The EMBO Journal vol.13 no.17 pp.4165–4173, 1994

Four nucleotides are the minimal requirement for RNA recognition by rotavirus non-structural protein NSP3

Didier Poncet[1], Sylvie Laurent and Jean Cohen

Laboratoire INRA de Virologie et d'Immunologie Moléculaires, 78350 Jouy-en-Josas, France

[1]Corresponding author

Communicated by J.Tooze

The interaction of the group A rotavirus non-structural protein NSP3 (NSP3A) with RNA has been studied *in vitro*. Using semi-purified NSP3A protein expressed by a recombinant baculovirus and *in vitro* synthesized RNA, we determined by UV cross-linking and gel retardation assays that NSP3A binds, in a sequence-specific manner, the consensus sequence (AUGUGACC) present on the 3′ ends of all group A rotavirus mRNAs. Using short oligoribonucleotides, we established that the minimal RNA sequence required for binding of NSP3A is GACC. Modifications of the UGACC oligonucleotide sequence impaired binding of the protein to the RNA. Furthermore, the recombinant NSP3 protein from rotavirus group C showed specificity for the 3′ end consensus sequence (AUGUGGCU) of only group C mRNAs. Sequence analysis of the NSP3 proteins did not reveal significant homologies with other RNA binding proteins, thus the NSP3 proteins of rotaviruses are the prototypes of a new kind of sequence-specific RNA binding protein.

Key words: 3′ end/non-structural protein/RNA binding/rotavirus/sequence-specific

Introduction

Rotaviruses are members of the Reoviridae family and their genome is composed of 11 molecules of double-stranded RNA (dsRNA), ranging in size from 3.3 to 0.6 kb. These RNAs encode six structural proteins and five non-structural proteins (Mattion *et al.*, 1994). Virus replication occurs in the cell cytoplasm; as the virus enters the cell, the viral transcriptase is activated and synthesizes capped, non-polyadenylated mRNAs. These viral mRNAs are either translated or used as templates for the synthesis of the genomic dsRNAs. Replication is non-conservative; viral mRNAs are exactly copied into the negative strand and the dsRNAs thus formed are encapsidated in new viral particles. During this cycle, the mRNAs are not linked together, as co-infection of the same cell with different strains of virus can result in the exchange of genes (reassortment) at high frequency (Ramig and Ward, 1991). It is thought that replication is a precise process, as viral particles released from the cells contain only one set of the 11 genes and only one copy of each gene.

One challenging question about the morphogenesis of the Reoviridae is how does the precise set of dsRNA segments become encapsidated in viral particles. RNA–protein interactions should be essential for proper encapsidation. In the Reoviridae, the viral mRNAs bear 5′ and 3′ untranslated regions of variable length and are bordered by two different sequences common to all genes. In the case of group A rotaviruses, the first four and the last five nucleotides are strictly conserved among all the segments. The 5′ consensus sequences for group A rotavirus RNAs (5′Acs) GGCU(A,U)$_n$ (n = 6–9) is more heterogeneous than the 3′ end consensus sequence UGUGACC (3′Acs). The latter is highly conserved in the 11 genes, except for slight differences in genes two (UAUGACC) and three (CGUGACC). The 3′ and 5′ consensus sequences are different from one rotavirus serogroup to another, for example in group C rotaviruses the 3′ end consensus sequence (3′Ccs) is UGUGGCU (Brémont *et al.*, 1990; Qian *et al.*, 1991a,b). Reassortment between rotavirus strains of different serogroups has not been observed (Yolken *et al.*, 1988) and is thought to be restricted by these different consensus sequences.

We have previously shown that in group A rotavirus-infected cells, the non-structural protein NSP3 is bound to the 3′ conserved sequence of rotavirus mRNAs (Poncet *et al.*, 1993). In the present report we have investigated the RNA sequence needed for binding of NSP3 *in vitro*.

Results

Recombinant NSP3A binds to the 3′ end of rotaviral mRNAs

The recombinant protein NSP3A (rNSP3A) was semi-purified from infected Sf9 cells. In non-reducing conditions, the recombinant protein forms multimers, like the natural protein made in rotavirus-infected cells (Mattion *et al.*, 1992; Aponte *et al.*, 1993; Poncet *et al.*, 1993). To determine if the semi-purified rNSP3A was able to bind viral mRNAs, we performed a UV cross-linking immunoprecipitation assay, similar to that used for rotavirus-infected cells (Poncet *et al.*, 1993). Purified rNSP3A was allowed to bind RNAs made *in vitro* and the RNA–protein mix was subjected to UV cross-linking. After immunoprecipitation, protein–RNA complexes were treated with RNase T1 (G specific), the RNAs in the protein–RNA complexes were labelled with T4 polynucleotide kinase and [γ-^{32}P]ATP and then analysed on denaturing polyacrylamide gels. Rotavirus mRNAs made *in vitro* by the viral transcriptase and RNA made *in vitro* by T7 polymerase from the plasmid p71Cat (Figure 1A) were used in this experiment. With the plasmid p71Cat linearized by *Kpn*I, we obtained by *in vitro* transcription a rotavirus-like RNA bearing the 3′ and 5′ non-coding sequences of the rotavirus gene 9 but where the major part of the coding sequence was replaced by the bacterial chloramphenicol acetyltransferase

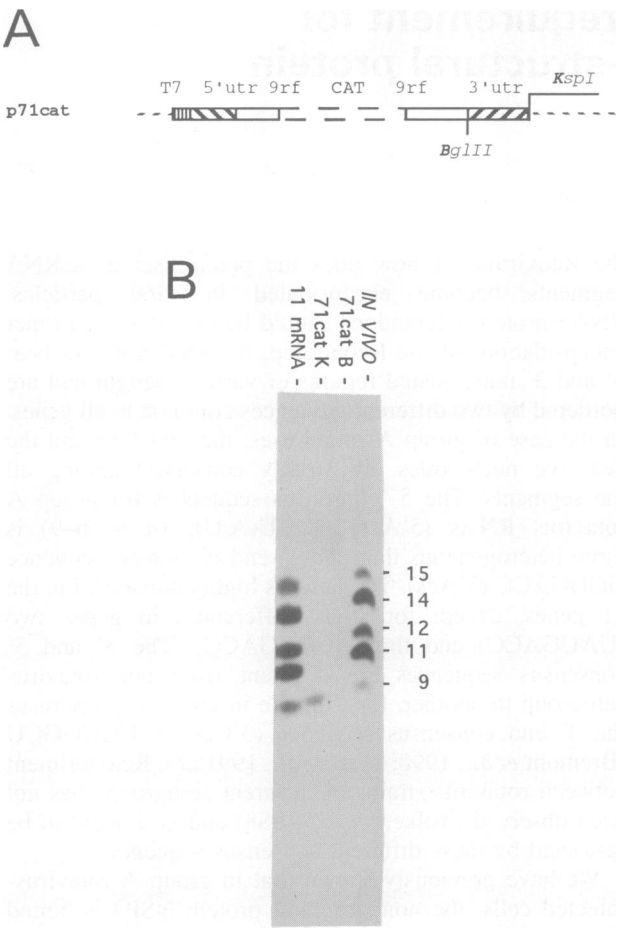

Fig. 1. Comparison of the RNA cross-linked to rNSP3A after binding with *in vitro* synthesized RNA. (**A**) Schematic representation of the p71cat plasmid used for RNA synthesis. The T7 promoter (T7), the 5′ and 3′ untranslated sequence of RF gene 9 (utr), the Cat gene coding sequence (CAT), the remaining 5′ and 3′ coding sequence of RF gene 9 (9rf), the plasmid sequence (dashed line) and the *Bgl*II and *Ksp*I restriction sites are indicated. (**B**) The RNA–protein complexes obtained after UV cross-linking of rotavirus-infected cells or after binding of rNSP3A with *in vitro* synthesized viral mRNA were immunoprecipitated with an anti-NSP3A monoclonal antibody and treated with RNase T1. After 5′ end labelling, RNA freed from NSP3A was analysed in denaturing SDS–PAGE. 11 mRNA, mRNAs made *in vitro* from viral particles; 71cat, RNA obtained by T7 RNA polymerase synthesis from p71cat digested with *Ksp*I (K) or *Bgl*II (B); *in vivo*, RNA obtained after UV cross-linking of rotavirus infected cells. The apparent sizes of the RNA fragments are indicated on the right.

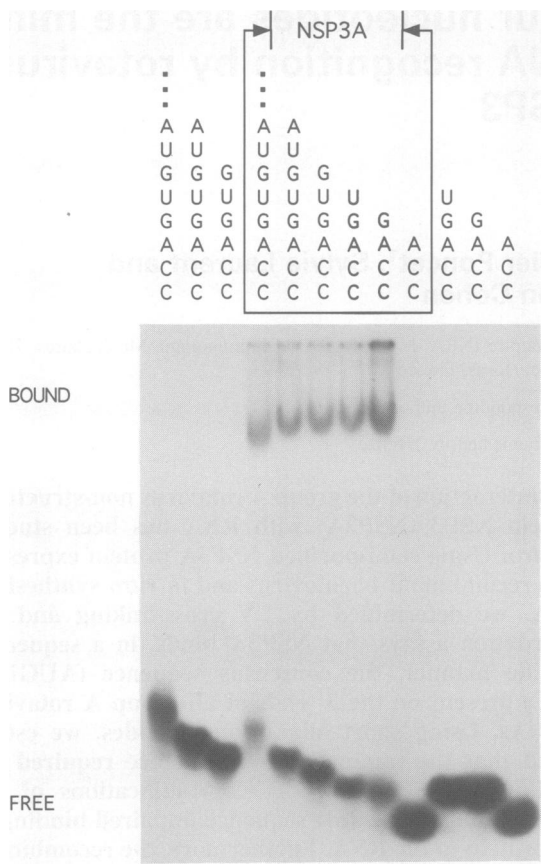

Fig. 2. Gel retardation assay with rNSP3A and synthetic 3′Acs RNAs of decreasing length. The base sequences of the different RNAs used are indicated at the top of the gel (the 20 base long RNA is symbolized by ... AUGUGACC). The position of the reactions containing the rNSP3A protein is framed.

gene. When the plasmid is linearized by *Bgl*II, a RNA with no rotaviral 3′ non-coding sequence is obtained (Figure 1A). The pattern of RNA obtained with rNSP3A and *in vitro* made rotavirus mRNA was identical to the one obtained from *in vivo* cross-linked NSP3A (Figure 1B). When a rotavirus-like RNA made from plasmid DNA was used, a RNA with an apparent length of nine nucleotides was obtained (Figure 1B). Conversely, with the RNA lacking the rotaviral 3′ non-coding sequence, no RNA was cross-linked to NSP3A.

Sequence specifity of NSP3A binding on RNA

We next examined whether rNSP3A bound directly to the 3′ consensus sequence or if a more complex sequence or a secondary structure was necessary. This was important because if binding occurs with a RNA lacking a large part of the rotaviral mRNA, as we have shown above, secondary structures in the 3′ non-coding sequence (Hua *et al.*, 1993) or between 5′ and 3′ non-coding sequences (Patton *et al.*, 1993) might still persist in this RNA.

We chemically synthesized oligoribonucleotides representing the 20 or the eight last nucleotides of rotavirus RF gene 9 and the eight last nucleotides of RF gene 2. Controls were provided by oligonucleotides bearing the rotavirus group A 5′cs in the plus or minus polarity, reverse 3′cs or shuffled 3′cs. Other controls included 3′cs of other Reoviridae. All were tested with the recombinant NSP3A protein in UV cross-linking and gel retardation assays. Representative results of these two assays are illustrated in Figures 2 and 3 and the results obtained with the probes tested are summarized in Table I (lines 1–11).

In the gel retardation assay (Figure 2), a single RNA–NSP3A complex was observed with only some of the probes used. The RNA–protein complexes observed after UV cross-linking and SDS–PAGE in non-reducing conditions (Figure 3) were of the sizes expected for NSP3A monomers and multimers. Moreover, all the probes leading to a NSP3A–RNA complex by UV cross-linking led to a RNA–protein complex by gel retardation. The results of the two tests indicated that the RNA–protein complexes observed in the gel retardation assay were due to NSP3A and not to another RNA binding protein present

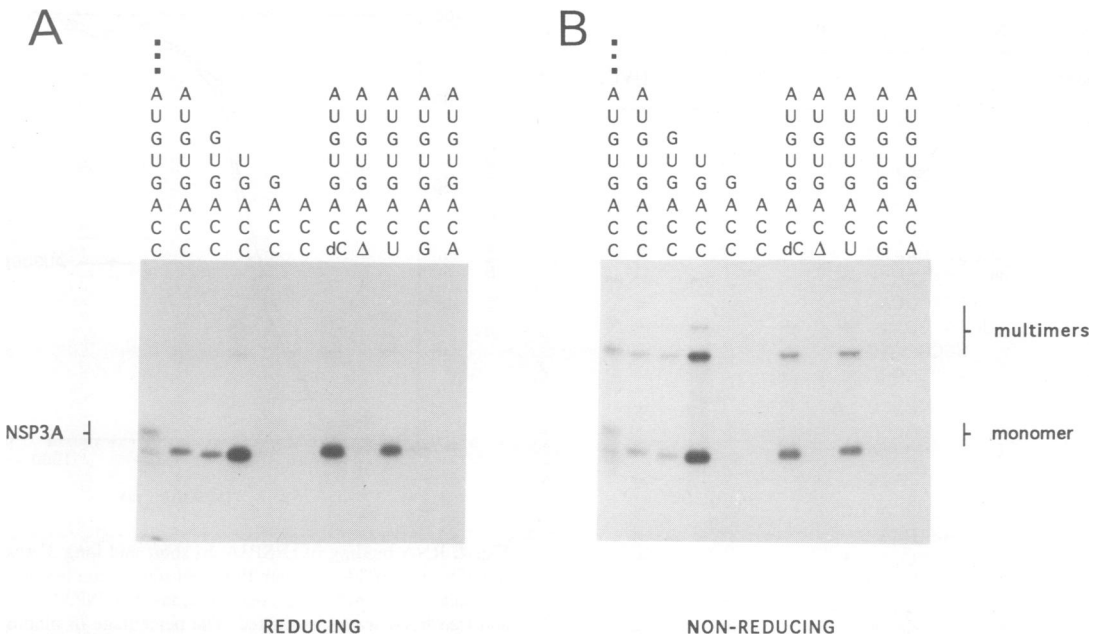

Fig. 3. UV cross-linking of rNSP3A with synthetic 3′cs RNAs of different sequence. NSP3A protein was allowed to bind to the labelled RNAs, cross-linked by UV irradiation and then analysed by SDS−PAGE in reducing (**A**) or non-reducing (**B**) conditions.

in our NSP3A preparation. By comparing the results obtained with the oligonucleotides that included the 3′cs and those corresponding to other sequences, it was evident that only group A rotavirus 3′cs could be bound by NSP3A. Probes bearing Orthoreoviridae, group C rotavirus or blue tongue virus (BTV) 3′cs or rotavirus 5′Acs did not lead to a NSP3A−RNA complex. Since an eight base long oligoribonucleotide could form a RNA−protein complex, it is clear that NSP3A does not need a complex RNA structure to bind to RNA. Furthermore, the absence of binding of NSP3A with the rotavirus group C 3′cs, which differ from group A by only two bases, clearly indicates that this protein possesses a remarkable sequence specificity.

It is also worth noting that a single RNA−rNSP3A retarded complex coud be observed by gel retardation (Figure 2), indicating that only one form of rNSP3A was able to bind to the RNA. However, monomers as well as multimers were observed after UV cross-linking and SDS−PAGE in non-reducing conditions (Figure 3). The dimeric and monomeric forms observed in SDS−PAGE probably result from the dissociation of the trimeric form of rNSP3A.

Minimum sequence required for binding and cross-linking

To determine the minimal RNA sequence length necessary for binding of rNSP3A, we synthesized a series of nested oligonucleotides corresponding to shorter (in the 5′−3′ direction) 3′Acs and tested their binding capacity (Figures 2 and 3). Gel retardation assay showed that a four base long oligonucleotide (GACC) was necessary and sufficient to lead to a retarded complex, but a five nucleotide long RNA (UGACC) was necessary to obtain a RNA−protein complex by UV cross-linking.

We then performed quantitative studies of the interaction between rNSP3A and the eight and five nucleotides long 3′Acs. First, the number of RNA−protein complexes as a function of rNSP3A concentration was established for the AUGUGACC and the UGACC oligonucleotides (Figure 4). It should be noted that only 60% of the RNA probes tested here was bound, even at the highest rNSP3A concentration. This reflects the fact that the synthetic RNAs were only partially active, probably due to partial deprotection. Under those conditions, half of the maximum binding was obtained for the AUGUGACC and UGACC oligoribonucleotides at a 25 nM concentration of rNSP3A. Furthermore, the shape of the curves indicated that binding of rNSP3A to RNA was cooperative (Senear and Brenowitz, 1991). The similarity of the results obtained with these two probes showed that all the information necessary for the binding of rNSP3A to RNA was contained in a five base long RNA. Binding constants were determined at equilibrium with the AUGUGACC and UGACC oligoribonucleotides. For this analysis we used a concentration of protein close to the plateau of maximum binding to circumvent the problem of cooperativity. The Scatchard plot analysis (Figure 5) gave a nearly equivalent apparent binding constant for the AUGUGACC oligoribonucleotide (K_d = 25 nM) and the UGACC oligoribonucleotide (45 nM), confirming that the shortest oligoribonucleotide contained most of the information necessary and sufficient for RNA−rNSP3A interaction.

Importance of the terminal C ribonucleotide

From here, the 3′Acs bases will be numbered in the 3′−5′ direction, C1 being the most 3′ nucleotide.

Depending on the binding test used, deletion of A8−U5 did not impair binding of rNSP3A to RNA. However, deletion of only the terminal C totally abolished binding (Figures 3 and 6), as did the addition of only one C (Table I). Together with our previous localization of the RNA−protein cross-linked site as the last nucleotide of the RNA (Poncet *et al.*, 1993), these observations prompted

Table I. Binding of rNSP3A to synthetic RNAs

	Sequences	GR	UVCL
1	UAGAGGUGUACGAUGUGACC	+	+
2	AU**A**UGACC	+	+
3	AUGUGACC	+	+
4	GGUCACAU	−	−
5	CCAGUGUA	−	−
6	GGCUUUUAAAG	−	−
7	CUUUUAAAGCC	−	−
8	UCACACUUAC	−	−
9	UCAACUCAUC	−	−
10	UUCGCGAG	−	−
11	AUGUGGCU	−	−
12	GUGACC	+	+
13	UGACC	+	+
14	GACC	+	−
15	ACC	−	−
16	AUGUGACC**C**	−	−
17	AUGUGAC	−	−
18	AUGUGACC*	+	+
19	AUGUGAC**d**C	+/−	+
20	AUGUGAC**U**	+	+
21	AUGUGAC**G**	−	−
22	AUGUGAC**A**	−	−
23	**C**GACC	+	−
24	**G**GACC	+	−
25	**A**GACC	+	−
26	U**A**ACC	−	−
27	U**U**ACC	−	−
28	U**C**ACC	−	−
29	UG**G**CC	+/−	−
30	UG**C**CC	−	−
31	UG**U**CC	+/−	−
32	UGA**A**C	−	−
33	UGA**U**C	−	−
34	UGA**G**C	−	−

The results obtained with the different oligoribonucleotides used in gel retardation (GR) and UV cross-linking (UVCL) with rNSP3A are indicated (+ indicates binding and − indicates absence of binding). The chemically synthesized oligoribonucleotides represent: the 20 or the eight last nucleotides of RF gene 9 (lines 1 and 3); the eight last nucleotides of RF gene 2 (line 2); the antiparallel 3′Acs (line 4); reverse 3′Acs (line 5); the rotavirus group A 5′cs in the plus (line 6) or minus polarity (line 7); shuffled 3′Acs (line 10). Other controls includes 3′cs of other Reoviridae, including BTV (line 8), orthoreovirus (line 9) and rotavirus group C (line 11). Lines 12–15 represent shorter 3′Acs. Lines 16–22 are modifications of the 3′end nucleotides and lines 23–34 are mutations of the U5−C2 bases. The mutations introduced are indicated in bold. Number 18 corresponds to line 17 labelled on the 3′ end with [^{32}P]pCp and RNA ligase; the 3′ phosphate is indicated by *.

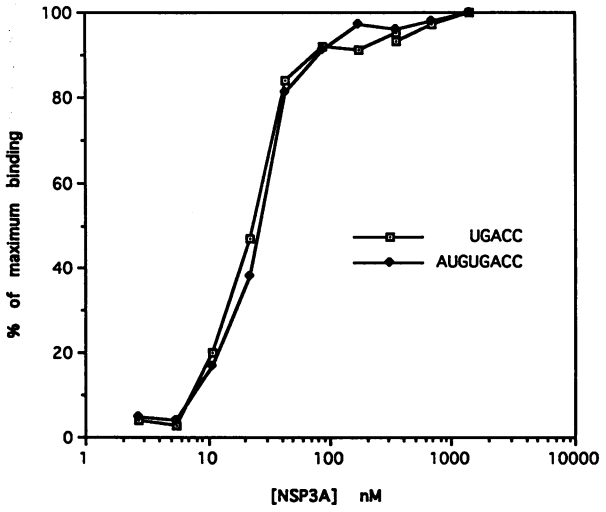

Fig. 4. RNA binding of rNSP3A on short and long 3′Acs as a function of NSP3A concentration. Gel retardation assays were conducted with increasing concentrations of rNSP3A and the bound and free RNA were quantitated. The percentage of binding relative to the maximum binding (obtained with the highest concentration of protein and corresponding to 60% of the input probe) is represented as a function of monomeric rNSP3A concentration.

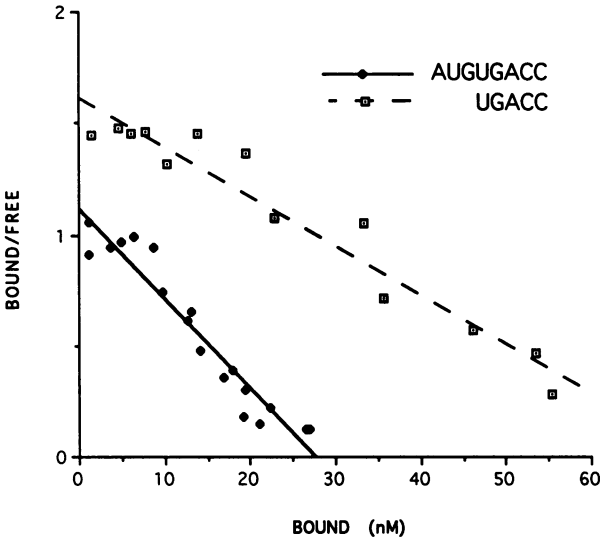

Fig. 5. Scatchard plot analysis of rNSP3A binding on 3′Acs. Oligoribonucleotide UGUGACC: y intercept = 1.115, x intercept = 27 nM and R^2 = 0.928; oligoribonucleotide UGACC: y intercept = 1.614, x intercept = 75 nM and R^2 = 0.939.

us to study whether important contacts between rNSP3A and the last nucleotide occur on the base or the sugar. First, three eight nucleotide long RNAs were synthesized with the last C base being changed for one of the three other possible bases. From the results obtained with those probes (Figures 3 and 6) it could be concluded that only a pyrimidine could be positioned in the 3′ position. Second, as the chemistry of RNA synthesis is very similar to the chemistry of DNA synthesis, we reasoned that synthesis of the RNA on a dC column should result in a change in the last ribose to a deoxyribose, without any change of the base sequence. Changing the last ribose to a deoxyribose did not impair cross-linking of the protein to RNA (Figure 3), but in the gel retardation assay the probe made a smear all along the gel, instead of a clear band (Figure 6). This is probably the effect of a weakened association between the RNA probe and the protein that does not withstand electrophoresis completely, resulting in dissociation of the RNA from the RNA−protein complex during running of the gel. Thus, contacts important for the stability of the complexes occur between NSP3A and the 3′ ribose 2′ OH group.

As shown above, deletion of the last nucleotide abolished binding, but adding it back in the form of a 3′ phosphate base by ligation of pCp to the AUGUGAC oligonucleotide restored binding (Table I, lines 17 and 18). Thus a phosphate group at the 3′ end of the oligonucleotide did not prevent binding of NSP3A.

Taken together, these results showed that, in contrast

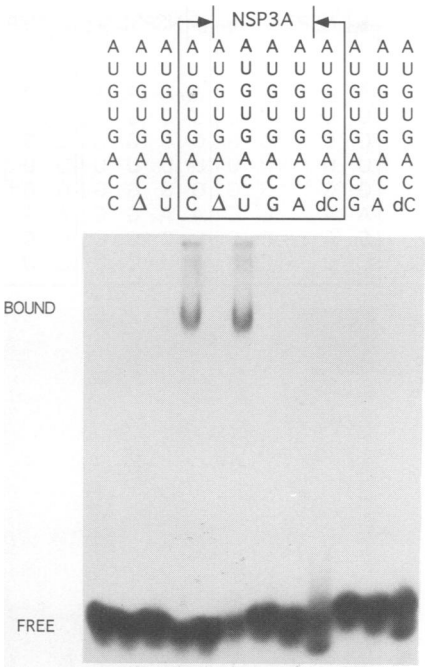

Fig. 6. Gel retardation assays with the NSP3A recombinant protein and synthetic 3′Acs RNA with different 3′ end nucleotides. The base sequences of the different RNAs used are indicated at the top of the gel (the last base deletion is indicated by Δ). The position of the reactions containing the rNSP3A protein is framed.

with the 5′ part of the 3′cs, which can be deleted up to U5 without affecting the binding of NSP3A, the 3′ nucleotide of the 3′Acs is very important for stability of the RNA–protein complex.

Mutational analysis of the 3′Acs

We have shown that a U could replace the C1 base of the 3′Acs with no change in rNSP3A binding characteristics and that the five nucleotide long RNA UGACC reacted in our two binding tests similarly to a full-length 3′Acs. We next carried out saturation mutagenesis on the first four bases of this five base long oligonucleotide (Table I, lines 23–34).

Changing any of the U5–C2 bases modified the binding characteristics in at least one of the binding tests we used (Table I, line 23–34). The most dramatic effect of mutations was observed in the cross-linking assay; none of the mutated oligoribonucleotides could be cross-linked to rNSP3A. Replacement of U5 by any other base were the only mutations that did not modify the binding of NSP3A as judged by gel retardation. Two other mutations, A3→G and A3→U, only partially impaired the gel retardation assay. As for the C1→dC change, mutation of A3 resulted in a RNA–protein interaction too weak to withstand electrophoresis completely. The mutations with the greatest effect were those on the C2 and G4 bases; any mutation of these two bases totally abolished binding, underlining the importance of these bases.

From these saturation mutagenesis experiments it appeared that the specificity of binding of rNSP3A was localized in the four last bases of the 3′Acs and that any base change in the U5–C2 bases modified the affinity of NSP3A for RNA.

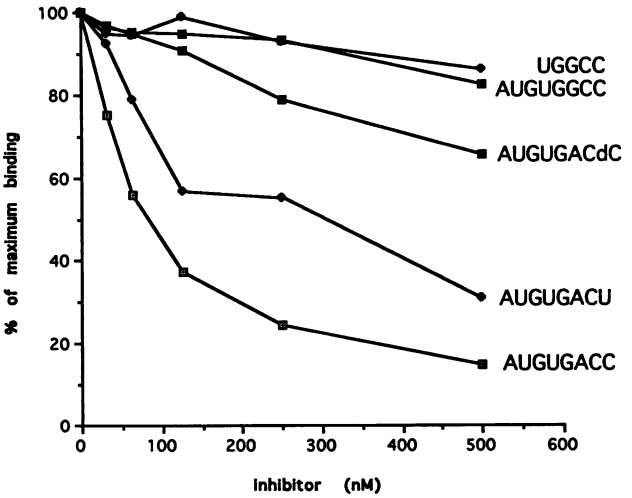

Fig. 7. Competition of different RNA for 3′Acs binding on rNSP3A. Labelled 3′Acs (AUGUGACC) bound by rNSP3A in the presence of increasing concentrations of unlabelled competitors was quantitated. The sequences of the cold competitors are indicated. Results are presented as a percentage of the maximum binding (obtained without competitors and corresponding to 60% of the input probe).

Competition of mutated oligoribonucleotide for rNSP3A binding to 3′Acs

In order to compare the effect of some of the mutations on the affinity of rNSP3A for RNA, we undertook quantitative studies. We first compared the C1→U change with the C1→dC mutation. The results of gel retardation competition experiments with a constant quantity of the labelled AUGUGACC probe and an increasing concentration of cold oligoribonucleotides (Figure 7) showed that the AUGUGACU oligoribonucleotide did not compete as well as the 3′Acs itself, but it was a better inhibitor than AUGUGACdC. The latter reduced binding to 30% at a 200-fold molar excess over the probe, while at the same concentration the binding of rNSP3A to the 3′Acs was reduced to only 70% with the AUGUGACU oligoribonucleotide. This result was consistent with the behaviour of the probes in the gel retardation assay, as AUGUGACdC led to a short smear above the free probe, in contrast to the AUGUGACU oligoribonucleotide which gave a discrete band (Figure 6). This comparison showed that changing the last ribose 2′ OH group affected binding more adversely than a base transversion on the last nucleotide.

To determine which base change, of the two bases that differ between the group C rotavirus 3′ conserved sequence and 3′Acs (C1→U and A3→G), affected binding of rNSP3A more adversely, the eight and five nucleotide long oligoribonucleotides bearing the A3→G mutation were used in a competition experiment (Figure 7). These two oligoribonucleotides were very poor inhibitors, regardless of their length. Comparison of the C1→U and A3→G mutations showed that the A3 base was the most discriminating between group A and C 3′cs for rNSP3A binding.

Rotavirus group C NSP3 also shows sequence-specific binding

The fact that the recombinant NSP3A protein was unable to bind to the 3′cs from group C rotaviruses (AUGUGGCU)

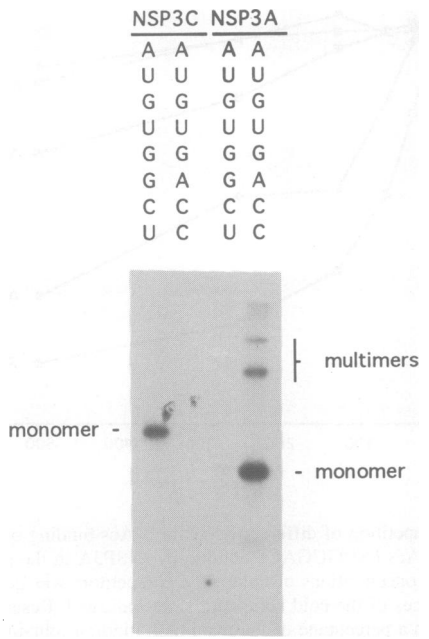

Fig. 8. UV cross-linking of rNSP3A and rNSP3C with 3′Acs and 3′Ccs. Extracts of Sf9 cells infected with recombinant baculovirus expressing NSP3A or NSP3C were used in a UV cross-linking experiment with labelled 3′Ccs and 3′Acs and analysed by non-reducing SDS–PAGE.

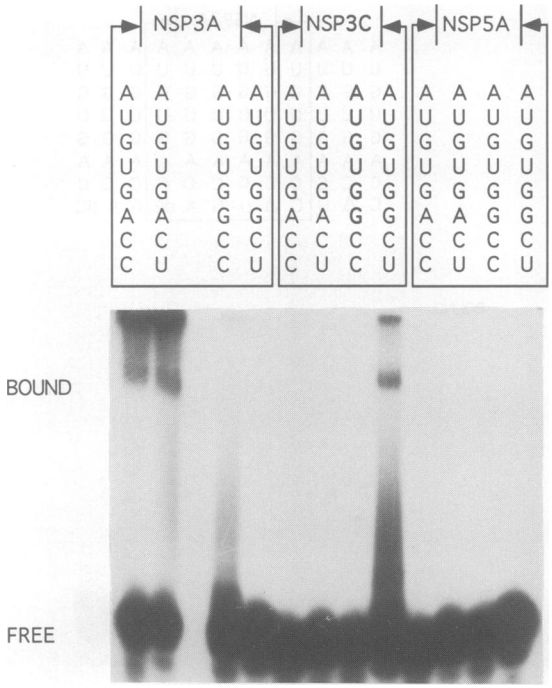

Fig. 9. rNSP3A and rNSP3C interactions with 3′Acs, 3′Ccs and mutated 3′Acs. Cell extracts of Sf9 cells infected with recombinant baculoviruses expressing NSP3A, NSP3C or NSP5A were used in a gel retardation assay with wild-type and mutated 3′Acs and 3′Ccs RNAs. The reactions containing extracts of Sf9 cells infected with recombinant baculoviruses expressing NSP3A, NSP3C or NSP5A are framed.

raised the question of whether the binding properties of NSP3A protein to the 3′Acs of viral mRNAs is particular to this group of rotaviruses or if it is a general property of the NSP3 protein. To test this possibility, the gene coding for NSP3 from group C rotaviruses was cloned, the protein (NSP3C) was expressed using the baculovirus expression system and an infected cell extract was used in a UV cross-linking assay. As a positive control, a cell extract obtained under the same conditions from Sf9 cells infected by a NSP3A recombinant baculovirus was used. The results (Figure 8) showed that the rNSP3C preparation bound only the 3′Ccs. The migration of the bands obtained after UV cross-linking corresponded to the size of NSP3A and NSP3C proteins, showing that the binding observed was due to the rNSP3 proteins and not to another protein present in the cell extract. It should be noted that no multimers were observed with rNSP3C, even under non-reducing conditions.

To compare further the sequence specificity of NSP3C and NSP3A, RNA probes bearing separately each of the two differences between 3′Acs and 3′Ccs were used in UV cross-linking and gel retardation assays. An extract prepared from Sf9 cells infected with the recombinant baculovirus expressing rNSP5A (another rotaviral non-structural protein; Mattion *et al.*, 1994) was used as a negative control. Figure 9 shows the results obtained with these three cell extracts in the gel retardation assay and Table II shows the results obtained in the two tests with rNSP3A and rNSP3C cell extracts.

A comparison of the results obtained with the purified rNSP3A (Figure 6) and the cells extract (Figure 9) validated the assay with the cell extract, because under those conditions NSP3A bound to the 3′Acs and AUGUGACU oligonucleotides and weakly to AUGUGGCC, like the purified protein. The only difference between the purified rNSP3A and the Sf9 cell extract was the material remaining in the wells of the gel; this material was only present with the probes leading to a specific binding and thus should represent aggregates of NSP3A with other cellular proteins or highly multimeric forms of NSP3A. Furthermore, no signal was obtained with the control NSP5A cell extract, showing that no RNA binding proteins other than the one studied here were detectable.

NSP3C also presents a sequence-specific affinity for the 3′cs of group C rotaviruses and thus the specific binding of NSP3 protein to the 3′cs is most probably a general property of this rotaviral protein.

Table II. Binding of rNSP3A and rNSP3C on the 3′Acs and 3′Ccs

		NSP3A		NSP3C	
		GR	UVCL	GR	UVCL
3′Acs	AUGUG**A**C**C**	+	+	−	−
	AUGUG**G**C**C**	+/−	+	−	−
	AUGUG**A**C**U**	+	+	−	−
3′Ccs	AUGUG**G**C**U**	−	−	+	+

The results obtained with the different oligoribonucleotides used in gel retardation (GR) and UV cross-linking (UVCL) with rNSP3A are indicated (+ indicates binding and − indicates absence of binding).

Discussion

The results presented in this study clearly demonstrate that the rotaviral non-structural protein NSP3 can, in the absence of all other viral proteins, bind the 3′ end of

rotaviral mRNAs and that the canonical UGACC sequence is necessary and sufficient for this interaction. The similarity of the RNA cross-linked to NSP3A in the course of rotaviral infection and the RNA cross-linked to rNSP3A *in vitro* suggests that the *in vitro* assays accurately reflect the properties of NSP3A.

Characteristics of binding

Deletions and mutagenesis of the 3′Acs have shown that a four nucleotide long RNA is sufficient for RNA–NSP3A interaction, but that a five nucleotide long RNA is required to observe a RNA–protein complex by UV cross-linking. Moreover, mutation of U5 did not impair binding when tested in gel retardation, but precluded UV cross-linking (Table I). This slight discrepancy is not surprising, because the two tests used in this study are not exactly equivalent. Gel retardation is dynamic and protein and RNA can dissociate during running of the gel, while UV cross-linking creates a covalent link between the molecules and even brief interactions occurring during the UV irradiation can be observed (Budowsky and Abdurashidova, 1989). Furthermore, UV is 'a zero length cross-linker' (Pashev *et al.*, 1991) and slight conformational changes can impair the cross-linking. Together with our previous localization of the cross-linked site on the C1 nucleotide of the 3′Acs (Poncet *et al.*, 1993), these observations lead us to favour the hypothesis that contact of NSP3A with the uridine at position five is important in conferring the right conformation to the RNA–protein complex. Modifying this contact by mutation or deletion slightly changed the conformation of NSP3A on the RNA and impaired cross-linking.

Our experiments also underline the importance of the C1 nucleotide of the 3′Acs for NSP3A binding. We have shown that contact of NSP3A with the C1 nucleotide occurs on the base and on the sugar. In other words, NSP3A binding to the linear sequence UGACC requires that this sequence be placed at the 3′ end of the RNA. This explains why binding occurs only on the 3′ end of the RNA and not internally. Contrary to the lupus antigen, which is also a 3′ end RNA binding protein (Stefano, 1984), the 3′ OH group is not involved in the mechanism of recognition of NSP3A and a 3′ phosphate as the RNA 3′ end does not impair binding. Introduction of modified bases (Hamy *et al.*, 1993) or sugars (Bardwell *et al.*, 1991) in the RNA sequence will be necessary to determine how NSP3A recognizes the 3′ end of RNA.

It has been shown that NSP3A forms multimers *in vivo* (Mattion *et al.*, 1992; Poncet *et al.*, 1993). After UV cross-linking and SDS–PAGE in non-reducing conditions, NSP3A monomers as well as multimers were seen cross-linked to RNA, but in gel retardation assays a single RNA–protein complex could be seen. Furthermore, the NSP3C protein did not show any multimerization after cross-linking and SDS–PAGE and in gel retardation assays the NSP3C binding on 3′Ccs was not as firm as the NSP3A binding on 3′Acs, as a smear was present above the free probe. One way to reconcile these observations is that multimers of NSP3 bind RNA and that partial (NSP3A) or total (NSP3C) dissociation of the proteins occurs in SDS–PAGE. This hypothesis is in accordance with the observation of cooperativity in the binding of NSP3A to RNA. Cooperativity in binding can result from multiple binding sites to the probes or from multimerization of the proteins on the nucleic acid (Ptashne, 1988; Kim and Little, 1992). We favour the second explanation, because cooperativity was observed with the UGACC probe, which seems too short to accommodate two binding sites.

Possible roles of NSP3 in the rotavirus replication cycle

The characteristics of NSP3 binding on RNA suggest different roles for this protein. As relatively few cellular mRNAs lack a poly(A) tail, and thus woud be susceptible to be bound by NSP3A, only viral mRNA can be bound by NSP3A. We propose that one role for this protein in the replication cycle is to discriminate between viral and cellular RNA. Secondly, as NSP3 binds RNA in the absence of any other viral protein, NSP3 may function in the replication cycle before the constitution of sub-viral particles (Patton and Gallegos, 1990). Gorziglia and Collins (1992) have shown that deletion of the last nine bases of rotaviral mRNA impairs its replication. Together with the data presented here, this observation suggests a role for NSP3 in the commitment of viral mRNAs into the replication pathway. Moreover, it has been observed by immunofluorescence that NSP3A forms fine punctuations in the cytoplasm (our unpublished data) or even filaments (Mattion *et al.*, 1992), instead of being localized in the viroplasms, where viral replication is thought to occur (Petrie *et al.*, 1984). Taken together, these results suggest that NSP3A could be involved in the transport of viral mRNA to the place of viral replication.

Comparisons with other RNA binding proteins

A number of RNA binding proteins have been identified and some have been extensively studied (Steitz, 1990; Kenan *et al.*, 1991; Mattaj, 1993), but few recognize linear RNA sequences. Most often, RNA binding proteins interact with RNA targets which contain various secondary structures, such as stems and loops, pseudo-knots and bulges, as well as three-dimensional shapes (Chastain and Tinoco, 1991; Kenan *et al.*, 1991). In contrast, NSP3 has the remarkable property of binding to a short sequence which is unlikely to be able to make a stable secondary structure. Furthermore, even limited changes in the base sequence impair NSP3–RNA interaction. In this respect, NSP3 is closer to cellular proteins than to other viral RNA binding proteins (Olsen *et al.*, 1990; Puglisi *et al.*, 1992). For example, the cleavage and polyadenylation factor (CPF), which is required for the cleavage and polyadenylation reactions during 3′ processing of mRNA precursor (Sachs and Wahle, 1993), binds to the AAUAAA sequence with base and sugar contacts (Bardwell *et al.*, 1991; Keller *et al.*, 1991). RNAs as short as 10 nucleotides are specifically bound by the CPF complex, the nucleotides outside the AAUAAA sequence being irrelevant. Two CPF polypeptides of 160 and 35 kDa have been cross-linked to the AAUAAA element, but unfortunately none of these proteins has yet been sequenced to allow comparison with NSP3. As mentioned earlier, the lupus antigen (La) shares similar properties with NSP3; this nuclear protein preferentially binds the 3′ end of polIII RNA possessing a stretch of three or four uridylate residues (Stefano, 1984). The sequence specificity of the La protein seems less stringent than NSP3A, as the La protein also

binds RNA with guanylate 3′ termini, although with less affinity. In contrast to NSP3A, the binding of La protein to RNA is abolished by the presence of a phosphate group at the 3′ end.

The amino acid sequence of NSP3A reveals no similarity to the RNP motif present in the La protein and other RNA binding proteins (Kenan *et al.*, 1991). No other RNA binding motifs, like the arginine-rich motif (Lazinski *et al.*, 1989; Puglisi *et al.*, 1992), the zinc finger (Clemens *et al.*, 1993), the KH domain (Siomi *et al.*, 1993) or the RGG and RS motifs (Zamore *et al.*, 1992; Birney *et al.*, 1993), are found in NSP3A. The only homology with other RNA binding proteins is a short motif reported by van Staden *et al.* (1991). These authors have identified a short sequence conserved between different Reoviridae RNA binding proteins; the blue tongue virus NS2, reovirus sigma NS and rotavirus NSP3. Binding of recombinant BTV NS2 to RNA does not seem to be sequence-specific (Zhao *et al.*, 1994) and while sigma NS purified from reovirus-infected cells reportedly prefers reovirus 3′ end mRNA (Stamatos and Gomatos, 1982), the *Escherichia coli* expressed protein does not (Richardson and Furuichi, 1985). Furthermore, mutations outside this motif abolished RNA binding by BTV NS2 (Zhao *et al.*, 1994). Thus, it does not seem that the presence of this motif is a signature of sequence-specific RNA binding proteins. Mutagenesis experiments and chimeric proteins should allow us to define the protein sequence involved in the sequence-specific RNA binding properties of NSP3. The existence of at least five different sub-groups of rotaviruses should allow the NSP3 family to be extended and the protein sequence necessary for sequence-specific RNA binding of this protein to be more precisely defined.

Materials and methods

Expression of recombinant NSP3 in baculovirus

The expression of bovine rotavirus recombinant NSP3A protein and the production of monoclonal antibodies against it have been described (Aponte *et al.*, 1993). The NSP3C coding sequence was obtained by reverse transcription and the polymerase chain reaction (PCR) of genomic dsRNA of the porcine Cowden strain (Quian *et al.*, 1991b) using appropriate primers and cloned in pBlueScript (Stratagene). The integrity of the PCR product was confirmed by *in vitro* transcription and then translation in rabbit reticulocyte lysates. NSP5A of the bovine rotavirus strain RF coding sequence was flanked by two *Bam*HI sites by PCR with appropriate primers. The integrity of the PCR product was confirmed by sequencing. Recombinant baculovirus were obtained as described for NSP3A except that the transfer vector used was pBacPAK1 (Clontech).

Purification of rNSP3A from baculovirus-infected cells took advantage of the insolubility of this protein in non-ionic detergents. Sf9 cells in monolayers were infected with the recombinant baculovirus at a multiplicity of infection of 10 p.f.u./cell and 48 h post-infection the cells were harvested and washed twice in phosphate-buffered saline (PBS). The cell pellets were processed immediately or frozen at −20°C. The cells were lysed by addition of 20 mM HEPES, pH 7.5, 1 μg/ml leupeptin and after centrifugation the pellet was extracted twice with 100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, 0.5% Triton X-100. The pellet of insoluble proteins was dissolved in CAPS buffer (50 mM CAPS, pH 10.8, 10 mM DTT, 2% betaine, 0.5% Brij35, 5 M urea). Protein aggregates were removed by ultracentrifugation (30 min at 50 000 *g*) and the proteins were separated by FPLC on a monoQ column eluted with a linear gradient of 0–1 M NaCl in CAPS buffer. Fractions were analysed on SDS−PAGE and by Coomassie blue staining. Fractions containing rNSP3A were pooled and dialysed for 24 h against renaturation buffer (50 mM Tris, pH 8, 10% glycerol, 5 mM reduced glutathione, 0.5 mM oxidized glutathione). Protein purity was examined on SDS−PAGE and concentration was estimated by Coomassie blue staining using dilutions of known concentrations of bovine serum albumin as standards.

Cell extracts were obtained as above except that the pellet of insoluble protein was dissolved in 50 mM Tris, pH 8, 6 M urea and directly dialysed against renaturation buffer.

RNA−protein interaction in vivo

The protocol described previously (Poncet *et al.*, 1993) was slightly modified with the aim of minimizing RNA degradation. Six hours post-infection, MA104 cells (in 20 cm^2 culture plates) were washed with 2 ml cold TBS (25 mM Tris−HCl, pH 8, 0.7 mM Na_2HPO_4, 5.1 mM KCl, 137 mM NaCl) and exposed on ice to a germicidal UV lamp for 13 min at 22 cm distance. The TBS was removed and the cells were lysed in 1 ml lysis buffer (50 mM Tris−HCl, pH 8.5, 500 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 20 mM EDTA, 2 mg/ml aprotinin, 10 mM vanadyl ribonucleoside). Cell debris was pelleted by centrifugation (3 min at 13 000 *g*) and the supernatant was stored at −80°C.

To immunoprecipitate the RNA−protein complexes cross-linked by UV irradiation, 1 μl mouse monoclonal ascites fluid was added to 20 μl of a 50% suspension of protein A−Sepharose (Pharmacia) and incubated overnight at 4°C. Protein A−Sepharose beads were spun down (10 s, 13 000 *g*), washed three times with RIPA buffer (50 mM Tris−HCl, pH 8.5, 500 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 20 mM EDTA) and added to cell lysates or to the RNA−protein mixture, followed by an overnight incubation at 4°C with end-over-end rotation. Then the protein A-Sepharose beads were washed three times with RIPA buffer and twice with RNase T1 buffer (10 mM Tris−HCl, pH 7.5, 1 mM EDTA, 0.1% Triton X-100). RNase T1 (Boehringer; 50 U in 20 μl RNase T1 buffer) was added and the RNA was digested by a 30 min incubation at 37°C. RNA not linked to the protein was removed by two washes with RIPA buffer. The protein A−Sepharose beads were washed twice with T4 polynucleotide kinase buffer (50 mM Tris−HCl, pH 8, 1.5 mM spermidine, 5 mM $MgCl_2$, 1 mM DTT, 5% glycerol) and RNA linked to the protein was labelled with [γ-^{32}P]ATP (1.85×10^9 MBq/mmol) and T4 polynucleotide kinase (3.7 MBq and 10 U respectively in 20 μl T4 polynucleotide kinase buffer per assay) for 20 min at 37°C. Unincorporated [γ-^{32}P]ATP was removed by two washes with RIPA buffer. RNA cross-linked to NSP3A was deproteinized by proteinase K treatment and analysed on 20% acrylamide−7 M urea denaturing gels.

RNA synthesis and labelling

The 11 full-length rotaviral RNAs were synthesized from purified viral single-shell particles as described (Cohen, 1977). The 71Cat RNA was obtained by *in vitro* transcription of p71Cat (Poncet *et al.*, 1993) linearized by *Ksp*I or *Bgl*II with T7 RNA polymerase using standard protocols (Sambrook *et al.*, 1989). Automated synthesis of oligoribonucleotides via dimethoxytrityl cyanoethyl RNA phosphoramidite chemistry was performed using an Applied Biosystems 380A DNA synthesizer. Cleavage and deprotection were carried out according to the supplier's (Applied Biosystems) instructions. Oligonucleotides were desalted by column chromatography on G10 Sephadex. RNA concentration was determined by UV absorption using the formula; A_{260} = (A×15.4) + (C×7.3) + (G×11.7) + (U×10). For the 20 base long oligonucleotide, an additional purification step on a denaturing gel was added.

Purified oligoribonucleotides used as probes for UV cross-linking and gel retardation were labelled *in vitro* with T4 polynucleotide kinase (Amersham; 10 units) and [γ-^{32}P]ATP (3.7 MBq). To control the quality of the probes, an aliquot was loaded on a 25% polyacrylamide−7 M urea gel and autoradiographed. The efficiency of labelling was assessed by counting the radioactivity present in the gel slice corresponding to the full-length labelled oligoribonucleotide. Specific activities between 0.5 and 1.5×10^6 c.p.m./pmol were obtained. In some case, oligonucleotides were labelled at their 3′ end by use of RNA ligase and [^{32}P]pCp as described (Sambrook *et al.*, 1989).

RNA binding in vitro

Semi-purified protein (175 nM) or cellular extracts (5 μl) were incubated with RNA (1 μg unlabelled RNA or 2.5 nM of 5′ end labelled oligonucleotide) in 20 μl RNA binding buffer (10 mM HEPES, pH 7.9, 3 mM $MgCl_2$, 40 mM KCl, 1 mM DTT, 5% glycerol) for 30 min at 30°C and then subjected to cross-linking (10 min at 4 cm from a germicidal UV bulb) followed by SDS−PAGE (after addition of protein dye with or without 2-mercaptoethanol) or immunoprecipitation (see above).

For gel retardation assays, the RNA binding buffer also contained traces of bromophenol blue. These reactions were conducted at 30°C

and then the products directly loaded onto a non-denaturing polyacrylamide gel. Conditions for electrophoresis for the gel retardation assay were as follows. Non-denaturing polyacrylamide gels (8% acrylamide−0.2% bisacrylamide in 27 mM Tris, pH 7.9, 4 mM EDTA, 13.2 mM Na acetate, pH 7.9, 10% glycerol) were pre-electrophoresed at 4°C at a constant current of 25 mA with recirculation of the low ionic strength buffer (27 mM Tris, pH 7.9, 4 mM EDTA, 13.2 mM Na acetate, pH 7.9). Samples were loaded with the current on. When the bromophenol blue had entered 1 cm into the gel, the current was raised to 50 mA and electrophoresis continued until the dye front had migrated to the middle of the gel (2–3 h). For qualitative studies, gels were directly autoradiographed (X-ray film, −80°C, plus intensifying screen). For quantitation, the gels were dried on DE52 paper before autoradiography and the bands were cut out from the dried gel and counted in a scintillation spectrometer. For competition experiments, serial dilutions of cold competitor RNA were mixed with the labelled probe before addition of protein and buffer mix.

Acknowledgements

We are indebted to Dr M.K.Estes for critical reading of the manuscript. We thank D.Mabon for technical assistance and Dr M.Brémont for the gift of the group C rotavirus dsRNA.

References

Aponte,C., Mattion,N.M., Estes,M.K, Charpilienne,A. and Cohen,J. (1993) *Arch. Virol.*, **133**, 85–95.
Bardwell,V.J., Wickens,M., Bienroth,S., Keller,W., Sproat,B.S. and Lamond,A.I. (1991) *Cell*, **65**, 125–133.
Birney,E., Kumar,S. and Krainer,A.R. (1993) *Nucleic Acids Res.*, **21**, 5803–5816.
Brémont,M., Chabanne-Vautherot,D., Vannier,P., McCrae,M.A. and Cohen,J. (1990) *Virology*, **178**, 579–583.
Budowsky,E.I. and Abdurashidova,G.G. (1989) *Prog. Nucleic Acid Res. Mol. Biol.*, **37**, 1–63.
Chastain,M. and Tinoco,I.,Jr (1991) *Prog. Nucleic Acid Res. Mol. Biol.*, **41**, 131–173.
Clemens,K.R., Wolf,V., McBryant,S.J., Zhang,P., Liao,X., Wright,P.E. and Gottesfeld,J.M. (1993) *Science*, **260**, 530–533.
Cohen,J. (1977) *J. Gen. Virol.*, **36**, 395–402.
Gorziglia,M.I. and Collins,P.L. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 5784–5788.
Hamy,F., Asseline,U., Grasby,J., Iwai,S., Pritchard,C., Slim,G., Butler,P.J.G., Karn,J. and Gait,M.J. (1993) *J. Mol. Biol.*, **230**, 111–123.
Hua,J., Mansell,E.A. and Patton,J.T. (1993) *Virology*, **196**, 372–378.
Keller,W., Bienroth,S., Lang,K.M. and Christofori,G. (1991) *EMBO J.*, **10**, 4241–4249.
Kenan,D.J., Query,S.C. and Keene,J.D. (1991) *Trends Biochem. Sci.*, **16**, 214–220.
Kim,B. and Little,J.W. (1992) *Science*, **255**, 203–206.
Lazinski,D., Grzadzielska,E. and Das,A. (1989) *Cell*, **59**, 207–218.
Mattaj,I.W. (1993) *Cell*, **73**, 837–840.
Mattion,N.M., Aponte,C., Cohen,J. and Estes,M.K. (1992) *Virology*, **190**, 68–83.
Mattion,N.M., Cohen,J. and Estes,M.K. (1994) In Kapikian,A.Z. (ed.), *Viral Infection of the Gastrointestinal Tract.* pp. 169–251.
Olsen,H.S., Nelbock,P., Cochrane,A.W. and Rosen,C.A. (1990) *Science*, **247**, 845–848.
Pashev,I.G., Dimitrov,S.I. and Dimitar,A. (1991) *Trends Biochem. Sci.*, **16**, 323–326.
Patton,J.T. and Gallegos,C. (1990) *J. Gen. Virol.*, **71**, 1087–1094.
Patton,J.T., Salter-Cid,L., Kalbach,A., Mansell,E.A. and Kattoura,M. (1993) *Virology*, **192**, 438–446.
Petrie,B.L., Greenberg,H.B., Graham,D.Y. and Estes,M.K. (1984) *Virus Res.*, **1**, 133–152.
Poncet,D., Aponte,C. and Cohen,J. (1993) *J. Virol.*, **67**, 3159–3165.
Ptashne,M. (1988) *Nature*, **335**, 683–689.
Puglisi,J.D., Tan,R., Calnan,B.J., Frankel,A.D. and Williamson,J.R. (1992) *Science*, **257**, 76–80.
Qian,Y.A., Jiang,B.M., Saif,L.J., Kang,S.Y., Ishimaru,Y., Yamashita,Y., Oseto,M. and Green,K.Y. (1991a) *Virology*, **182**, 562–569.
Qian,Y.A., Jiang,B.M., Saif,L.J., Kang,S.Y., Ojeh,C.K. and Green,K.Y. (1991b) *Virology*, **184**, 752–757.
Ramig,R.F. and Ward,R.L.R (1991) *Adv. Virus Res.*, **39**, 163–207.
Richardson,M.A. and Furuichi,Y. (1985) *J. Virol.*, **56**, 527–533.
Sachs,A. and Wahle,E. (1993) *J. Biol. Chem.*, **268**, 22955–22958.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual.* 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Senear,D.F. and Brenowitz,M. (1991) *J. Biol. Chem.*, **266**, 13661–13671.
Siomi,H., Matunis,M.J., Matthew,W. and Dreyfuss,G. (1993) *Nucleic Acids Res.*, **21**, 1193–1198.
Stamatos,N.M. and Gomatos,P.J. (1982) *Proc. Natl Acad. Sci. USA*, **79**, 3457–3461.
Stefano,J.E. (1984) *Cell*, **36**, 145–154.
Steitz,T.A. (1990) *Q. Rev. Biophys.*, **23**, 205–280.
van Staden,V., Theron,J., Greyling,B.J., Huismans,H. and Nel,L.H.A. (1991) *Virology*, **185**, 500–504.
Yolken,R., Arango-Jaramillo,S., Eiden,J. and Vonderfecht,S. (1988) *J. Infect. Dis.*, **158**, 1120–1123.
Zamore,P.D., Patton,J.G. and Green,M.R. (1992) *Nature*, **355**, 609–614.
Zhao,Y., Thomas,C., Bremer,C. and Roy,P. (1994) *J. Virol.*, **68**, 2179–2185.

Received on May 3, 1994; revised on June 10, 1994