA polymerase are located in NS4 (Fig. 2), while NS1 carries the sequences involved in RNA capping [17]. Pro-

cessing of the polyprotein precursors is catalyzed autolytically by the NS2 protease.

The NS4 RNA polymerase is integrated into transcription/replication complexes associated with the infected cell membranes. This form of NS4 polymerase is metabolically stable, but unassembled RNA polymerase is rapidly degraded by the ubiquitin-dependent pathway, as it carries the N-end rule-based degradation signal or N-degron [19]. Some incomplete SND RNA species contain, at their 5' termini, sequences identical to a cellular tRNA^{Asp} [37], whereas their 3'-terminal sequences are identical to that of the parental infectious viral RNA. This tRNA species might serve as a primer for vRNA synthesis, but be normally removed at the end of polymerization.

Coronavirus RNA polymerase family

Coronaviruses contain a single molecule of large RNA (more than 20 kb in length) of positive-polarity as the genome (Fig. 1). The expression of the viral genes takes place through transcription of cRNA into gene-specific subgenomic mRNAs. All mRNA species carry a common sequence at the 5' terminus. This 5' coterminal structure is a transcript of the cRNA 3' terminal region. The RNA polymerase genes of two coronaviruses sequenced so far, avian coronavirus IBV (infectious bronchitis virus) and murine coronavirus (MCV) MH-A59, consists of two ORFs of about 12 (ORF1a) and 8 (ORF1b) kb in length, encoded in the 5'-proximal region of vRNA [10, 11]. The complete RNA polymerase is formed by a translation frameshift from the 3' end of ORF1a to the 5' end of ORF1b (Fig. 2). This process involves specific RNA sequence elements at the ORF1a/ORF1b boundary, including an RNA pseudoknot, and efficient ribosome frameshifting takes place even in an in vitro translation system [12]. This unusual expression strategy of the coronavirus RNA polymerase gene may be related to the generation of a mosaic molecule, with a picornavirus-like N-terminal region and alphavirus-like C-terminal region [45]. The complete translation product of 250 kDa is, however, subsequently cleaved again to yield 28- and 220-kDa functional proteins [5].

Berne virus (BEV), the prototype of the *Toroviridae*, is similar to coronaviruses in genome organization and expression mode [77]. Sequence analysis revealed at least five open reading frames on the BEV genome. The amino acid sequence identity between BEV and IBV/MHV is the highest in the RNA polymerase gene. The RNA polymerase gene is again split into two ORFs (ORF1a and ORF1b) which overlap by 12 nucleotides, indicating that the RNA polymerase protein is translated by ribosomal frameshifting. In contrast to coronaviruses, however, the BEV mRNAs are 3' coterminal but no common leader sequence has been detected at their 5' end.

RNA polymerases of plus-strand plant viruses

The vast majority of plant viruses (about 70%) are plus-strand RNA viruses, which can be broadly divided on the basis of RNA polymerase structure into

picornavirus-like and alphavirus-like groups [21]. The hypothesis that an RNA-dependent RNA polymerase present in uninfected plant cells is activated by virus infection and is used for virus replication, was disproved by the finding of RNA polymerase genes in the viral genomes (for reviews see [18, 66]). Isolation and characterization of viral RNA polymerases from infected cells, however, have been hindered by two factors: first, the presence of an inducible host RNA-dependent RNA polymerase activity; second, the low extractability of active RNA polymerases from cellular membranes and their tight association with endogenous RNA templates.

Tripartite virus RNA polymerase family

The viral RNA synthesis observed in cell extracts of virus-infected plants involves mainly elongation of preexisting nascent RNA chains. RNA-free RNA polymerases have so far been solubilized and partially purified from plants infected with several tripartite viruses including cowpea chlorotic mottle virus (CCMV) [53], turnip yellow mosaic tymovirus (TYMV) [56], brome mosaic virus (BMV) [67] and alfalfa mosaic virus (AMV) [69]. These enzymes are, however, able to catalyze only vRNA-directed cRNA synthesis (the first-step reaction of RNA replication), although the initiation of RNA synthesis depends upon the addition of exogenous vRNA templates. Complete replication in vitro, i.e., the synthesis not only of minus-strand cRNA but also of plus-strand vRNA, has been successfully achieved only in the case of RNA polymerase prepared from cucumber mosaic virus (CMV)-infected tobacco leaves [30]. None of these enzymes were, however, pure enough to relate the observed functional difference between different enzyme preparations to any structural difference.

RNA polymerases of tripartite viruses contain two viral proteins, which are encoded in two large vRNA segments. The RNA polymerase prepared from BMV-infected cells, however, contained up to 12 polypeptides, among which two polypeptides corresponded to virus-encoded non-structural proteins, 1a (the RNA 1 gene product) and 2a (the RNA 2 gene product) [67]. RNA polymerases encoded by two other tripartite viruses, CCMV and AMV, also contain two gene products: proteins 1a and 2a (CCMV); and P1 and P2 (AMV), respectively. Sequence comparisons indicate that the RNA 1 product corresponds to the NS1 and NS2 proteins of alphaviruses, while the RNA 2 protein corresponds to NS4 [3] (Fig. 2). The conserved structural domains for NTP-binding/helicase activities are located in the former, while the GDD motif is present in the latter [2], confirming the requirement of both components for the RNA polymerase functions. The RNA 1 protein also carries a sequence similar to alphavirus NS1 with the motifs for RNA capping functions. A novel approach has been developed recently for the identification of viral proteins involved in transcription and replication: complementation of a defective virus could be achieved by heterologous expression of viral RNA polymerase proteins in transgenic plants for BMV [54] and AMV [84].

Besides the two viral coded subunits, several host proteins seem to be associated with transcription/replication complexes. A cellular protein of 50 kDa

copurifies along the CMV RNA polymerase [30] and loss of this 50 kDa host factor always correlates with loss of the CMV RNA polymerase activity. The 3' regions of some plant virus RNAs can be folded to form tRNA-like structures which can be aminoacylated by cellular enzymes, and thus any host protein interacting with aminoacyl-tRNA may also bind to aminoacylable viral RNAs. In fact, an affinity-purified 41 kDa protein on BMV 2a protein-column, with stimulation activity of the BMV RNA polymerase, was recently identified as translation elongation factor eIF3 [68]. The successful propagation of a BMV-derived replicon, established in the yeast *S. cerevisiae* by expressing the BMV RNA polymerase proteins [40], indicates that any host factor necessary for viral RNA replication must be conserved in evolution. On the other hand, the similarity of the cis-acting sequence at the 5' termini of BMV RNAs with the internal control region of tRNAs raises the possibility that a host RNA polymerase III transcription factor could be the trans-acting factor binding this BMV sequence [49].

Mono- and bipartite virus RNA polymerase family

In the mono- and bipartite virus genomes, the gene for RNA polymerase is encoded in a single RNA molecule. For instance, monopartite tobamoviruses contain a single large RNA molecule as their genome, from which two translation products with the RNA polymerase motifs are formed. In the case of tobacco mosaic virus (TMV), a 126 kDa polypeptide is translated from the vRNA 5'-terminus, while a 183 kDa polypeptide is produced by read-through of the translation termination codon of the 126 kDa protein. The NTP-binding site is located in the 126 kDa protein, but the GDD motif is present only in the C-terminal region of 183 kDa read-through protein (Fig. 2). In agreement with the sequence analysis, photoaffinity labeling experiments indicated that nucleoside triphosphates can be cross-linked to p126 [24]. A methyltransferase activity was also detected in this polypeptide [23]. Membrane-associated transcription/replication complexes contain both p126 and p183 [90]. The putative RNA polymerase gene of the monopartite tymovirus TYMV is first translated into a 206 kDa polyprotein, which is then processed into three fragments, an N-terminal 120 kDa protein with the methyltransferase domain, a 30 kDa protein with the helicase motif, and a C-terminal 78 kDa protein with the GDD motif.

Several lines of evidence support the involvement of host factors in the functions of this group RNA polymerases. For example, molecular interaction between TMV RNA polymerase and a host protein encoded by the tomato resistance gene Tm1 has been suggested, because resistance-breaking mutations maps to p126 [52].

In the case of dipartite cowpea mosaic virus (CPMV), vRNA segment 1 is translated to yield a polyprotein of 200 kDa, which is then processed into a 58 kDa putative helicase, a 4 kDa VPg and a C-terminal 110 kDa protein with the GDD motif [22] (Fig. 2). The 110 kDa protein produced using a baculovirus expression system is, however, inactive in RNA synthesis [83]. In virus-infected cells, the 110 kDa RNA polymerase is further processed into the 87 kDa form.
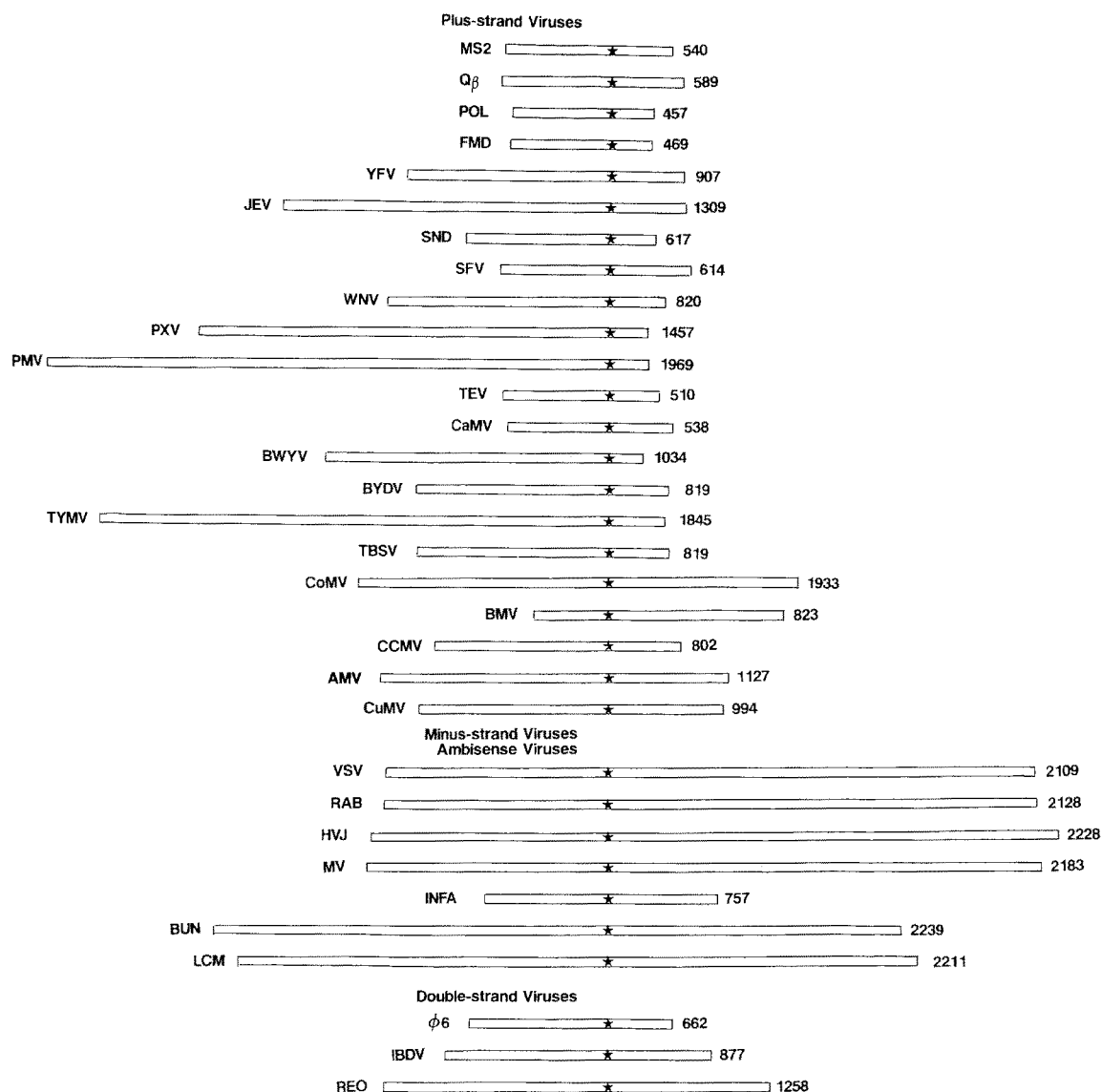
　　　　　　　　A. Ishihama and P. Barbier

Plus-strand Viruses

MS2 ⊏━━━★━━⊐ 540
Qβ ⊏━━━━★━━⊐ 589
POL ⊏━━━★━⊐ 457
FMD ⊏━━━★━⊐ 469
YFV ⊏━━━━━━★━━━⊐ 907
JEV ⊏━━━━━━━★━━━⊐ 1309
SND ⊏━━━━★━⊐ 617
SFV ⊏━━━━★━⊐ 614
WNV ⊏━━━━━★━⊐ 820
PXV ⊏━━━━━━━━★━⊐ 1457
PMV ⊏━━━━━━━━━━★━⊐ 1969
TEV ⊏━━━★━⊐ 510
CaMV ⊏━━━━★━⊐ 538
BWYV ⊏━━━━━━★━⊐ 1034
BYDV ⊏━━━━★━⊐ 819
TYMV ⊏━━━━━━━━━★━⊐ 1845
TBSV ⊏━━━━★━⊐ 819
CoMV ⊏━━━━━━★━━━━⊐ 1933
BMV ⊏━━━★━━━⊐ 823
CCMV ⊏━━━━★━⊐ 802
AMV ⊏━━━━━★━━⊐ 1127
CuMV ⊏━━━━━★━━⊐ 994

Minus-strand Viruses
Ambisense Viruses

VSV ⊏━━━━━★━━━━━━━━━⊐ 2109
RAB ⊏━━━━━★━━━━━━━━━⊐ 2128
HVJ ⊏━━━━━★━━━━━━━━━━⊐ 2228
MV ⊏━━━━━★━━━━━━━━━⊐ 2183
INFA ⊏━━━━★━━⊐ 757
BUN ⊏━━━━━★━━━━━⊐ 2239
LCM ⊏━━━━━★━━━━━⊐ 2211

Double-strand Viruses

φ6 ⊏━━━★━⊐ 662
IBDV ⊏━━━━★━⊐ 877
REO ⊏━━━━★━⊐ 1258

Fig. 2. Viral RNA-dependent RNA polymerases. Viral proteins involved in RNA replication are aligned from N to C terminal direction as to match the GDD core motif (marked by star). The numbers shown at the C termini represent the total length (in amino acid residues). In the case of plus-strand virus RNA polymerases, the GDD motif is located in the C-terminal region, while that is located in the middle of RNA polymerase from minus-strand viruses. *MS2* Phage MS2; *Qβ* phage Qβ; *POL* polio virus; *FMD* foot-and-mouth disease virus; *YFV* yellow fever virus; *JEV* Japanese encephalitis virus; *SND* Sindbis virus; *SFV* Semliki Forest virus; *WNV* West Nile virus; *PXV* potato virus X; *PMV* potato virus M; *TEV* tobacco etch virus; *CaMV* carnation mottle virus; *CoMV* cowpea mosaic virus; *BMV* brome mosaic virus; *CCMV* cowpea chlorotic mottle virus; *AMV* alfalfa mosaic virus; *CuMV* cucumber mosaic virus; *VSV* vesicular stomatitis virus; *RAB* rabies virus; *HVJ* hemagglutination virus of Japan (Sendai virus); *MV* measles virus; *INFA* influenza virus A; *BUN* bunyamwaera virus; *LCM* lymphatic choriomeningitis virus; φ6 phage φ6; *IBDV* infectious bronchitis virus; *REO* human reovirus

At present, it is not known which is the functional form for transcription and replication. As in the case of animal picornaviruses, the replication complexes are tightly associated with the infected cell membranes and the synthesis of VPg-linked RNA is coupled with the polyprotein processing [85]. According to the picornavirus model, a host factor such as uridyltransferase would also be necessary for formation of primers for RNA synthesis.

RNA polymerases of double-strand RNA viruses

A small number of bacterial, animal and plant viruses carry double-stranded RNA as their genome. A feature in common to all double-strand RNA viruses is a virion-associated RNA polymerase activity, indicating that viral proteins can not be expressed directly from double-strand vRNAs. The RNA polymerase sequences of double-strand RNA viruses are, however, close to those of plus-strand viruses.

RNA polymerases of double-strand RNA phages

Phage φ6 is an enveloped phage carrying three double-strand RNA segments as its genome. It infects a number of phytopathogenic *Pseudomonas* strains. During phage maturation, the genome RNA segments are packaged into pre-formed empty particles (or procapsids) consisting of four early viral proteins, P1, P2, P4 and P7. The RNA-filled procapsids, which have RNA polymerase activity, are then coated with a shell of the viral protein P8, to form the nucleocapsids, and concomitantly the RNA polymerase activity is shut off. Uncoating of the nucleocapsids is an essential step in the initial stage of phage infection for the expression of the virion-associated polymerase activity [59]. Transcription of φ6 dsRNA resembles DNA replication, as being semiconservative and in sharp contrast to the conservative transcription mode of reo- and rotaviral double-strand RNA (see below).

Protein P2 has some sequence similarity to other viral RNA polymerases. It is present in about 10 to 20 molecules per virion, thus 3 to 7 molecules for each RNA segment. This protein is necessary not only for RNA synthesis but also for packaging of the genome RNA. One possible explanation is that P2 forms complexes with newly synthesized plus-strand copies as an assembly core, to which other capsid proteins associate to form nucleocapsid cores [27]. The synthesis of minus-strand copies may then take place, forming dsRNA, prior to association of the P8 shell. In any case, the P2 protein has dual functions as the catalytic unit for RNA polymerization and the nucleation core for virus assembly.

RNA polymerases of double-strand animal RNA viruses

Three genera of the *Reoviridae* infect animals, and three genera infect insects and plants. Reoviruses lack an envelope, but instead have an outer capsid composed of hexagonal and pentagonal arrangements of protein subunits. The double-

strand RNA genomes of all these genera are divided into ten or more segments. Transcription of the double-strand RNAs is catalyzed by a virion-associated RNA polymerase, expressing its intrinsic enzyme activity only after infection [37]. At early stages of infection, transcription is asymmetric, i.e., only plus-strand RNA is synthesized, and conservative, i.e., the newly synthesized RNA is released without strand separation of the template duplex. The synthesis of progeny double-strand RNA occurs through transcription of the newly synthesized plus-strand RNA as template. This conservative mechanism of double-strand RNA replication is completely different from the semiconservative replication mode of phage $\phi6$ RNA.

Either heat shock treatment or partial protease digestion allows the virion-associated RNA polymerase to express in vitro its enzymatic activities. K^+ triggers this switch-on of RNA polymerase functions, while Na^+ antagonizes it; divalent cations modulate the process. During activation of the core-associated RNA polymerase, outer capsid proteins are removed. The RNA polymerase activity disappears when virions are completely disrupted. The isolated inner core containing all RNA segments is inactive, suggesting that the RNA polymerase is located inside the core shell and is disassembled after removal of the outer shell. Thus, it has been difficult to isolate the RNA polymerase in an active form [37].

Bovine rotavirus RNA segment 1 encodes VP1 of 1088 amino acid residues, which carries the conserved sequences of RNA-dependent RNA polymerase [15], but VP2 encoded by RNA segment 2 is also an essential component of RNA polymerase [48]. VP2 of infectious bursal disease virus (IBDV) carries guanyltransferase and methyl transferase activities [78].

RNA polymerases of double-strand plant RNA viruses

As in the case of animal viruses, RNA polymerase activity has been found associated with purified particles of Phytoreoviruses [47, 50], Fijiviruses and Cryptoviruses [1, 9, 51]. Enzymes involved in RNA cap formation such as methyltransferase are also associated with double-strand RNA viruses [71]. Addition of the methyl donor, S-adenosylmethionine, stimulates transcription in vitro, indicating that the reactions of cap formation and RNA synthesis are coupled.

Since the 5' and 3' terminal sequences are conserved between RNA segments of Phytoreoviruses and exhibit some complementarity, they may function as both promoter and origin for transcription and replication, respectively, by the same RNA polymerase. Analysis of these cis-acting signals by site-directed mutagenesis is, however, precluded by the lack of a reconstituted in vitro system.

RNA polymerases of minus-strand RNA viruses

In this group of single-strand RNA viruses, the presence of RNA polymerase in virions is mandatory to initiate infection, since their RNA genomes are in the anti-sense form, and the genomic RNA cannot serve as mRNA for translation.

Thus, the naked vRNA alone is insufficient to initiate an infection. One of the unique characteristics of the RNA polymerase of negative-strand RNA viruses is that it transcribes the RNA genome only in assembled nucleocapsid form. Negative-strand animal RNA viruses fall into two major groups, one carrying the non-segmented RNA genomes and the other containing segmented RNA genomes. Up to the present time, the RNA polymerase has been purified from vesicular stomatitis virus (VSV) of non-segmented animal RNA viruses and influenza virus of segmented animal RNA viruses.

RNA polymerases of minus-strand animal RNA viruses

Rhabdovirus RNA polymerase family

Most of the information on transcription and replication of non-segmented viruses has come from studies with VSV, the prototype virus of the *Rhabdoviridae*. The genome of VSV is a linear single-strand RNA of about 11.2 kb (Fig. 1). The ribonucleoprotein (RNP) core contains the genome RNA, which is tightly bound with nucleocapsid protein (NP), and two other minor proteins, L (RNA polymerase β subunit) and NS (α subunit). Both L and NS are individually inactive, but when combined, they regain the activity of N-vRNA complex-dependent RNA synthesis, indicating that both L and NS are essential subunits of the VSV RNA polymerase holoenzyme (reviewed in [37]).

The VSV genome contains, reading from the 3′ to the 5′ terminus, a leader sequence, five genes in the order N-NS-M-G-L, and the 5′ untranslated trailer. Using purified virions, five monocistronic mRNAs are synthesized in vitro in a sequential fashion, due to obligatory entry of the RNA polymerase at the promoter located at the 3′ end of the vRNA. Two contradictory mechanisms of transcription have been proposed [37]: either the five genes are transcribed into a single polycistronic transcript, which is later processed to monocistronic RNA ("RNA processing model"); or transcription is initiated at the beginning of each gene and terminated at its end ("termination and reinitiation model"). Transcription initiated from the template 3′ terminus is attenuated at the end of the leader sequence, releasing short leader RNA; concurrent encapsidation of these nascent RNAs by the NP protein affects the level of read-through beyond this point. In this connection, it is interesting that the host protein La, a component of ribonucleoprotein complexes made from RNA polymerase III transcripts, specifically binds to this VSV leader RNA [87].

The product RNAs of in vitro transcription are capped and polyadenylated. Hence, purified VSV virions contain enzyme systems for RNA modification. The RNA modification activities are associated with the RNA polymerase holoenzyme (NS-L or αβ complex). The L protein is multifunctional: the catalytic site for RNA polymerization is located on the L protein [60], and thus the L protein is involved in both transcription and replication; the site on L for polyadenylation is, however, different from the catalytic site for RNA polymerization [34], and the site involved in cap formation is also different from the site for RNA synthesis. In addition, the L protein carries a protein kinase activity

for phosphorylation of the NS protein [75]. Protein phosphorylation controls not only RNA synthesis but also virus uncoating [88]. Since an oligopeptide with the same amino acid sequence as the C-terminal proximal region of NS interferes with transcriptase activity of the L protein [89], the C-terminal domain of NS may be involved in protein-protein contact with L. Expression of the cDNA clones for both L and NS [76] will enable detailed mapping of the functional sites on these RNA polymerase subunits.

An alternative function of the RNA polymerase is the synthesis of full-length cRNA, which serves as the template for vRNA synthesis. The virus infection cycle involves a balance between the two functions of the RNA polymerase, transcription and replication. Several different mechanisms have been proposed for the switching of RNA polymerase from transcription to replication [37]. The N protein is needed for vRNA to function as the template for RNA synthesis; in addition, excess N associates with nascent RNA products and the formation of N-RNA complexes prevents transcription attenuation, allowing to produce the genome-size read-through cRNA. The NS protein also plays a role in the switching control, because the NS protein interacts not only with L but also with NP. NS is an exceptional protein as to the extent of its phosphorylation: there are, in all, 33 potential sites of phosphorylation, 12 threonine and 21 serine residues, which mostly reside in the N-terminal proximal region. The L-NS interaction is controlled by the level of NS phosphorylation. A heavily phosphorylated and extremely acidic N-terminal domain constitutes a novel structure that resembles the phosphate backbone of RNA, to which NP binds tightly. The formation of NS-NP complexes prevents self-aggregation of free NP. Furthermore, NS displaces RNA-bound NP in a localized manner, thereby allowing the RNA polymerase to gain access to the template RNA.

Paramyxovirus infections share many characteristics with those of rhabdoviruses. The RNP cores contain one molecule of non-segmented vRNA, approximately 2500 molecules of NP, 100–300 molecules of P (equivalent to VSV NS) with a molecular weight of 50–55 kDa, and 50–100 molecules of L protein with a molecular weight of about 200 kDa. The functional RNA polymerase is composed of L(β) and P(α) proteins [37]. The addition of cellular factors stimulate transcription in vitro by detergent-treated paramyxovirus virions or RNP cores isolated from virus particles. Tubulin is a host factor for the RNA polymerase activity associated with VSV and HVJ (Sendai virus). On the other hand, microtubule-associated proteins (MAPs) stimulates in vitro RNA synthesis by HVJ (or Sendai virus). The synthesis and assembly of RNP cores in infected cells take place in association with the cellular cytoskeletal framework.

Orthomyxovirus RNA polymerase family

The genome of influenza virus is composed of one molecule each of eight (A and B type viruses) or 7 (C type virus) RNA segments of negative polarity (Fig. 1). In the early stages of virus infection, two types of vRNA copy are produced:

mRNA with a 5'-cap and 3'-poly(A) tail and full-length cRNA without cap structure and poly(A) tail. At late stages, however, progeny vRNA is synthesized by copying cRNA. The growth of influenza virus is unique because continued synthesis of host cell nuclear RNA is required for viral transcription to occur, as RNA molecules synthesized by RNA polymerase II are used as primers for primary transcription of influenza virus [37]. The first step in capped RNA-primed transcription is the cleavage of existing capped RNA by a "capped RNA endonuclease" associated with influenza viral RNA polymerase. Viral RNA polymerase recognizes the cap-1 structure at the RNA 5'-end, measures a distance of 10–11 nucleotides from the cap structure, and cleaves next to purine bases or prior to either A or U residues. Transcription of viral genome RNAs is initiated using the resulting capped oligonucleotides as primers. This unusual "cap snatching mechanism" of transcription is employed by other enveloped segmented minus-strand RNA viruses and ambisense RNA viruses such as bunyaviruses and arenaviruses. Besides its priming function for viral transcription, the cap-1 structure acts as an allosteric effector for activation of virion-associated RNA polymerase.

Virion-associated RNA polymerase catalyzes only mRNA synthesis [58]. Isolated RNP cores with RNA-synthesizing activity contain four virus-specific proteins: NP, which represents over 95% of the total core proteins, and three P proteins, i.e., two basic, PB1 and PB2, and one acidic PA protein. The cleavage of capped RNA, primer-dependent initiation of RNA synthesis, elongation of RNA chains, and termination and polyadenylation of the transcripts are all carried out by these complexes. Moreover, the influenza virus RNA polymerase is capable of replacing bases at the growing ends of nascent RNA molecules [36]. Such apparent proof-reading function has never been observed for any RNA polymerases.

Centrifugation of RNP cores in CsCl leads to dissociation of NP, leaving the P proteins bound to vRNA [33]. The isolated RNA-P protein complexes devoid of NP are active in the synthesis of short RNA chains, but further elongation of RNA chains requires the presence of NP [32]. The dissociation of P proteins can be achieved by centrifugation in CsTFA. The solubilized RNA polymerase is composed of each one of the three P proteins [33] and exhibits RNA synthesis activity dependent on exogenously added templates [62]. Clear evidence for the requirement of all three P proteins for enzymatic activity was given by enzyme reconstitution experiments using individual P proteins, which were isolated from virions by SDS-gel electrophoresis followed by thioredoxin renaturation [81] or purified from insect cells expressing cDNA for each P protein carried on baculovirus vectors [43].

Chemical cross-linking experiments have been carried out to identify the functional sites. For example, pyridoxal 5'-phosphate (PLP) was cross-linked to PB1, indicating that PB1 has a nucleotide-binding site [72]. Genetic analyses indicate that PB1 plays a major role in RNA synthesis, while PB2 is involved in endonucleolytic cleavage of capped RNA. Mutations in the RNA polymerase genes affect virus growth markedly. Hence, the growth attenuation of cold-

adapted influenza viruses carry mutations in one of the P protein genes. Likewise, the attenuation of avian influenza viruses in humans is due to the reassortment of one of the genes encoding either RNA polymerase subunits or NP.

RNA polymerases of minus-strand plant RNA viruses

Virion-associated RNA-dependent RNA polymerase activity has been found in plant rhabdoviruses. In the case of wheat rosette stunt virus (WRSV) [80], both detergent-treated virions and isolated nucleocapsids exhibit RNA polymerase activity. Like animal rhabdoviruses, the enzyme activity can be regained upon mixing of L and NS proteins and using N-associated RNA template.

Products synthesized in vitro by the virion-associated RNA polymerase of plant rhabdoviruses contain genome-length and single-strand cRNA, indicating that the RNA polymerase acts not only as transcriptase but also as replicase.

RNA polymerases of ambisense RNA viruses

Sequence analysis of RNA virus genomes revealed that a group of viruses previously recognized as minus-strand viruses form a different class, ambisense viruses, including viruses belonging animal bunyaviruses and arenaviruses, and plant bunyaviruses and tenuiviruses. The strategy of gene expression of ambisense viruses is completely different from that of minus-strand viruses as both genome (vRNA) and anti-genome (cRNA) code for specific viral proteins.

RNA polymerases of ambisense animal RNA viruses

Viruses belonging to the *Bunyaviridae*, the largest family of animal RNA viruses, contain three RNA segments, L, M and S, in their genome. As also noted for the negative-strand viruses, the 3'- and 5'-terminal sequences of the three RNA segments show sequence homology of between 11 to 13 nucleotides, and are complementary to each other, allowing the formation of panhandle structures. RNA polymerase is found associated with virus particles. As in the case of influenza virus transcription, the bunyaviruses employ a cap-snatch mechanism for transcription initiation. These observations altogether indicated that bunyaviruses could be at least negative-strand viruses.

Sequence analysis of the S segment of Punta Toro virus, however, indicated that the 5'-proximal half of S-vRNA codes for a non-structural protein NS_S, although the 5'-proximal half of S-cRNA (antigenome RNA) encodes N protein. This type of RNA was proposed to be designated as "ambisense RNA", i.e., both vRNA and cRNA carry coding sequences (reviewed in [37]). Such an ambisense coding strategy was later identified in Pichinde virus, a member of the *Arenaviridae*. The virus-associated RNA polymerase is responsible for transcription of vRNA to allow expression of cRNA-encoded NS_S at early stages of infection. Like the influenza virus RNA polymerase, the RNA polymerase associated with La Crosse virus, a bunyavirus, carries the activities not only of primer-dependent RNA synthesis but also of capped RNA endonuclease. The RNA polymerase has been isolated from none of animal bunyaviruses.

RNA polymerases of ambisense plant RNA viruses

The plant bunyavirus, tomato spotted wilt virus (TSWV), and tenuiviruses, such as rice stripe virus (RSV) and rice grassy stunt virus (RGSV), carry single-strand ambisense RNAs as their genomes (Fig. 1). In the case of RSV, for example, at least three out of four RNA segments appear to use an ambisense coding strategy [82]. Expression of the putative open reading frames on both vRNA and cRNA strands has been demonstrated using in vitro translation systems [29].

As with animal ambisense viruses, an RNA polymerase activity is associated with purified virions of tenuiviruses. The only RNA polymerase gene so far sequenced is that of TSWV [20]. Barbier et al. [6] isolated the RNA polymerase activity by CsCl centrifugation of purified RSV virions. The active fraction contained two viral structural proteins, a 30 kDa nucleocapsid protein and a 230 kDa putative polymerase protein. An in vitro RNA synthesis system was reconstituted using this RNA-free protein fraction and short model templates carrying the conserved 5' and 3' terminal sequences. This showed that, as in the case of influenza virus, a minimum promoter function resides in the panhandle secondary structure formed by the complementary termini or in the 3' terminal sequence of 11–14 nucleotides in length.

The mechanism and regulation of transcription/replication of ambisense viruses are poorly known. Several lines of evidence indicate that, for gene expression of an ambisense RNA, subgenomic mRNA is formed from each RNA strand; and full-length vRNA and genome-size cRNA are synthesized only during replication. During virus maturation, cRNA is also assembled, albeit with low efficiency into virus particles, which after extraction generates double-strand forms by hybridizing with vRNA.

Molecular architecture of RNA-dependent RNA polymerases

Amino acid sequence comparisons of viral RNA polymerases have highlighted conserved domains corresponding to the following three sets of functions: RNA polymerase functions including RNA polymerization, nucleoside triphosphate (NTP) binding, and template and product binding [13, 42, 45, 64], RNA helicase functions for unwinding intramolecular secondary structures or template-product intermolecular duplexes [26, 31]; and RNA capping functions including methyltransferase [74]. These domains are usually composed of several blocks of conserved structural motifs.

The canonical GDD tripeptide originally proposed by Kamer and Argos [42] in the domain for RNA polymerization is now considered to be the hallmark of the viral RNA-dependent RNA polymerases. In many cases, the GDD motif is preceded by Tyr and is located within a hydrophobic amino acid stretch. Sequences surrounding the (Y)GDD motif show some similarity among each group of viruses sharing the same replication strategy. As might be expected from the difference in replication strategies, RNA polymerases from plus- and minus-strand viruses differ in their sequences even in the most conserved GDD motif; the consensus sequence is SDD for segmented minus-strand viruses, and GDN(Q) for unsegmented minus-strand viruses [64]. In contrast, double-strand

viruses carry conserved sequences rather similar to those of plus-strand viruses [13, 45]. In the case of RNA-dependent DNA polymerases (or reverse transcriptases), this motif is modified to (Y)MDD.

The RNA helicase and RNA methyltransferase domains of plus-strand viruses are not found in all virus families [28]. For instance, the helicase domain including the GKT/S motif is absent from the small genomes of plant carmo- and sobemoviruses; while the methyltransferase domain is a characteristic N-terminal component of the alphavirus RNA polymerase family, but is not required for RNA viruses which employ the cap-snatch mechanism for transcription initiation.

Poch et al. [65] showed that the RNA polymerase domain is composed of 4 major motifs. Motif A, DxxxxxD, is an acidic motif [corresponding to motif 1 of Bruenn [13] and motif VI of Koonin [45]]; motif B, GxxxTxxx(N/E)(S/T), is the core motif for nucleotide binding [motif 2 of Bruenn and motif V of Koonin]; motif C, (Y)GDD, is the core motif for catalytic function [motif 3 of Bruenn and motif VI of Koonin]; and motif D, LKR, is a basic motif [motif 6 of Bruenn and motif VII of Koonin]. These motifs are often spaced regularly with more variable hinges, altogether forming a chain of "concatenated motifs" [65].

In the crystal structure of the human immunodeficiency virus type I (HIV-I) reverse transcriptase at 3.5A resolution, the conserved (Y)MDD motif exists on a β-hairpin in the enzyme active center [44]. Inokuchi et al. (unpubl.) postulated that this motif is involved in metal ion binding at the catalytic site of the enzyme. As demonstrated for phage Q_β and poliovirus RNA polymerases, mutations in this motif result in significant loss of enzymatic activity [35, 38, 39]. However, some viral RNA polymerases carry different amino acid residues at one or two positions within this motif sequence: the Tyr residue in (Y)GDD motif is strictly conserved in picornavirus-like RNA polymerases, but is variable in alphavirus-like RNA polymerases. Substitution of Met for the Tyr residue in the poliovirus (Y)GDD motif results in a loss-of-function mutation, which can be suppressed by an intramolecular secondary mutation [38]. In this position, Cys is found in black beetle virus (BBV) and TMV; Ile is found in SND and Middleburg virus (MDV); Met is found in southern bean mosaic virus (SBMV); and Ser is found in yellow fever virus (YFV), BMV and AMV [4, 65]. In these cases, amino acid changes in the motif may be suppressed by secondary mutations elsewhere within the same molecule.

Two well-conserved sequence motifs have been found to be associated with purine nucleotide triphosphate (NTP)-binding activity [25, 79], i.e., the first motif A (G/AXXXXXGKS/T) and the second motif B (DEAD) occurring approximately 20–40 amino acids downstream (3′) from site A. The latter site is believed to interact with the Mg^{2+} cation of the Mg-NTP complexes for RNA- or DNA-dependent NTPase activity. Two superfamilies of DNA helicases involved in replication, recombination and DNA repair have been described with consensus features similar to those described here for NTP-binding [26, 31].

Evolution of viral RNA-dependent RNA polymerases

Sequence comparison of viral RNA polymerases is an approach suitable for defining the phylogenetic relationships between RNA viruses, because all viruses carry the RNA polymerase gene. Rates of evolution of cellular genes average 10^{-9} substitutions per site per year. On the other hand, RNA viruses evolve at rates a millionfold higher than their hosts with DNA genomes [70]. Quantitative estimation indicates that the error frequency of poliovirus (POL) and vesicular stomatitis (VSV) RNA polymerase ranges from 10^{-3} to 10^{-4} both in vitro and in vivo. The rapid fixation of mutations in RNA viruses implies, for instance, that the similarity observed between viruses which infect insect vectors and either animal or plant hosts indicates that their divergence took place recently. It can be best explained if all of the viruses in these groups originally infected insect cells and subsequently evolved to infect either plants or animals. Supporting this prediction, plant arboviruses such as plant bunya- and tenuiviruses share regions of sequence similarity with insect-transmitted animal bunyaviruses, arguing for a common origin [20, 82].

Both the gene organization within viral RNA genomes and the sequence organization within single viral genes sometimes display mosaic structures. The fact that different phylogenies are obtained from the amino acid sequences of various viral genes indicates that virus evolution also proceeds by recombination of modules [21]. Recombination by means of RNA polymerase jumping or template switching during replication has been well documented by the RNA sequence analysis of DI RNAs of several viruses, including polioviruses and influenza viruses.

Perspectives

The viral RNA-dependent RNA polymerases are extraordinarily complex, acting as not only RNA replicase but also transcriptase, and catalyzing not only RNA polymerization but also RNA modifications. Sequence analysis of the genes coding for the RNA polymerases from various viruses indicate mosaic structures consisting of multiple functional domains and motifs. The domain organization of RNA polymerase proteins seems to correlate with the strategy of gene expression and replication. Research is being focussed towards two directions: mapping of the functional domains involved in each reaction; and searching for the molecular mechanisms of interconversion between replicase and transcriptase.

Acknowledgements

References

1. Accotto GP, Marzachi C, Black LM, Banerjee AK (1990) Molecular characterization of double-strand RNA by carnation cryptic virus-associated RNA-dependent RNA polymerase. J Gen Virol 71: 433–437
2. Ahlquist P (1992) Bromovirus RNA replication and transcription. Curr Opin Genet Dev 2: 71–76
3. Ahlquist P, Strauss EG, Rice CM, Strauss JH, Zimmern D (1985) Sindbis virus proteins 1a and 1b contain homology to nonstructural proteins from several plant viruses. J Virol 53: 536–542
4. Argos P (1988) A sequence motif in many polymerases. Nucleic Acids Res 16: 9909–9916
5. Baker SC, Shieh C-K, Chang M-F, Vannier DM, Lai MMC (1989) Identification of a domain required for autocatalytic cleavage of murine coronavirus gene A polyproteins. J Virol 63: 3693–3699
6. Barbier P, Takahashi M, Toriyama S, Ishihama A (1992) solubilization and promoter analysis of RNA polymerase from rice stripe virus. J Virol 66: 6171–6174
7. Biebricher CK, Luce R (1992) In vitro recombination and terminal elongation of RNA by Q_β replicase. EMBO J 11: 5129–5135
8. Blumenthal T, Carmichael GG (1979) RNA replication: Function and structure of Qβ replicase. Annu Rev Biochem 48: 525–548
9. Boccardo G, Accotto GP (1988) RNA-dependent RNA polymerase activity in two morphologically different white clover cryptic viruses. Virology 165: 115–121
10. Boursnell MEG, Brown TD, Foulds IJ, Green PH, Tomley FM, Binns MM (1987) Compilation of the sequence of the genome of the coronavirus avian infectious bronchitis virus. J Gen Virol 68: 57–77
11. Bredenbeck PJ, Pachuk CJ, Noten AF, Carit'e J, Lsythes W, Weiss SR, Spaan WJ (1990) The primary structure and expression of the second open reading frame of the polymerase gene of the coronavirus MHV-A59: a highly conserved polymerase is expressed by an efficient ribosomal frameshifting mechanism. Nucleic Acids Res 18: 1825–1832
12. Brierly I, Boursnell ME, Binns MM, Bilimoria B, Blok VC, Brown TD, Inglis SC (1987) An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. EMBO J 6: 3779–3785
13. Bruenn JA (1991) Relationships among the positive-strand and double-strand RNA viruses as viewed through their RNA-dependent RNA polymerases. Nucleic Acids Res 19: 217–226
14. Burns CC, Lawson MA, Semler B, Ehrenfeld E (1989) Effects of mutations in poliovirus $3D^{pol}$ on RNA polymerase activity and on polyprotein cleavage. J Virol 63: 4866–4874
15. Cohen J, Chapilienne A, Chilmonczyk S, Ester MK (1989) Nucleotide sequence of bovine rotavirus gene 1 and expression of the gene product in baculovirus. Viorology 171: 131–140
16. Cole PE, Sawchyn I, Guerrier-Takada C (1982) Qβ replicase containing altered forms of ribosomal protein S1. J Biol Chem 257: 12929–12934
17. Cross RK (1983) Identification of a unique guanine 7-methyltransferase in Semliki Forest virus (SFV) infected cell extracts. Virology 130: 452–463
18. David C, Gargouri-Bouzid R, Haenni AL (1992) RNA replication of plant viruses containing an RNA genome. Prog Nucleic Acid Res Mol Biol 42: 157–227
19. De Groot RJ, Rumenaph T, Kuhn RJ, Strauss EG, Strauss JH (1991) Sindbis virus RNA polymerase is degraded by the N-end rule pathway. Proc Natl Acad Sci USA 88: 8967–8971
20. De Hann P, Kormelink R, de Oliveira Resende R, van Poelwijk F, Peters D, Goldnach

R (1989) Tomato spotted wilt virus L RNA encodes a putative RNA polymerase. J Gen Virol 72: 2207–2216

21. Dolja VV, Carrington JC (1992) Evolution of positive-strand RNA viruses. Semin Virol 3: 315–326

22. Dorssers L, van der Krol S, van der Meer J, van Kammen A, Zabel P (1984) Purification of cowpea mosaic virus RNA replication complex: Identification of a virus-encoded 110,000-dalton polypeptide responsible for RNA chain elongation. Proc Natl Acad Sci USA 81: 1951–1955

23. Dunigan DD, Zaitlin M (1990) Capping of tobacco mosaic virus RNA: Analysis of viral-coded guanyltransferase-like activity. J Biol Chem 265: 7779–7786

24. Evans RK, Haley BE, Roth DA (1985) Photoaffinity labeling of a viral induced protein from tobacco. J Biol Chem 260: 7800–7804

25. Gentry GA (1985) Locating a nucleotide binding site in the thymidine kinase of vaccinia virus and of herpes simplex virus by scoring triply aligned protein sequences. Proc Natl Acad Sci USA 82: 6815–6819

26. Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM (1989) Two related super-families of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. Nucleic Acids Res 17: 4713–4730

27. Gottliev P, Strassman J, Qiao X, Frilander M, Frucht A, Mindich L (1992) In vitro packaging and replication of individual genomic segments of bacteriophage φ6 RNA. J Virol 66: 2611–2616

28. Habili N, Symons RH (1989) Evolutionary relationships between luteoviruses and other plant RNA viruses. Nucleic Acids Res 17: 4543–4555

29. Hamamatsu C, Toriyama S, Ishihama A (1993) Ambisense coding strategy of rice stripe virus: In vitro translation studies. J Gen Virol 74: 1125–1131

30. Hayes RJ, Buck KW (1990) Complete replication of a eukaryotic virus RNA in vitro by a purified RNA-dependent RNA polymerase. Cell 63: 363–368

31. Hodgman TC (1988) A new superfamily of replicase proteins. Nature 333: 22–23

32. Honda A, Ueda K, Nagata K, Ishihama A (1988) RNA polymerase of influenza virus: Role of NP in RNA chain elongation. J Biochem 104: 1021–1026

33. Honda A, Mukaigawa J, Yokoiyama A, Kato A, Ueda S, Nagata K, Krystal M, Nayak DP, Ishihama A (1990) Purification and molecular structure of RNA polymerase from influenza virus A/PR8. J Biochem 107: 624–628

34. Hutchinson KL, Bouknight DP, Fan WM, Hunt DM (1990) Revertants of a mutant of vesicular stomatitis virus which has an aberrant polyadenylation activity and a temperature-sensitive transcription. Virology 194: 444–449

35. Inokuchi K, Hirashima A (1987) Interference with viral infection by defective RNA replicase. J Virol 61: 3946–3949

36. Ishihama A, Mizumoto K, Kawakami K, Kato A, Honda A (1986) Proofreading function associated with the RNA-dependent RNA polymerase from influenza virus. J Biol Chem 261: 10417–10421

37. Ishihama A, Nagata K (1988) Viral RNA polymerases. CRC Crit Rev Biochem 23: 27–76

38. Jablonski SA, Morrow CD (1993) Enzymatic activity of poliovirus RNA polymerases with mutations at the tyrosine residue of the conserved YGDD motif: Isolation and characterization of poliovirus containing RNA polymerases with FGDD and MGDD sequences. J Virol 67: 373–381

39. Jablonski SA, Luo M, Morrow CD (1991) Enzymatic activity of poliovirus RNA polymerase mutants with single amino acid changes in the conserved YGDD amino acid motif. J Virol 65: 4565–4572

40. Janda M, Ahlquist P (1993) RNA-dependent replication, transcription, and persistence of brome mosaic virus RNA replication in *S. Cerevisiae. Cell* 72: 961–970

41. Kajitani M, Ishihama A (1991) Identification and sequence determination of the host factor gene for bacteriophage Qβ. Nucleic Acids Res 19: 1063–1066

42. Kamer G, Argos P (1984) Primary structural comparison of RNA-dependent RNA polymerases from plant, animal and bacterial viruses. Nucleic Acids Res 12: 7269–7282

43. Kobayashi M, Tuchiya K, Nagata K, Ishihama A (1992) Reconstitution of influenza virus RNA polymerase from three subunits expressed using recombinant baculovirus system. Virus Res 22: 235–245

44. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA (1992) Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexes with an inhibitor. Science 256: 1783–1790

45. Koonin EV (1991) The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. J Gen Virol 72: 2197–2206

46. Koonin EV (1992) Evolution of double-stranded RNA viruses. Semin Virol 3: 327–340

47. Lee SY (1987) Characterization of RNA polymerase associated with rice ragged stunt virus. Intervirology 27: 189–195

48. Mansel EA, Patton JT (1990) Rotavirus RNA replication: VP2, but not VP6, is necessary for viral replicase activity. J Virol 64: 4988–4996

49. Marsh LE, Pogue GP, Huntley CC, Hall TC (1980) Insight to replication strategies and evolution of (+) strand RNA viruses provided by brome mosaic virus. Oxf Surv Plant Biol Cell 7: 297–334

50. Marzachi C, Accotto GP, d'Aquilio M, Caciagli P, Boccardo G (1990) In vitro transcription of the double-strand RNA genome of maize rough dwarf virus (Reoviridae). J Gen Virol 71: 707–711

51. Marzachi C, Milne RG, Boccardo G (1988) In vitro synthesis of double-strand RNA by carnation cryptic virus-associated RNA-dependent RNA polymerase. Virology 165: 115–121

52. Meshi T, Motoyoshi F, Adachi A, Watanabe Y, Takamatsu N, Okada Y (1988) Two concomittant base substitutions in the putative replicase gene of TMV confer the ability to overcome the effects of a tomato resistant gene, Tm-1. EMBO J 7: 1575–1581

53. Miller WA, Hall TC (1984) RNA-dependent RNA polymerase isolated from cowpea chlorotic mottle virus-infected cowpeas is specific for bromoviral RNA. Virology 132: 53–60

54. Mori M, Mise K, Okuno T, Furusawa I (1992) Expression of brome mosaic virus-encoded replicase genes in transgenic tobacco plants. J Gen Virol 73: 169–172

55. Morrow CD, Warren B, Lentz MR (1987) Expression of enzymatically active poliovirus RNA-dependent RNA polymerase in Escherichia coli. Proc Natl Acad Sci USA 84: 6050–6054

56. Mouches C, Candresse T, Bove JM (1984) Turnip yellow mosaic virus RNA-replicase contains host and virus-encoded subunits. Virology 134: 78–90

57. Munishkin AV, Voronin LA, Chetverin AB (1988) An in vivo recombinant RNA capable of autocatalytic synthesis by Qβ replicase. Nature 333: 473–475

58. Nagata K, Takeuchi K, Ishihama A (1989) In vitro synthesis of influenza viral RNA: Biochemical complementation assay of factors for influenza virus replication. J Biochem 106: 205–208

59. Olkkonen VM, Ojala PM, Bamford DH (1991) Generation of infectious nucleocapsids by in vitro assembly of the shell protein on to the polymerase complex of the dsRNA bacteriophage Φ6. J Mol Biol 218: 569–581

60. Ongradi J, Cunningham C, Szilagyi JF (1985) Temperature sensitivity of the transcriptase of mutants tsB1 and tsF1 of vesicular stomatitis virus New Jersey is a consequence of mutation affecting polypeptide L. J Gen Virol 66: 1507–1513

61. Palasingam P, Shaklee PN (1992) Reversion of Qβ RNA phage mutants by homologous RNA recombination. J Virol 66: 2435–2442
62. Parvin JD, Palese P, Honda A, Ishihama A, Krystal M (1989) Promoter analysis of influenza virus RNA polymerase. J Virol 63: 5142–5152
63. Plotch SJ, Palant O, Gluzman Y (1989) Purification and properties of poliovirus RNA polymerase expressed in *Escherichia coli*. J Virol 63: 216–225
64. Poch O, Blumberg BM, Bougueleret L, Tordo N (1990) Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. J Gen Virol 74: 1153–1162
65. Poch O, Sauvaget I, Delarue M, Tordo N (1989) Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. EMBO J 8: 3867–3874
66. Quadt R, Jaspars EM (1989) RNA polymerases of plus-strand RNA viruses of plants. Mol Plant Microb Int 2: 219–223
67. Quadt R, Jaspars EM (1990) Purification and characterization of brome mosaic virus RNA-dependent RNA polymerase. Virology 178: 189–194
68. Quadt R, Kao CC, Browning KS, Hershberger RP, Ahlquist P (1993) Characterization of a host protein associated with brome mosaic virus RNA-dependent RNA polymerase. Proc Natl Acad Sci USA 90: 1498–1502
69. Quadt R, Rosdorff HJM, Hunt TW, Jaspars EMJ (1991) Analysis of the protein composition of alfalfa mosaic virus RNA-dependent RNA polymerase. Virology 182: 309–315
70. Ramirez BC, Barbier P, Scron K, Haenni AL, Bernardi F (1993) Molecular mechanisms of point mutations in RNA viruses. In: Gibbs A, Callister CH (eds) Molecular basis of viral evolution. Cambridge University Press, Cambridge
71. Rhodes DP, Reddy DVR, MacLeod R, Black LM, Banerjee AK (1977) In vitro synthesis of RNA containing 5'-terminal structure $m^7G(5')ppp(5')Ap^m$ by purified wound tumor virus. Virology 76: 554–559
72. Romanos MA, Hay AJ (1984) Identification of the influenza virus transcriptase by affinity-labeling with pyridoxal 5'-phosphate. Virology 132: 110–117
73. Rothstein MA, Richards OC, Amin C, Ehrenfeld E (1988) Enzymatic activity of poliovirus RNA polymerase synthesized in *Escherichia coli* from viral cDNA. Virology 164: 301–308
74. Rozanov M, Koonin EV, Gorbalenya AE (1992) Conservation of the putative methyltransferase domain: a hallmark of the Sindbis-like supergroup of positive-strand RNA viruses. J Gen Virol 73: 2129–2134
75. Sanchez AS, De BP, Banerjee AK (1985) In vitro phosphorylation of NS protein by the L protein of vesicular stomatitis virus. J Gen Virol 66: 1025–1032
76. Schubert M, Harmison GG, Richardson CD, Meier E (1985) Expression of a cDNA encoding a functional 241-kilodalton vesicular stomatitis virus RNA polymerase. Proc Natl Acad Sci USA 82: 7984–7988
77. Snijder EJ, Boon JA, Bredenbeek PJ, Horzinek M, Rijanbrand R, Spaan WJM (1990) The carboxy-terminal part of the putative Berne virus polymerase is expressed by ribosomal frameshifting and contains sequence motifs which indicate that toro- and coronaviruses are evolutionarily related. Nucleic Acids Res 18: 4535–4542
78. Spies U, Muller H (1990) Demonstration of enzyme activities required for cap formation in infectious bursal disease virus, a member of the birnavirus group. J Gen Virol 71: 977–981
79. Strauss JH, Strauss EG (1988) Evolution of RNA viruses. Annu Rev Microbiol 42: 657–683
80. Sun W, Gong ZX (1988) In vitro studies on the nucleocapsid associated with RNA polymerase of wheat rosette stunt virus. Intervirology 29: 154–161

81. Szewczyk B, Laver WG, Summes DF (1988) Purification, thioredoxin renaturation, and reconstituted activity of the three subunits of the influenza A virus RNA polymerase. Proc Natl Acad Sci USA 85: 7907–7911

82. Takahashi M, Toriyama S, Hamamatsu C, Ishihama A (1991) Nucleotide sequence and possible ambisense coding strategy of rice stripe virus RNA segment 2. J Gen Virol 74: 769–773

83. Van Bokhoven H, Mulders M, Wellink J, Vlak JM, Goldbach R, van Kammen A (1991) Evidence for dissimilar properties of comoviral and picornaviral RNA polymerases. J Gen Virol 72: 567–572

84. Van Dun CMP, van Vloten-Doting L, Bol JF (1991) Expression of alfalfa mosaic virus cDNA1 and 2 in transgenic tabacco plants. Virology 163: 572–578

85. Van Kammen A, Eggen HIL (1987) The replication of cowpea mosaic virus. BioEssays 5: 261–265

86. Vidal S, Kolakofsky D (1989) Modified model for the switch for Sendai virus transcription to replication. J Virol 63: 1951–1958

87. Wilusz J, Kurilla MG, Keene JD (1983) A host protein (La) binds to a unique species of minus-sense leader RNA during replication of vesicular stomatitis virus. Proc Natl Acad Sci USA 80: 5827–5832

88. Witt DJ, Naeve CW, Summers DF (1981) Phosphorylation of vesicular stomatitis virus proteins as a possible contributing factor in virion uncoating. J Gen Virol 56: 383–391

89. Yamashita T, Kawai A (1990) Vesicular stomatitis virion-associated transcriptase activity was suppressed in vitro by a synthetic 21 amino acid oligopeptide prepared to mimic the carboxy-terminus of NS protein. Virology 178: 166–173

90. Young N, Fomey J, Zaitlin M (1987) Tobacco mosaic virus replicase and replicative structures. J Cell Sci [Suppl] 7: 277–285

Authors' address: Dr. A. Ishihama, National Institute of Genetics, Mishima, Shizuoka 411, Japan.