The EMBO Journal vol.13 no.13 pp.3158–3165, 1994

Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent

Florence Baudin, Christian Bach[1], Stephen Cusack and Rob W.H.Ruigrok[2]

EMBL Grenoble Outstation, c/o ILL, BP 156, F-38042 Grenoble Cedex 9, France and [1]Institute of Virology, Justus Liebig Universität, Frankfurterstrasse 107, D-35392 Giessen, Germany
[2]Corresponding author

Communicated by C.Scholtissek

The influenza virus genome consists of eight segments of negative-sense RNA, i.e. the viral (v) RNA forms the template for the mRNA. Each segment is encapsidated by the viral nucleoprotein to form a ribonucleoprotein (RNP) particle and each RNP carries its own polymerase complex. We studied the interaction of purified nucleoprotein with RNA *in vitro*, by using a variety of enzymatic and chemical probes for RNA conformation. Our results suggest that the nucleoprotein binds to the vRNA backbone without apparent sequence specificity, exposing the bases to the outside and melting all secondary structure. In this way, the viral polymerase may transcribe the RNA without the need for dissociating the nucleoprotein and without being stopped by RNA secondary structure, and the viral RNPs are ready to start transcription as soon as they enter the host cell.

Key words: influenza nucleoprotein/replication/RNA–protein interaction/RNP/transcription

Introduction

Influenza viruses are spherical, enveloped viruses with a genome composed of eight separate RNA segments of negative polarity (Palese, 1979). In the virus, each RNA segment occurs as a separate ribonucleoprotein particle (RNP; Pons *et al.*, 1969; Compans *et al.*, 1972) in which the RNA is covered with nucleoprotein (NP, 56 kDa) and also has a polymerase complex attached (consisting of three subunits PA, PB1 and PB2). The RNPs are independent transcription active units. The same polymerase complex is responsible for both transcription and replication of viral (v) RNA (Krug, 1983; Lamb and Choppin, 1983; McCauley and Mahy, 1983). However, full-length RNA copies, in contrast to mRNA copies, are only made in the presence of soluble NP, implying that a proper arrangement of NP on the nascent RNA is needed for efficient elongation (Beaton and Krug, 1986; Shapiro and Krug, 1988). Furthermore, the polymerase acts efficiently only on intact RNP; removal of NP from the RNP results in abortive transcription (Honda *et al.*, 1988).

Influenza vRNAs have two particular features which may be important for the interaction with NP and polymerases. They show extensive conservation of the terminal sequences: the 12 nucleotides (nt) at the 3′ terminus show only a single base difference in half of the eight genome segments and the 13 nt at the 5′ terminal end are completely conserved (Skehel and Hay, 1978; Robertson, 1979; Desselberger *et al.*, 1980; Stoeckle *et al.*, 1987). A second aspect of the influenza vRNA segments is the inverted complementarity between the termini, extending beyond the conserved nucleotides. These termini may serve as regulatory elements for influenza virus RNA transcription and replication, and possibly also for the packaging of vRNAs into virions (Desselberger *et al.*, 1980; Hsu *et al.*, 1987). In the intact RNP complexes, the complementary termini can be cross-linked with psoralen and it has been suggested that they may be associated in a double-stranded structure, called a panhandle (Hsu *et al.*, 1987).

In this paper, we investigate the mode of binding of NP to RNA and in particular address firstly the question as to whether there is any sequence specificity in NP binding that might for instance indicate a site of nucleation of RNA coating. In this respect, we have tested whether either of the conserved terminal oligoribonucleotides have higher affinity for NP. Secondly, we have determined the change in RNA conformation that occurs upon NP binding by studying an artificial mini vRNA of 81 nt that contains the conserved 3′ and 5′ ends and the poly(U) stretch which occurs near the 5′ end of all eight vRNA fragments. We have determined the conformation of this 'panhandle' RNA in solution, both free and complexed with NP, by using different nucleases and chemical probes that give information on local RNA conformation or the accessibility of various atomic positions on the RNA (Ehresmann *et al.*, 1987). The data indicate that in the free RNA, the 3′ and 5′ ends are indeed base paired, forming a panhandle. However, upon association of NP with the RNA, apparently through backbone contacts, all secondary structure was disrupted and the Watson–Crick positions of the bases were exposed to the solvent. This suggests that one of the roles of NP is to present the RNA bases in an easily accessible way to the polymerase which can then transcribe without disrupting the RNP.

Results

Binding of isolated NP to RNA

The affinity of NP for RNA was determined using a nitrocellulose filter binding assay in which labelled RNA, at a constant concentration, was titrated with increasing concentrations of purified NP. The results are given in Figure 1 and Table I. Our preparation of NP contained ~1% contaminating viral polymerase proteins, but here and later we assume that all binding to, and effects on, RNA are caused by NP only.

We found no significant difference between the affinities for the conserved 3′ and 5′ ends of the vRNA. Binding

to oligonucleotides with a degenerate sequence from 16 to 20 nt showed no significant preference for length in the range tested and the K_d values were of the same order of magnitude as for binding to intact segment 8 vRNA (890 nt) and panhandle RNA (Figure 2). When the two oligoribonucleotides representing the 3′ and 5′ ends were incubated together in solution, they formed a double-stranded structure, as judged from gel retardation (not shown), although NP bound to the mix with an affinity intermediate to the individual affinities, as if NP had reacted with the separate single strands. The oligoribonucleotides corresponding to the 3′ and 5′ conserved ends do not form a very stable double-stranded structure due to a bulge and an inner loop in the duplex (see below for the structure of the double strand in the free panhandle RNA). We therefore designed two perfectly complementary oligoribonucleotides of 20 nt based on the conserved vRNA ends (see Table I). Binding to these single-stranded oligoribonucleotides resulted in values similar to those of the affinities for the real 3′ and 5′ ends, but the affinity for the double-stranded complex was considerably lower. We also measured binding of NP to yeast $tRNA^{Ser}$ which contains a high percentage of GC base pairs and represents a very compact and stable structure. NP was found to bind with an affinity similar to that for the 20 nt double-stranded oligoribonucleotide.

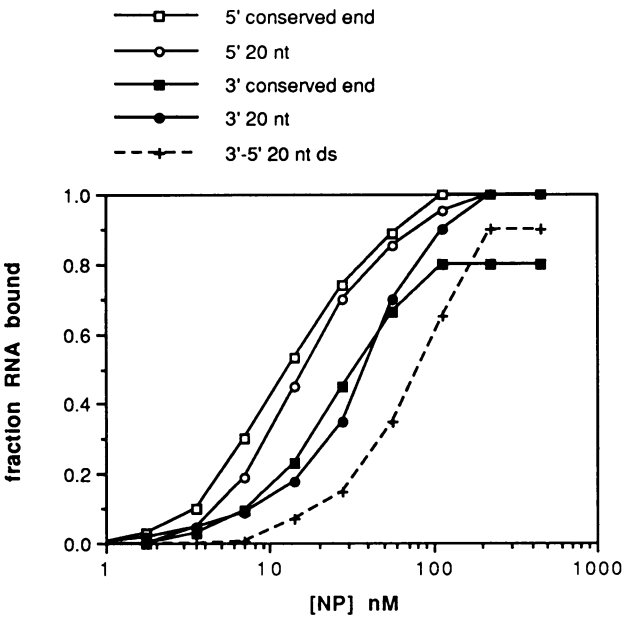

Fig. 1. Titration curves from filter binding assays of various oligoribonucleotides with NP. The mid-points of the curves represent the K_d values.

Table I. Affinity of purified influenza NP for RNA: relative binding strengths of various RNAs and oligoribonucleotides to influenza A NP

Type of RNA	Number of nt	K_d in 10^{-8} M
Segment 8 vRNA (–) strand	890	2.3 ± 0.4
Panhandle RNA	81	3.7 ± 0.2
5′ Conserved end of vRNA	16 ss	0.9 ± 0.2
3′ Conserved end of vRNA	15 ss	2.0 ± 0.6
3′ 5′ Conserved ends mixed	15 + 16 ds	1.7 ± 0.1
Degenerate sequence	16	3.2 ± 0.7
	17	2.2 ± 0.2
	18	2.6 ± 0.6
	19	3.8 ± 1.2
	20	3.2
5′ AGUAGAACAGGGUGACAAAG	20 ss	1.1 ± 0.4
3′ UCAUCUUGUCCCACUGUUUC	20 ss	3.0 ± 1.0
5′ AGUAGAACAGGGUGACAAAG 3′ UCAUCUUGUCCCACUGUUUC	20 ds	7.1 ± 0.1
$tRNA^{Ser}$ from yeast	88	9.0

The values are the averages of at least two experiments ± SD, except for the tRNA and 20 nucleotide degenerate oligo experiments which were single determinations.

Secondary structure of the naked panhandle RNA

Examples of autoradiograms of several enzymatic and chemical probing experiments on free panhandle RNA are

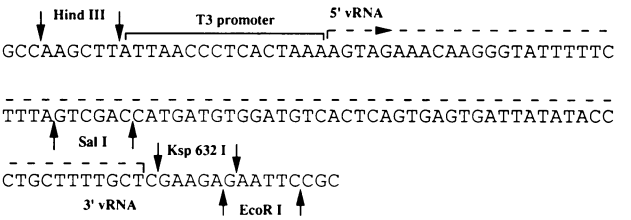

Fig. 2. Construction of the DNA template for the panhandle model RNA (vRNA sense). An oligodeoxyribonucleotide was synthesized containing the T3 promotor sequence, the first 26 nt of the 5′ end of segment 5 (NP) of fowl plague virus (strain FPV.GI, H7N1), a spacer of 6 nt (*Sal*I restriction site), followed by the last 48 nt of the 3′ end of the NP sequence, a spacer of 1 nucleotide (C) and a *Ksp*632I restriction site. The sequence was extended by an additional 9 nt at the 5′ end containing a *Hin*dIII restriction site and an additional 8 nt at the 3′ end containing an *Eco*RI restriction site.

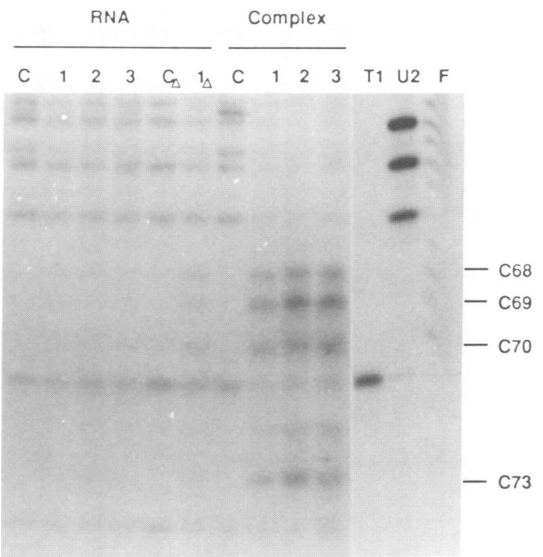

Fig. 3. PAGE autoradiogram of the DMS experiment on 3′-labelled RNA. The RNA was cut by treatment with hydrazine and aniline. The reactivities of the N-3 positions of cytosines 68–70 and 73 are shown on the naked panhandle RNA (RNA) and on the panhandle RNA–NP complex (complex). Lanes C show the incubation controls for the RNA and complex experiments, △ indicates the experiment under semi-denaturing conditions (absence of $MgCl_2$). Lanes 1, 2 and 3: incubation with DMS for 4, 7 and 10 min, respectively. T1 and U2 show cleavage products of RNase treatment of the naked RNA, showing the G and A residues, respectively. F shows the formamide ladder.

shown in Figures 3–7. The experimental data obtained by the probing experiments were used to select from a series of possible secondary structure models that which was most consistent with the results. The model can be divided into three domains (Figure 8A).

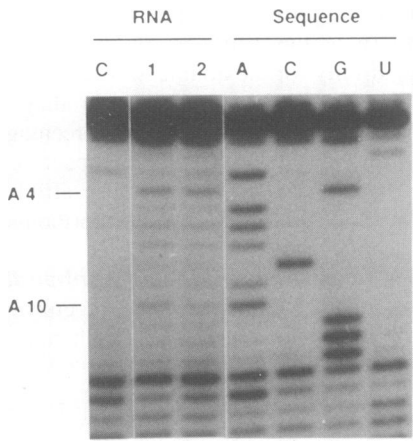

Fig. 4. PAGE autoradiogram of the DMS experiment on naked RNA; the reactive positions were revealed by the reverse transcriptase method. Lane C is the control, lanes 1 and 2 represent 4 and 7 min incubation, respectively, and A, C, G, U give the sequence. Note the reactivity of adenines 4 and 10, suggesting that they are not base paired, and the absence of reactivity for adenines 6–8 indicating that they are part of the helix.

Domain I. Domain I consists of a helix formed by base pairing of the conserved 3′ and 5′ ends, this conclusion being based on the following evidence. Figure 3 shows that the Watson–Crick N-3 positions of C68, C69 and C70 are not reactive towards dimethylsulphate (DMS), suggesting base pairing with G12, G13 and G14 which are also not reactive at their Watson–Crick positions. Both stretches are cut by RNase V1, also suggesting a double strand (Figure 8A). A16 is not reactive to diethylpyrocarbonate (DEPC) (Figure 5B) and A65 and 67 are only mildly reactive (Figure 5A). Owing to the size of the adduct, DEPC is sensitive to stacking and all adenines in helices are expected to be unreactive. The Watson–Crick positions of A6, A7 and A8 are not reactive towards DMS (Figure 4) as revealed by the reverse transcriptase technique. The reactivities for the counterparts of these adenines could not be identified by the same technique since positions 69–81 were used for primer hybridization. However, the fact that two RNase V1 cuts at positions 75 and 76 were revealed by the direct method, and that the N-7 positions of A6, A7 and A8 were not reactive to DEPC (Figure 5B), would suggest that both strands are involved in a helix.

Domain I contains an inner loop and a bulge. *Inner*

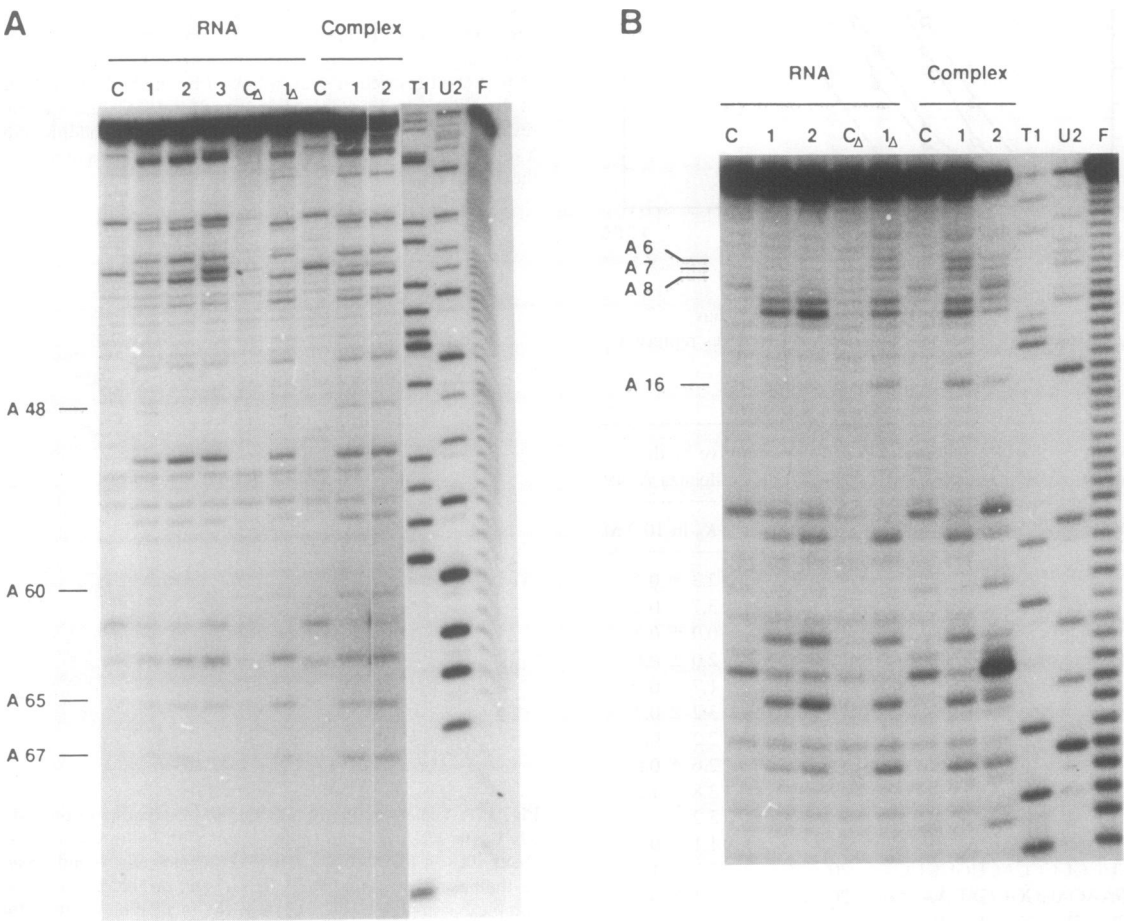

Fig. 5. PAGE autoradiogram of the DEPC experiment on 3′-labelled RNA, showing the reactivities of the N-7 positions of adenines in the naked panhandle RNA (RNA) and on the panhandle RNA–NP complex (complex). Note that A10 and 11 are reactive in the naked RNA, A6–8 and A16 are not reactive in the naked RNA, but are reactive in the complex. T1 and U2 show cleavage products of RNase treatment of the naked RNA, showing the guanine and adenine residues, respectively, and F shows the formamide ladder. △ indicates semi-denaturing conditions and all C lanes are the respective controls. Lanes 1 and 2: incubations for 10 and 20 min, respectively. (**A**) short migration; (**B**) long migration.

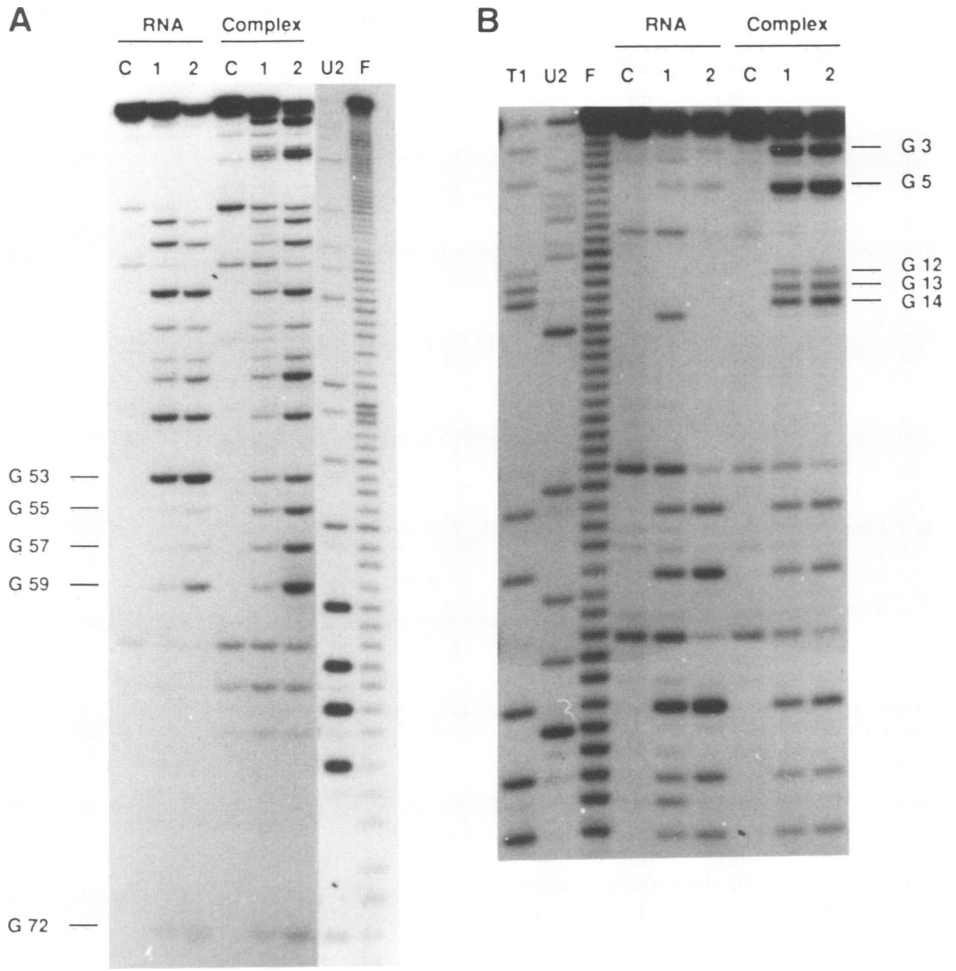

Fig. 6. PAGE autoradiogram of RNase T1-digested 3′ end-labelled panhandle RNA, free in solution (RNA) or in an RNA–NP complex (complex). RNase T1 is specific for single-stranded RNA, note the absence of cuts in the RNA lanes and their presence in the complex lanes, for guanines 2, 5, 12–14. Lanes C, T1, U2 and F, see the legend to Figure 5; lanes 1 and 2: 1.25×10^{-3} and 2.5×10^{-3} U of enzyme, respectively. (**A**) short migration; (**B**) long migration.

loop: A10 is reactive towards DMS (Figure 4) and DEPC (Figure 5B), and the fact that a cut by RNase T1 occurs at G72 (Figure 6A) suggests that the O-6 and N-1 atomic positions of the guanine are free (Heinemann and Saenger, 1982). C9 and C73 are not reactive at their Watson–Crick N-3 positions (Figures 3 and 4). It is possible that both cytosines interact together by non-canonical hydrogen bonds, which have been shown to exist in poly(C) (Cantor and Schimmel, 1980) and in anticodon–anticodon interaction in $tRNA^{Gly}$ (Romby *et al.*, 1986). All this suggests that C9, A10 and G72, C73 are forming an inner loop. *Bulge*: Figure 4 shows that A4 is reactive towards DMS and, therefore, cannot be base paired with U77. The fact the the N-7 position of the adenine is not reactive towards DEPC (Figure 5B) may suggest that the base is buried inside the helix. Models of the base-paired ends of the influenza vRNAs are usually drawn with base pairing between A4 and U77, leaving A8 bulging out [see for instance Hsu *et al.* (1987)]. Our data suggest that the stretch GAAA(5–8) interacts with UUUU (74–77), as shown by the non-reactivities of N-7 of the purines (Figure 5A) and of the Watson–Crick positions (Figure 8A). Our model is consistent with the fact that U77 of influenza A virus RNA segments is sometimes replaced by C in segments (Skehel and Hay, 1978; Robertson, 1979). In this case, C77 could base pair with G5, resulting in the same structure for the helix as proposed in Figure 8A.

Domain II. Nucleotides 45–61 fold into a stable stem–loop structure, with the N-7 of adenines 48 and 60 being unreactive towards DEPC (Figure 5A), and the Watson–Crick positions of the bases in the stem not reactive (Figure 8A). The Watson–Crick positions of G45 and U61, located at the terminal base pair, are only marginally reactive (Figure 8A). This behaviour reveals the dynamic properties of G-U pairs or A-U pairs near junctions, as already pointed out for other RNAs (e.g. Baudin *et al.*, 1987; Philippe *et al.*, 1990). The exterior loop is fully accessible to single-strand-specific nucleases (G53, Figure 6A). Nucleotides in loops are expected to be reactive at their Watson–Crick positions. This is true for 4/5 nt (51–54). The absence of reactivity of G55 at N-1 could suggest some intrinsic structure or long-range interactions, although the high accessibility of the loop to the enzymes would argue against long-range interactions.

Domain III. Domain III consists of nt 18–45. The overall reactivities of the Watson–Crick positions, together with the accessibility to single-strand-specific nucleases, sug-

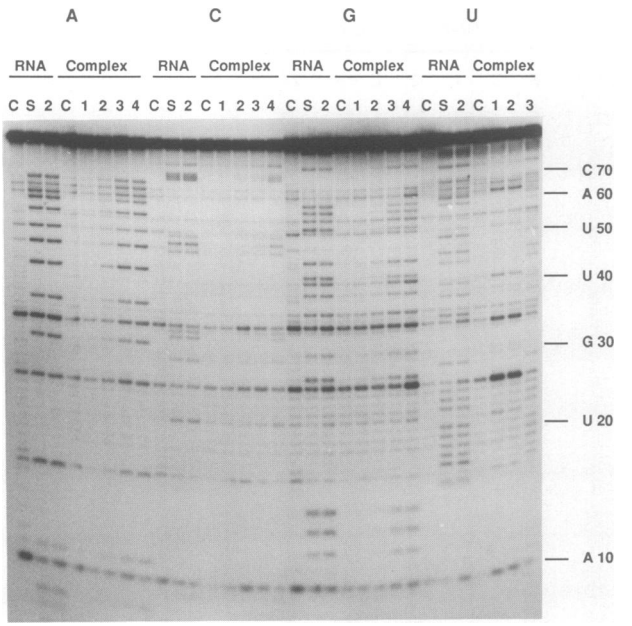

Fig. 7. PAGE autoradiogram of 5′ end-labelled RNA showing the reactivities of the phosphates of naked panhandle RNA and the RNA–NP complex. A, C, G and U stand for the four transcription reactions in which one of the NTPs was supplemented by 5% of the corresponding NTP[αS]. Lanes 1, 2, 3 and 4 represent iodine concentrations of 0.1, 0.3, 0.6 and 1 mM, respectively. C stands for control and S for sequence.

gests that this whole region is single stranded. Such a large loop would not be very stable thermodynamically, but the sequence does not show suitable base pairing possibilities. The fact that the N-7 position of the adenines is reactive towards DEPC (Figure 5A and B) confirms the idea that the nucleotides in this region are disorganized.

None of the phosphates in the naked RNA are protected (Figure 7), suggesting that there are no tertiary interactions between the domains nor protection by bound Mg^{2+} ions. Further, the N-7 positions of the guanines were all found to be reactive, also indicating the absence of tertiary interactions (not shown).

Conformation of the RNA inside the panhandle RNA–NP complex

We subsequently determined the structure of the panhandle RNA when complexed with nucleoprotein. For formation of the complex, we used the RNA–protein ratio that resulted in 100% retention in the filter binding curve, as shown in Figure 1. Specifically, 0.8 μg panhandle RNA was incubated with 8.6 μg NP (20 min at 37°C), corresponding to an input stoichiometry of 15 nt/NP. The results of the probing experiments on the complex are shown in Figure 8B.

In domain I, C68, C69, C70 as well as C73, that were not reactive in the naked RNA, have become fully reactive at their N-3 positions towards DMS in the complex (Figure 3). The counterparts of the cytosines in the helix, G12, G13 and G14, are strongly cut by RNase T1 in the presence of the protein (Figure 6B). The N-7 positions of all the adenines present in the panhandle structure have become reactive towards DEPC (Figure 5A and B). Since all Watson–Crick positions of the nucleotides in domain I have become reactive, these observations suggest that

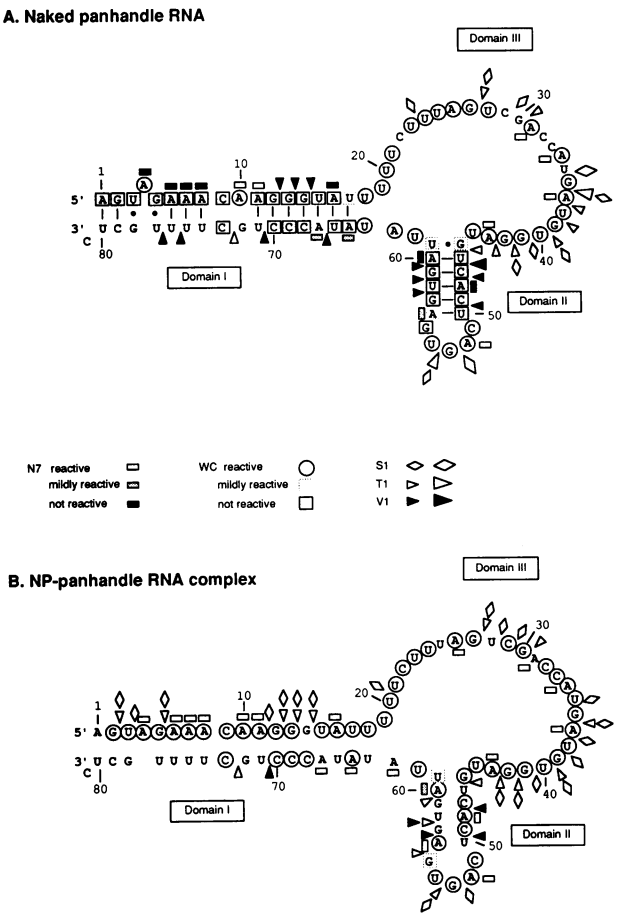

Fig. 8. Secondary structure model for the naked panhandle RNA (**A**) and for the RNA in the complex with NP (**B**). Circles and squares around bases indicate the reactivity of Watson–Crick positions of the bases as indicated; the absence of circles or squares means that the reactivity could not be determined. The cleavage by nucleases is indicated by arrowheads and diamonds as indicated, and the reactivity of the adenine N-7 positions by bars above or below the bases.

the 3′ and 5′ conserved ends are no longer base paired in the complex. The RNase V1 sensitivities that existed in domain I of the naked RNA are no longer seen in the complex (Figure 8B). We observed that all the guanines become accessible towards RNase T1 (which implies access to O6 and N1), even though steric hindrance due to the presence of bound NP could have disturbed the accessibility of the bases to the RNase (Figure 6A and B). The results with nuclease S1 gave the same pattern as with RNase T1. Similar results were found for domain II. The results of chemical probing showed that the Watson–Crick positions were accessible for modification, indicating that the panhandle RNA had become completely unfolded due to the fixation of the NP. However, some residual V1 reactivities remained, even if at the same site RNase T1 reactivity was observed, possibly due to heterogeneity caused by incomplete complex formation. A similar explanation could be given for some remaining RNase V1 reactivity at residue C70 in domain I. No big changes were observed in domain III since the reactivity of the nucleotides towards the different probes already existed. All the cuts by the single-strand-specific RNases remained at the same intensity (Figure 6A and B).

When the phosphate reactivities were tested, we observed a relative protection of all phosphates at a concentration of 300 μM iodine (Figure 7, compare lanes 2 for the naked RNA and the complex). This result suggests that the nucleoprotein interacts with the phosphate-sugar backbone of the RNA. Since we observe an overall protection of the phosphates, in contrast to a footprint of protected areas interspaced by non-protected regions, the NP binds either to specific sites on the RNA without free phosphates in between, or the NP binds randomly to no precise site on the RNA molecule. The fact that we see an appearance of phosphate reactivities at 1 mM iodine would favour the latter possibility.

Discussion

For the replication of influenza vRNA, during which full-length copies of template and vRNA are made, the presence of soluble NP is essential (Beaton and Krug, 1986; Shapiro and Krug, 1988). It is thought that NP binds to the nascent RNA and in this way prevents somehow the formation of the poly(A) tail by the 'stuttering' viral polymerase; when NP associates with the newly formed RNA chain, the polymerase produces a full-length copy without stuttering. Since the first piece of RNA to be synthesized is the 5′ end, and since the sequences of the first 13 nt of the 5′ ends of all RNA segments are conserved, a plausible candidate for the initiation of encapsidation would be the 5′ end sequence. Furthermore, the 5′ terminus is complementary to the sequence of the 3′ end, which makes the 5′ ends of viral and template RNAs homologous (template RNA is also encapsidated by NP in the infected cell and is an intermediate for replication).

Our results show that NP has a preference for single-stranded rather than for double-stranded structures, in agreement with previous studies (Scholtissek and Becht, 1971; Kingsbury *et al.*, 1987). However, although the affinities for the 5′ conserved sequence and for the 20 nt long sequence derived from the 5′ end seem to be the strongest, the difference with the affinities for the other single-stranded RNAs is hardly significant. The affinity of vesicular stomatitis virus (VSV) N protein for leader RNA, containing the VSV encapsidation signal, is ~10 times higher than that for non-specific sequences (Blumberg *et al.*, 1983).

Our results on the NP–panhandle RNA complex suggest that NP binds to the RNA phosphate-sugar backbone. This mode of binding is consistent with the fact that the protein has a charge of +14 at pH 6.5 (Winter and Fields, 1981). Our results also suggest that the RNA is on the outside of the RNP structure. This is in agreement with the fact that the RNA in RNP is sensitive to RNases without disrupting the RNP structures (Duesberg, 1969; Kingsbury and Webster, 1969; Pons *et al.*, 1969) and that the RNA in the RNP can be easily displaced by polyvinylsulphate (Pons *et al.*, 1969; Goldstein and Pons, 1970). Each NP monomer interacts with ~20 nt (Duesberg, 1969; Pons *et al.*, 1969; Kingsbury and Webster, 1969; Jennings *et al.*, 1983) and has the positive charges divided over the primary structure without clustering. We suggest that the RNA is coiled around the outside of NP, with the bases exposed, so that the polymerase can walk over the RNA without the need for dissociation of NP from RNA. Sequence determination of subgenomic RNAs led Jennings *et al.* (1983) to propose a similar model, in which the polymerase could jump from a 'down-going strand to an up-going strand', leading to a deletion. Honda *et al.* (1988) showed that polymerase–RNA complexes devoid of NP were active in transcription, but only produced short transcripts. Addition of NP resulted in restoration of the production of full-length copies, also suggesting that NP is needed to avoid stops in transcription, probably by melting out secondary structure. Although the removal of secondary structure in the panhandle RNA by NP is reminiscent of the action of helicases, NP is clearly not a helicase. None of the conserved sequence elements of helicases (Pause and Sonenberg, 1993) are present in NP and NP does not have the ATPase activity of helicases which permits them to bind to and dissociate from the RNA. NP is thought to bind to the nascent single-stranded RNA and remain bound.

Hsu *et al.* (1987) have suggested that the 3′ and 5′ ends of the RNA in the RNP are base paired, whereas those in the naked vRNA are not. They suggested that a protein component in the RNP probably stabilized the interaction between the two ends. Our results with the panhandle RNA show the opposite situation, with the ends in the naked RNA paired, but single-stranded in the complex. However, although the results of Hsu *et al.* (1987) clearly show that the two ends of the RNA in RNP can be cross-linked with a psoralen reagent, the experiments to show that the ends are actually base paired were not quite conclusive since only a very small percentage of the 3′ and 5′ probes they used was indicative for base pairing. We are aware that the work presented here does not necessarily reflect the situation in the native RNP. However, our approach has the advantage that we now know the effect of NP binding to an RNA molecule with several important features of true vRNA. Further, preliminary experiments on native RNP show similar results as presented here on the NP–panhandle RNA complex, with the bases available for modification and the phosphates protected in a general manner without well-defined footprints (F.Baudin *et al.*, unpublished results). Information on the structure at the 3′ and 5′ ends of the vRNA is obscured, however, by the presence of the polymerase complex.

Immediately after entering the host cell, negative-strand viruses depend on transcription of the vRNA in order to start the replication process. Our results suggest that the influenza RNPs in the virus are ready to go directly after infection. The situation is probably similar for other negative-strand viruses (rabies, measles, respiratory syncytial virus, etc.), but different for the positive-strand viruses which start with translation and for which rapid interaction with the ribosomes is essential.

Materials and methods

Chemicals and enzymes

1-Cyclohexyl-*N*′-[2-(*N*-methylmorpholino) ethyl-carbodiimide-*p*-toluenesulphonate (CMCT) was obtained from Merck; aniline, DMS, DEPC from Fluka; kethoxal from USB; RNasin and avian myeloblastosis virus reverse transcriptase from Appligène (France). [γ-^{32}P]ATP was from ICN (France). Restriction enzymes were from Biolabs (France), Boehringer (France), Stratagene (Heidelberg) or Eurogentec (Belgium). RNases T1,

U2 and V1, nuclease S1, nucleotides, T4 polynucleotide kinase and T4 RNA ligase were obtained from Pharmacia.

Virus, RNP and NP preparations

Influenza virus A/PR/8/34 was grown in embryonated hen's eggs and obtained in purified form from Pasteur-Mérieux, Marcy L'Etoile, France. Virus was treated with Triton X-100 (1%) and lysolecithin (1 mg/ml) in 5 mM $MgCl_2$, 100 mM KCl, 1.5 mM dithiothreitol (DTT), 5% glycerol and 10 mM Tris–HCl (pH 8) as described by Honda *et al.* (1987). The sample was incubated for 10 min at 30°C and centrifuged through a linear 10–30% glycerol gradient in 100 mM NaCl, 1 mM DTT and 50 mM Tris–HCl (pH 8) (SW27 rotor, 25 000 r.p.m., 16.5 h, 4°C). RNP-containing fractions were pooled, dialysed and concentrated in 100 mM NaCl, 1 mM DTT and 20 mM Tris–HCl (pH 8) at 4°C. The RNP suspension was then layered onto a CsCl/glycerol step gradient as described by Honda *et al.* (1988). NP fractions were collected and dialysed against 50 mM Tris–HCl (pH 7.5), 100 mM NaCl, 10 mM $MgCl_2$, 2 mM DTT and 30% glycerol. In our hands, the NP and polymerase proteins were found in the same gradient fractions, the RNA was at the bottom of the tube.

Panhandle RNA, NS vRNA and oligoribonucleotides

Plasmid pVN1 was constructed for the production of a short 'panhandle RNA' in vRNA-sense under the control of a T3 RNA polymerase promotor by *in vitro* transcription (see Figure 2). The resulting 120-mer was purified by polyacrylamide gel electrophoresis (PAGE). Another oligodeoxyribonucleotide of 17 bases, GCGGAATTCTCTTCGAG, complementary to the 3′ end of the 120-mer, was used as a primer to produce double-stranded DNA with Klenow enzyme (Stratagene, Heidelberg). The resulting double-stranded 120 bp DNA and pUC 19 DNA were digested with *Eco*RI and *Hin*dIII (Stratagene, Heidelberg), followed by ligation with T4 ligase.

In vitro transcription was carried out using T3 RNA polymerase (for the panhandle RNA) and T7 RNA polymerase (for intact segment 8 (NS) vRNA transcribed from pHgaNS, a gift from Dr Mark Krystal) as described in Baudin *et al.* (1993). Prior to transcription, each double-stranded pUC DNA containing the desired insert was digested with the restriction enzyme *Hga*I for the NS fragment and with *Ksp*632I for the panhandle vRNA. These restriction enzymes define the 3′ terminus of the transcripts as nt +890 and +81 of the gene, respectively.

A set of oligoribonucleotides was synthesized (EMBL, Heidelberg) representing the 3′ conserved end of influenza A NS vRNA [3′ UCGUUUUCGUCCCAC 15 nt] and the 5′ conserved end [5′ AGUAGAAACAAGGGUG 16 nt]. Two oligoribonucleotides of 20 bases were also synthesized, based on the conserved viral 3′ and 5′ sequences, but perfectly complementary, and were used to determine the affinity of NP for binding to single-stranded and double-stranded RNAs (see Table I). Oligonucleotides ranging from 16 to 20 ribonucleotides with a degenerate sequence were also synthesized.

RNA labelling

RNA was dephosphorylated at its 5′ end as described by Shinagawa and Padmanabhan (1979) and then labelled using [γ-^{32}P]ATP and T4 polynucleotide kinase according to Silberklang *et al.* (1977). [5′-^{32}P]pCp and T4 RNA ligase were used for 3′ labelling (England and Ulhenbeck, 1978).

RNA filter binding assay

The standard reaction buffer for the binding assay was 10 mM HEPES (pH 8.0), 100 mM NaCl, 10 mM $MgCl_2$ and 2.5 mM DTT. The usual reaction volume was 45 µl. Various concentrations of NP protein were incubated with 5′-labelled vRNA or oligoribonucleotides for 15 min at 37°C. The mixture was then filtered through Whatman 0.45 µm pore size nitrocellulose filters (Baudin *et al.*, 1989, 1991). Only the RNA complexed with NP was retained on the filters and was detected by scintillation counting. Since the amount of RNA used was extremely small (0.03 pmol of molecules, 10^4 c.p.m.), the NP concentration at 50% retention corresponds with the apparent dissociation constant K_d. Complex formation was carried out at 37°C after denaturation and renaturation of the RNA probe as described below for the probing experiments.

Probing experiments

Enzymatic digestions were performed on 3′ or 5′ end-labelled RNA (~50 000 c.p.m.). Chemical modifications were performed either on end-labelled RNA or on unlabelled RNA, depending on the detection procedure. Prior to modification, the RNA was denatured in double-distilled water for 1 min at 95°C, quickly cooled on ice (2 min), then slowly brought back to 37°C (20 min) in the appropriate buffer in the presence or absence of nucleoprotein. We also tried denaturation and renaturation of the RNA before addition of the protein with the same results.

Enzymatic probing

Enzymatic probing was carried out with nuclease S1 (32 kDa) which is specific for single-stranded structures without base specificity, RNase T1 (11 kDa) specific for G in single-stranded RNA, and RNase V1 (16 kDa) specific for double-stranded structures or for stacked bases in single strands [for more information, see Ehresmann *et al.* (1987)]. Digestion with RNase T1 (1.25×10^{-3} or 2.5×10^{-3} U), RNase V1 (1.2×10^{-3} or 6×10^{-4} U) or nuclease S1 (3 or 6 U) was for 5 and 15 min at 37°C in 20 µl buffer A [50 mM sodium cacodylate buffer (pH 7.5), 20 mM $Mg(acetate)_2$, 0.3 M KCl, 5 mM DTT]. For nuclease S1, the reaction assay was supplemented with 0.2 mM $ZnCl_2$. For each reaction, a control without RNase was treated in parallel.

Chemical probing

The chemical reagents used allowed us to test whether the Watson–Crick positions (A N-1 and C N-3 with DMS, G N-1 and U N-3 with CMCT, and G N-1 and N-2 with kethoxal) were available or occupied, giving information on the single- or double-stranded nature of the RNA. The reactivity of positions N-7 on purines (G N-7 with DMS and A N-7 with DEPC) gives information on the tertiary structure. The modifications of cytosine residues at N-3 and purine residues at N-7 were revealed by further chemical treatment of the end-labelled modified RNA, resulting in cleavage of the RNA. Modification of purine residues at N-1 and uridine residues at N-3 was identified by the primer extension method. The reactivities of nt 71–81 of the panhandle RNA could not be revealed since this region was used as the hybridization site for the DNA primer. The different probes and their respective uses have been described elsewhere (Ehresmann *et al.*, 1987).

DMS modification. Native conditions, for naked RNA and RNA–NP complex: 0.5 µl DMS was added to the sample in buffer A. The incubation times were 4, 7 or 10 min at 37°C. Semi-denaturing conditions, for naked RNA only: 0.5 µl of DMS in buffer B [50 mM sodium cacodylate buffer (pH 7.5), 1 mM EDTA], 4 min at 37°C. Under such conditions, the tertiary interactions are expected to melt, as are the weakest secondary interactions, which allows determination of the stability of helices. When 3′ end-labelled RNA was used, after modification of cytosine N-3, the strand scission was induced by incubation of the RNA in 10 µl 10% hydrazine for 5 min on ice, followed by incubation in 20 µl 10% aniline (pH 6.5) for 10 min at 60°C in the dark. To detect the modification on G N-7, the RNA was first incubated in 1 M Tris–HCl (pH 8.3) plus 10 µl 0.2 M $NaBH_4$ for 10 min on ice in the dark. The ribose-phosphate chain was then split by treatment with aniline.

DEPC modification. DEPC (20 µl) was added to the sample in buffer A (for native conditions) or in buffer B (for semi-denaturing conditions). The incubation was for 10 or 20 min at 37°C. The 3′ end-labelled RNA was then cut by aniline treatment.

CMCT modification. CMCT (25 µl; 42 mg/ml) was added to the sample in buffer C [50 mM sodium borate (pH 8.0), 20 mM $Mg(acetate)_2$, 300 mM KCl, 5 mM DTT]. The incubation time was 10 or 20 min at 37°C.

Kethoxal modification. Kethoxal (5 µl; 20 mg/ml in 20% ethanol) was added to the sample in buffer A and the incubation time was 10 or 20 min at 37°C. The reaction was then stopped by the addition of 20 µl 50 mM potassium borate (pH 7.0).

Modification of the phosphates

The method we used to test whether the phosphates are involved in tertiary structures in the naked RNA, or whether they interact directly with the NP in the RNA–NP complex, is based on the reactivity of phosphorothioate linkages with iodine. The cleavage depends on the accessibility of the sulphur to the reagent and on the spatial arrangement of the 2′-OH group (Gish and Eckstein, 1988; Schatz *et al.*, 1991). Briefly, four transcription reactions in which one of the NTPs was supplemented by 5% of the corresponding NTP[αS] were performed in parallel with the use of T3 RNA polymerase. They were stopped by the addition of Na_2EDTA to a final concentration of 40 mM. This modified panhandle RNA molecule was then labelled at the 5′ end with [γ-^{32}P]ATP and purified on PAGE. The RNA–NP complex was then made

as described before in 50 µl 10 mM NaCl, 10 mM $MgCl_2$ and 10 mM HEPES (pH 7.5). The naked RNA (also in the above buffer) was treated with a final concentration of 300 µM iodine and the vRNA–NP complex with 100 µM, 300 µM, 600 µM or 1 mM iodine. The incubation time was 1 min at room temperature. The sequence of the RNA was determined in parallel in 10 mM HEPES (pH 7.5) (hence without salt or $MgCl_2$) by incubation with 300 µM iodine.

Detection and analysis of the modified positions

After enzymatic and chemical reactions, the RNA was eventually phenol extracted and precipitated with ethanol in the presence of 0.3 M sodium acetate (pH 6.8). The precipitated RNA was washed with 70% ethanol, vacuum dried and redissolved in double-distilled water. Cuts in the RNA or modified positions were detected by the primer extension method using reverse transcriptase, as described in Baudin *et al.* (1987, 1993). An oligodeoxyribonucleotide complementary to nt 81–69 of the panhandle RNA was labelled at its 5′ end with [γ-^{32}P]ATP and used as a primer for reverse transcription. Hybridization and reverse transcription were carried out as described in Baudin *et al.* (1993), except that the RNA template was hydrolysed after the elongation reaction by the addition of 3 µl 3 M KOH and incubation at 90°C for 3 min, followed by 1 h incubation at 37°C. The reaction was stopped by adding 6 µl 3 M acetic acid. The cDNA fragments were then precipitated and sized at nucleotide resolution by PAGE on 12% acrylamide/0.5% bis(acrylamide)/8 M urea slab gels (40 × 30 × 0.04 cm) at 1500 V for 2 h. Dideoxy sequencing reactions were carried out in parallel on unmodified RNA as described by Sanger *et al.* (1977). Incubation controls were run in parallel in order to detect nicks in the unmodified RNA and pauses of the reverse transcriptase due to secondary RNA structure.

Alternatively, the terminal parts of the vRNA were probed directly by using end-labelled RNA, and the resulting RNA fragments were analysed by PAGE sizing. Cleavage positions were identified by running in parallel RNase T1 and U2 digests plus a formamide ladder.

Acknowledgements

We thank Brian Sproat (EMBL, Heidelberg) for oligoribonucleotide synthesis, Hans Hegemann (Institut für Mikrobiologie und Molekularbiologie, Giessen) for oligodeoxyribonucleotide synthesis, Mark Krystal (Bristol-Myers Squibb, Princeton, NJ) for the gift of the pHgaNS clone, Christoph Scholtissek for suggesting a collaboration between his laboratory and the EMBL Grenoble Outstation, Franck Borel for the gift of yeast $tRNA^{Ser}$, Elisabeth DiCapua for stimulating discussions and Reuben Leberman for critical comments on the text. This work was financially supported in part by the Deutsche Forschungsgemeinschaft, SFB 272.

References

Baudin,F., Ehresmann,C., Romby,P., Mougel,M., Colin,J., Lempereur,L., Bachellerie,J.P., Ebel,J.P. and Ehresmann,B. (1987) *Biochimie*, **69**, 1081–1096.

Baudin,F., Mougel,M., Romby,P., Eyermann,F., Ebel,J.P., Ehresmann,B. and Ehresmann,C. (1989) *Biochemistry*, **28**, 5847–5855.

Baudin,F., Romaniuk,P.J., Romby,P., Brunel,C., Westhof,E., Ehresmann,B. and Ehresmann,C. (1991) *J. Mol. Biol.*, **218**, 69–81.

Baudin,F., Marquet,R., Isel,C., Darlix,J.L., Ehresmann,B. and Ehresmann,C. (1993) *J. Mol. Biol.*, **229**, 382–397.

Beaton,A.R. and Krug,R.M. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 6282–6286.

Blumberg,B.M., Giorgi,C. and Kolakofsky,D. (1983) *Cell*, **32**, 559–567.

Cantor,C.R. and Schimmel,P.R. (1980) *Biophysical Chemistry Part III*. W.F.Freeman and Co., San Francisco, CA, pp. 1146–1160.

Compans,R.W., Content,J. and Duesberg,P.H. (1972) *J. Virol.*, **10**, 795–800.

Desselberger,U., Racaniello,V.R., Zazra,J.J. and Palese,P. (1980) *Gene*, **8**, 315–328.

Duesberg,P. (1969) *J. Mol. Biol.*, **42**, 485–499.

Ehresmann,C., Baudin,F., Mougel,M., Romby,P., Ebel,J.P. and Ehresmann,B. (1987) *Nucleic Acids Res.*, 15, 9109–9128.

England,T.E. and Ulhenbeck,O.C. (1978) *Nature*, **275**, 560–604.

Gish,G. and Eckstein,F. (1988) *Science*, **240**, 1520–1522.

Goldstein,E.I and Pons,M.W. (1970) *Virology*, **41**, 382–384.

Heinemann,U. and Saenger,W. (1982) *Nature*, **299**, 27–31.

Honda,A., Ueda,K., Nagata,K. and Ishihama,A. (1987) *J. Biochem.*, **102**, 1241–1249.

Honda,A., Ueda,K., Nagata,K. and Ishihama,A. (1988) *J. Biochem.*, **104**, 1021–1026.

Hsu,M.T., Parvin,J.D., Gupta,S., Krystal,M. and Palese,P. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 8140–8144.

Jennings,P.A., Finch,J.T., Winter,G. and Robertson,J.S. (1983) *Cell*, **34**, 619–627.

Kingsbury,D.W. and Webster,R.G (1969) *J. Virol.*, **4**, 219–225.

Kingsbury,D.W., Jones,I.A. and Murti,K.G. (1987) *Virology*, **156**, 396–403.

Krug,R.M. (1983) In Palese,P. and Kingsbury,D.W. (eds), *Genetics of Influenza Viruses*. Springer-Verlag, Berlin, pp. 70–98.

Lamb,R.A. and Choppin,P.W. (1983) *Annu. Rev. Biochem.*, **52**, 467–506.

McCauley,J.W. and Mahy,B.M. (1983) *Biochem. J.*, **211**, 281–294.

Palese,P. (1979) *Cell*, **10**, 1–10.

Pause,A. and Sonenberg,N. (1993) *Curr. Opin. Struct. Biol.*, **3**, 953–959.

Philippe,C., Portier,C., Mougel,M., Grunberg-Manago,M., Ebel,J.P., Ehresmann,B. and Ehresmann,C. (1990) *J. Mol. Biol.*, **211**, 415–426.

Pons,M.W., Schulze,I.T. and Hirst,G.K. (1969) *Virology*, **39**, 250–259.

Robertson,J.S. (1979) *Nucleic Acids Res.*, **6**, 3745–3757.

Romby,P., Westhof,E., Moras,D., Giege,R., Houssier,C. and Grosjean,H. (1986) *J. Biomol. Struct. Dyn.*, **4**, 193–203.

Sanger,F.S., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.

Schatz,D., Leberman,R. and Eckstein,F. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 6132–6136.

Scholtissek,C. and Becht,H. (1971) *J. Gen. Virol.*, **10**, 11–16.

Shapiro,G.I. and Krug,R.M. (1988) *J. Virol.*, **62**, 2285–2290.

Shinagawa,M. and Padmanabhan,R. (1979) *Anal. Biochem.*, **95**, 458–464.

Silberklang,M., Gillum,A.M. and Rajbhandary,C.L. (1977) *Nucleic Acids Res.*, **4**, 4091–4108.

Skehel,J.J. and Hay,A.J. (1978) *Nucleic Acids Res.*, **5**, 1207–1219.

Stoeckle,M.Y., Shaw,M.W. and Choppin,P.W. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 2703–2707.

Winter,G. and Fields,S. (1981) *Virology*, **114**,423–428.

Received on February 22, 1994; revised on April 20, 1994