Nucleic Acids Research, 1996, Vol. 24, No. 11 **2133–2142**

Structure–infectivity analysis of the human rhinovirus genomic RNA 3′ non-coding region

Stephen Todd and Bert L. Semler*

Department of Microbiology and Molecular Genetics, College of Medicine, University of California, Irvine, CA 92717, USA

Received January 24, 1996; Revised and Accepted April 10, 1996

ABSTRACT

The specific recognition of genomic positive strand RNAs as templates for the synthesis of intermediate negative strands by the picornavirus replication machinery is presumably mediated by *cis*-acting sequences within the genomic RNA 3′ non-coding region (NCR). A structure–infectivity analysis was conducted on the 44 nt human rhinovirus 14 (HRV14) 3′ NCR to identify the primary sequence and/or secondary structure determinants required for viral replication. Using biochemical RNA secondary structure probing techniques, we have demonstrated the existence of a single stem–loop structure contained entirely within the 3′ NCR, which appears to be phylogenetically conserved within the rhinovirus genus. We also report the *in vivo* analysis of a number of 3′ NCR deletion mutations engineered into infectious cDNA clones which were designed to disrupt the stem–loop secondary structure to varying degrees. Large deletions (up to 37 nt) resulted in defective growth phenotypes, although they were not lethal. We propose that the absolute requirements for initiation of negative strand synthesis are less stringent than previously postulated, even though defined RNA secondary structure determinants may have evolved to facilitate and/or regulate the process of viral RNA replication.

INTRODUCTION

RNA–protein interactions are ubiquitous in nature and direct numerous cellular processes intimately involved in the regulation of gene expression. The ultimate expression of a eukaryotic gene product requires not only ribosome recognition and utilization of a suitable mRNA template, but also pre-mRNA splicing as well as mRNA capping, transport, stabilization and eventual degradation in a cell. Positive strand (mRNA sense) RNA viruses utilize much of the host cell machinery to express viral-encoded gene products required to complete the virus life cycle. In addition to some of the RNA–protein interactions characteristic of cellular messages, most RNA viruses must also maintain a mechanism for the specific replication of the virus genome in the presence of an abundance of cellular cytoplasmic RNAs. In the case of picornaviruses, an important *cis*-acting molecular genetic determinant for this recognition process is believed to reside in the 3′ non-coding region (3′ NCR) of positive strand genomic RNA. Primary sequence determinants and/or secondary structure motifs in this region, in the context of the polyadenylated RNA molecule, are presumably recognized by the viral RNA replication complex to initiate the synthesis of negative strand RNA intermediates (1,2).

The prototypic member of the *Picornaviridae*, poliovirus type 1 (PV1), has a 3′ NCR of 72 nt immediately downstream of the two stop codons at the end of the polyprotein coding region (3,4). Computer-generated RNA secondary structure predictions suggest the existence of a pseudoknot structure in the PV1 3′ NCR (5), which has been partially biochemically confirmed (6). The genomic RNA of human rhinovirus type 14 (HRV14), a closely related picornavirus, contains a 3′ NCR that is 44 nt in length and follows a single stop codon at the end of the polyprotein coding sequence (7). The small size of the HRV14 3′ NCR makes it a particularly attractive target for genetic manipulation in order to identify the molecular features of this region of RNA required for replication complex recognition to initiate the synthesis of viral negative strand RNAs. In addition, an understanding of these features may ultimately aid in the design of antiviral strategies against HRV, a major causitive agent of the common cold.

In the following study, we have investigated the RNA secondary structure of the HRV14 3′ NCR and attempted to identify the primary sequence and/or secondary structure determinants in the region required for viral RNA replication. Using biochemical RNA secondary structure probing techniques, we have demonstrated the existence of a single stem–loop structure, similar to that predicted by Pilipenko and colleagues (5). This single stem–loop structural motif appears to be conserved among different members of the rhinovirus genus based on phylogenetic comparison (A.C.Palmenberg, personal communication) and computer-predicted RNA secondary structure determination (8). We also report the *in vivo* analysis of a number of 3′ NCR deletion mutations engineered into infectious cDNA clones. Surprisingly, large deletions are tolerated within this region of RNA, although they result in defective growth phenotypes. The largest engineered deletion eliminates 37 nt of the HRV14 44 nt 3′ NCR, eliminating the possibility of formation of any higher order RNA structure resembling the wild-type stem–loop. These findings suggest that while the rhinoviruses have evolved a highly conserved, predicted stem–loop in the 3′ NCR of their genomic RNAs which may affect the efficiency of template utilization by the viral replication complex, the absolute requirements for viral RNA replication are much less stringent than previously proposed. Our data support a mechanism for picornavirus negative strand RNA replication initiation which utilizes functions that do not absolutely

* To whom correspondence should be addressed

require a highly specific *cis*-acting RNA recognition determinant at the 3′-end of virus genomic positive strand RNAs.

MATERIALS AND METHODS

Mutagenesis of HRV14 viral cDNA constructs

The construction of a nested series of 3′ co-terminal, subgenomic, T7-based transcription vectors using the full-length HRV14 cDNA construct pT7RV(F.L.) has been described, as well as the construction of an 8 nt deletion mutation in the cDNA corresponding to the HRV14 3′ NCR (2). Transcription vectors containing larger deletions in the cDNA sequence corresponding to the HRV14 RNA 3′ NCR (i.e. 18 and 21 nt) were constructed in a similar manner (9) using the previously described synthetic oligonucleotide, RV10Δ(–) (5′-TGTTAACCTAAAAGAGGTCC-3′) and an additional oligonucleotide, RV14Δ(+) (5′-GAGTAGAAGTAGGAGTTTAT-3′). A mutagenesis cassette was engineered into the HRV14 3′ NCR cDNA sequence by site-directed mutagenesis at nt 7196 using the heteroduplex method (10) with the mutagenesis oligonucleotide RVU7196A/G (5′-CACTTAATTTGAGRAGAAGTAGG-3′, where R is A or G). The existing sequence, 5′-GAGTAG-3′, in the wild-type cDNA was changed to 5′-GAGGAG-3′ to create the recognition sequence of the restriction endonuclease *Bse*RI (5′-GAGGAG$(N)_{10/8}$-3′). Digestion of the resulting plasmid with *Hpa*I (7169) and *Bse*RI (7193), followed by repair with the Klenow enzyme, religation and transformation into *Escherichia coli* C600 cells resulted in the elimination of a 37 bp fragment from the 3′ NCR cDNA sequence. Mutations in subgenomic cDNA constructs were subsequently cloned into pT7-HRV14(ST), a reconstructed full-length HRV14 cDNA clone (S. Todd and B. L. Semler, unpublished results). The resulting cDNA sequences of these mutations are described in Results. All plasmids were sequenced using the modified T7 DNA polymerase.

In vitro RNA synthesis

Synthesis of non-radiolabeled RNAs was performed using the MEGAshortscript kit (Ambion) with ~4 µg of the appropriate *Pst*I- or *Cla*I-linearized plasmid DNA templates. The RNA was phenol/chloroform extracted and ethanol precipitated. Aliquots (0.5 µg/µl) were then stored at –70°C.

5′-End-labeling of oligonucleotides

Approximately 120 pmol of oligonucleotide RVoligoT+9(–) were incubated in the presence of 100 µCi 6000 Ci/mmol [γ-^{32}P]ATP and 60–100 U T4 polynucleotide kinase in 100 µl for 45 min at 37°C. Unincorporated radiolabeled nucleotides were removed using a Sephadex G-50 spin column. The resulting 5′-end-labeled oligonucleotides had a specific activity of ~3×10^5 c.p.m./pmol.

RNA secondary structure probing by primer extension

RNA secondary structure probing was performed using conditions modified from published methods (6,11,12). Briefly, 0.5 µg HRV14 3′ NCR-specific RNA was incubated in the presence of 40 µg *E.coli* tRNA in 0.7× TMK buffer (30 mM Tris–HCl, pH 7.4, 10 mM $MgCl_2$, 270 mM KCl) or in TM(–K) buffer supplemented with different amounts of KCl or NaCl under reducing conditions (18 mM 2-mercaptoethanol), in a total volume of 40 µl. The RNA was then incubated successively at 68°C and 37°C and room temperature for 5 min each to allow RNA secondary structure to form. Enzymatic treatment with RNase V_1, T_1, A, U_2 or an RNase isolated from *Baccillus cereus* was performed at room temperature for 5 min or on ice as detailed in Results. The enzyme reactions were stopped by the addition of 155 µl enzyme stop solution (0.3 M NaOAc, 10 mM EDTA, 0.3% SDS) followed by phenol/chloroform extraction and precipitation with 2.5 vol. ethanol. Primer extension was performed essentially as described in Eisenberg *et al.* (13) using 2–4 pmol (~10^6 c.p.m.) 5′-end-labeled oligonucleotide RVoligoT+9(–) and 10 U AMV reverse transcriptase (Life Sciences) at 50°C for 20–30 min in primer extension buffer (20 mM Tris–HCl, pH 7.4, 10 mM $MgCl_2$, 6 mM DTT, 300 µM each dNTP, 80 µg/ml actinomycin D). The extension reactions were phenol/chloroform extracted, ethanol precipitated, resuspended in a formamide loading buffer and resolved on 8% polyacrylamide–7 M urea gels. Sequencing ladders were generated from the appropriate transcription vector using the modified T7 DNA polymerase with 5′-end-labeled oligonucleotide RVoligoT+9(–) as a primer. The gels were then dried and exposed to Kodak AR or MR X-ray film.

Direct RNA secondary structure probing

Direct RNA secondary structure probing was performed using *in vitro* transcribed RNA which was dephosphorylated with intestinal alkaline phosphatase and 5′-end-labeled with T4 polynucleotide kinase in the presence of 50–100 µCi 6000 Ci/mmol [γ-^{32}P]ATP. Unincorporated radiolabeled nucleotides were removed using a Sephadex G-50 spin column in TE8 buffer (10 mM Tris, 1 mM EDTA, pH 8.0) with 0.1% SDS followed by a final ethanol precipitation in 2.5 M ammonium acetate. Enzymatic treatment of ~0.5 µg RNA was performed as described above, also in the presence of 40 µg/reaction *E.coli* tRNA and 18 mM 2-mercaptoethanol. Reactions were stopped with enzyme stop buffer, phenol/chloroform extracted, ethanol precipitated, resuspended in formamide or 8 M urea loading dye and analyzed as described above.

RNA transfection, virus propagation and sequencing of viral RNA

In vitro transcription of full-length virus-specific RNAs from wild-type and mutated *Pst*I-linearized pT7-HRV14(ST)-based cDNA constructs was performed as described previously (14). DEAE-mediated RNA transfection of R19 HeLa cells and rhinovirus propagation in tissue culture have also been described (2,14). Asymmetric RT-PCR sequencing of viral RNAs from total cytoplasmic RNA harvested from infected monolayers (15) was performed as described in Todd *et al.* (2) using oligonucleotide primer RV7035(+) (5′-GCATGTTAGCATGGCACTCAGG-3′), which is identical to nt 7035–7056 within the polymerase coding sequence of HRV14, along with either RVoligoT+9(–) (5′-TTTTTTTTATAAACTCC-3′) or RVoligoT+2(–) (5′-TTTTTTTTTTTTTTAT-3′), which are complementary to the 3′-end of virus positive strand RNAs.

RESULTS

Computer predictions of the HRV14 3′ NCR secondary structure

The computer-predicted RNA secondary structures of two HRV14 3′-end-specific RNA sequences using the Zuker Fold-RNA algorithm (16) are shown in Figure 1. In order to allow the

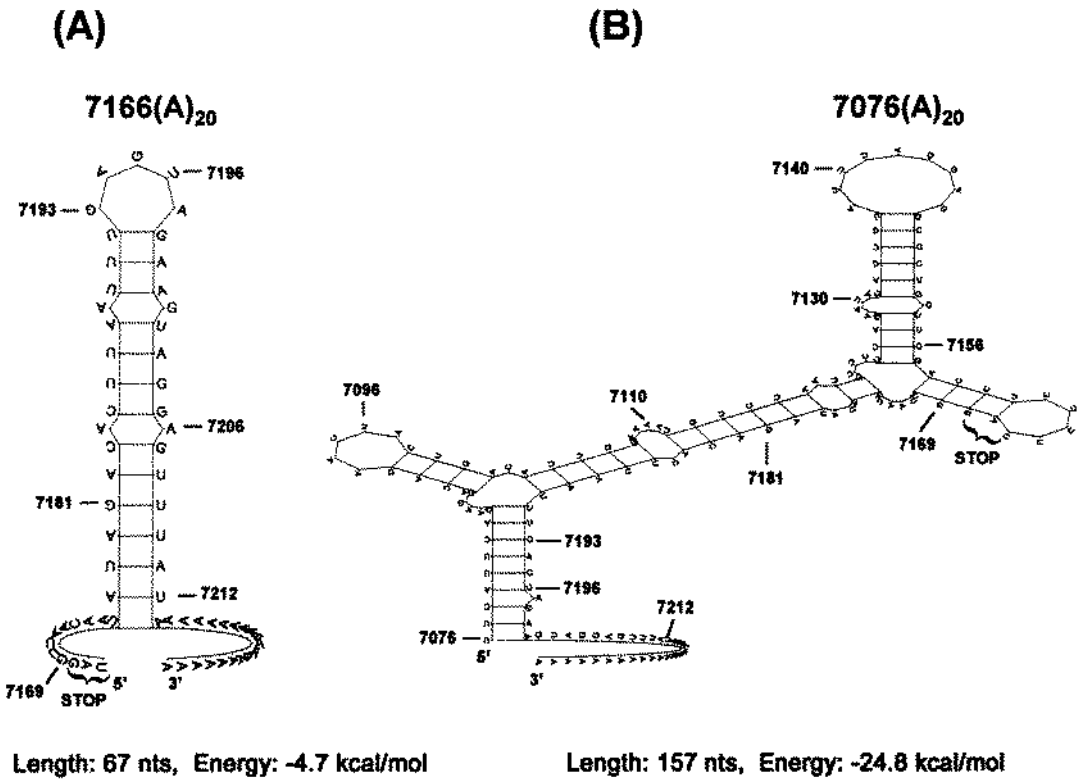

Figure 1. Computer-predicted RNA secondary structures of the (**A**) 7166A+ RNA and (**B**) 7076A+ RNA using the University of Wisconsin FoldRNA package and Squiggles output with no additional constraints. The sequences included a 20 nt poly(A)$_n$ tract to simulate authentic viral RNA molecules.

direct comparison of secondary structures predicted using the computer algorithms with subsequent biochemical probing data, the sequence of the RNAs initially folded by computer were similar to those which could be synthesized from available transcription vectors *in vitro*. *Pst*I-linearized, T7-based plasmids pT7RV7168A+, pT7RV7136A+, pT7RV7076A+ and pT7RV6338A+ served as transcription templates to generate RNAs which contain the HRV14 3′ NCR with poly(A)$_{60–80}$ and 0, 32, 98 or 831 nt of 5′-proximal (3D^{pol}) RNA sequence, respectively (2). RNAs with no sequences upstream of the 3′ NCR or 36 nt of 3D^{pol} coding sequence give rise to the type of secondary structure shown in Figure 1A, which contains a single predominant stem–loop structure within the 3′ NCR. The stop codon lies outside the stem structure (nt 7166–7168). Larger RNAs, with 98 or 831 nt of 5′-proximal sequence, give rise to a different computer-predicted secondary structure, in which the stop codon lies outside a different helical region, as shown in Figure 1B. The predicted free energy change (ΔG) for the formation of the RNA secondary structure shown in Figure 1A is –4.7 kcal/mol. The computer-predicted ΔG value using longer RNAs is more favorable (–10.4 kcal/mol for the 7136A+ RNA), although the 3′ NCR stem–loop structure is unchanged. These modest $-\Delta G$ values made the computer predictions for RNA secondary structure unconvincing by themselves, prompting the biochemical investigation of the RNA higher order structure. The Zuker MFold algorithm (8), which is used for subsequent computer-generated RNA secondary predictions, predicts an RNA secondary structure for the 7076A+ RNA which contains a helical region corresponding to the stem–loop shown in Figure 1A.

Secondary structure probing of HRV14 3′ NCR-containing RNAs by primer extension

The 3′ NCR of HRV14 was subjected to biochemical analysis by enzymatic treatment followed by primer extension analysis of the resulting RNA cleavage products. The enzymes (and their specificities) used in this study included RNase V_1 (dsRNA), T_1 (Gp↓N), A (Cp↓N, Up↓N), U_2 (Ap↓N) and an RNase isolated from *B.cereus* (herein refered to as *Bc*) (Cp↓N, Up↓N). Figure 2A shows the results of secondary structure probing with all the above listed enzymes on 7076A+ RNA generated from the pT7RV7076A+ transcription vector. The reactions were performed either at room temperature for 5 min (odd numbered lanes) or on ice for 20 min (even numbered lanes). The 5′-end-labeled 3′ NCR-specific deoxyoligonucleotide primer RVoligoT+9(–) was used for both primer extension analysis of the resulting RNA fragments as well as for generating the accompanying sequencing ladder by dideoxy sequencing of the pT7RV7076A+ plasmid DNA (lanes 1–4). As a result of the different mechanisms of cDNA termination (the incorporation of a dideoxynucleotide nucleotide analog compared with template scission), the reverse transcribed primer extension products are 1 nt shorter than the corresponding cDNA product of the sequencing ladder [e.g. the primer extension product resulting from cleavage after G(7193) by RNase T_1 co-migrates with position 7194 of the sequencing ladder].

An RNase V_1-sensitive region of the 7076A+ RNA was detected between nucleotides 7184 and 7192 (inclusive). The sequences immediately upstream of nucleotide 7184 (also

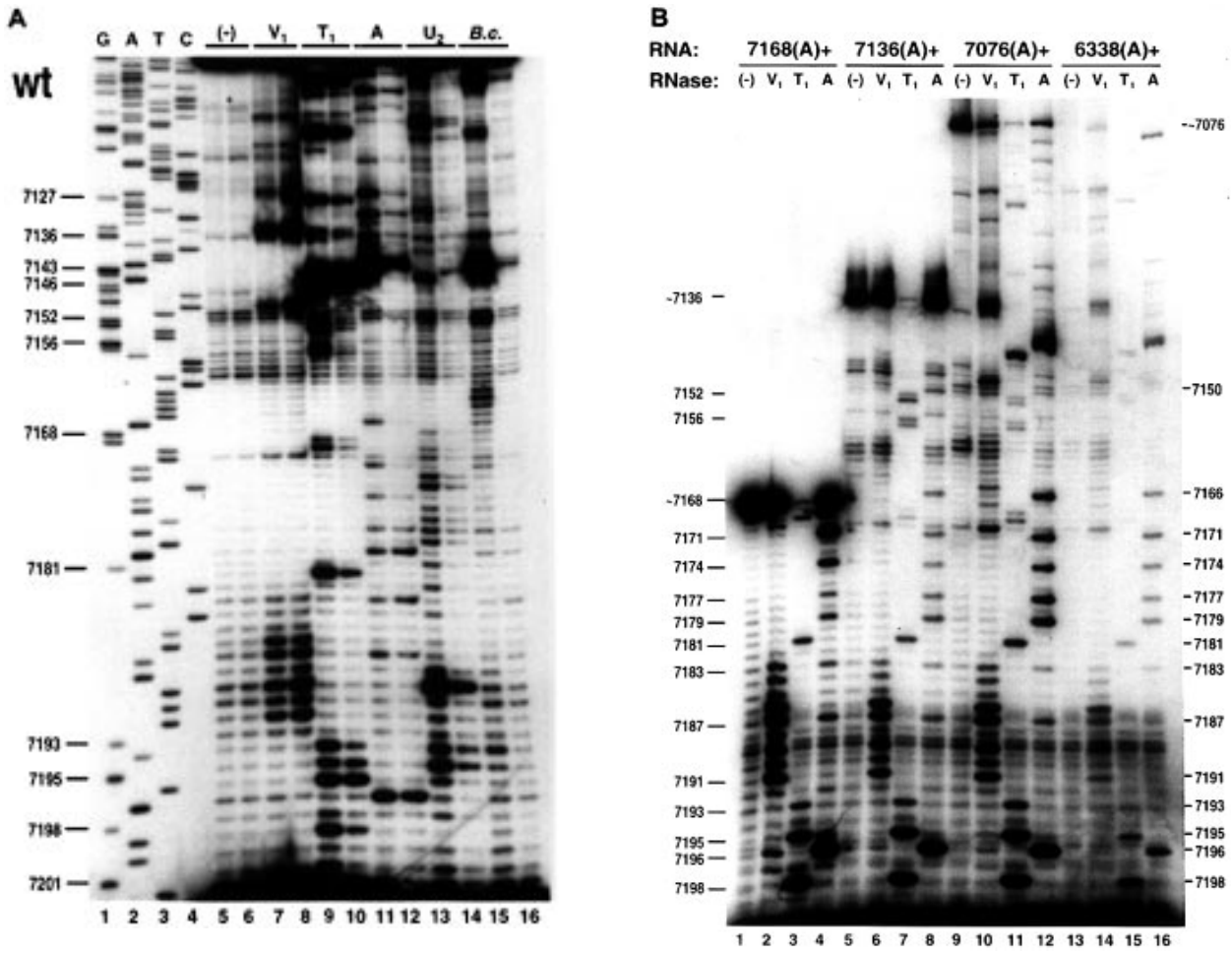

Figure 2. Enzymatic secondary structure probing of the wild-type HRV14 3′ NCR by primer extension analysis. (**A**) Approximately 0.5 µg 7076A+ RNA and 40 µg tRNA were either mock incubated (lanes 5 and 6) or incubated in the presence of 0.7 U RNase V_1 (lanes 7 and 8), 20 U RNase T_1 (lanes 9 and 10), 1 µg RNase A (lanes 11 and 12), 10 U RNase U_2 (lanes 13 and 14) or 10 U RNase *B.cereus* (lanes 15 and 16) at room temperature for 5 min (lanes 5, 7, 9, 11, 13 and 15) or on ice (lanes 6, 8, 10, 12, 14 and 16) for 15 min in 0.7× TMK buffer (30 mM Tris–HCl, 10 mM $MgCl_2$, 270 mM KCl, pH 7.4) supplemented with 18 mM 2-mercaptoethanol. Total reaction volumes were 40 µl. Primer extension was performed using ~10^6 c.p.m. [γ-^{32}P]ATP 5′-end-labeled oligonucleotide RVoligoT+9(–) (5′-TTTTTTTTATAAACTCC-3′) in the presence of AMV reverse transcriptase. The reactions were then phenol/chloroform extracted, ethanol precipitated, resuspended in formamide loading buffer and analyzed on an 8% polyacrylamide–7 M urea gel. The sequencing ladder (lanes 1–4) was generated from the pT7RV7076A+ plasmid using 5′-end-labeled oligonucleotide RVoligoT+9(–) with the modified T7 DNA polymerase. (**B**) RNAs 7168A+ (lanes 1–4), 7136A+ (lanes 5–8), 7076A+ (lanes 9–12) and 6338A+ (lanes 13–16) (containing 0, 32, 98 or 831 nt of 5′-proximal $3D^{pol}$ coding sequence, respectively) were either mock treated (lanes 1, 5, 9 and 13) or treated with 0.7 U RNase V_1 (lanes 2, 6, 10 and 14), 20 U RNase T_1 (lanes 3, 7, 11 and 15) or 1 µg RNase A (lanes 4, 8, 12 and 16) at room temperature for 5 min as described in (A) above.

predicted to be involved in base pairing) were not sensitive to RNase V_1 (lanes 7 and 8). Efficient RNase T_1-mediated cleavages at G(7193) and G(7195) were consistent with the existence of the computer-predicted distal stem–loop structure (lanes 9 and 10); there was even nuclease sensitivity at A(7194). The G at the base of the predicted loop region (7198) was also RNase T_1 sensitive, consistent with limited duplex breathing (note that cleavage at this site was reduced at the lower temperature; lane 10). The bulge G(7201) computer-predicted for the 7166A+ RNA (Fig. 1A) was also cleaved by RNase T_1, although this nucleotide is in close proximity to the primer and may be over-represented among the various primer extension products (see below). Although computer-predicted to lie within a helical region, G(7181) was clearly RNase T_1 sensitive at either temperature, suggesting that the lower portion of the stem structure either does not form, is in equilibrium with a single-stranded structure or also contains a non-paired nucleotide. Guanylate residues at 7168 and 7169 were only slightly RNase T_1 sensitive, suggesting that they are involved in base pairing, while G(7152), G(7153), G(7156) and G(7157) are apparently in an unpaired region of the structure.

Efficient cleavage by RNase A at U(7196) (lanes 11, 12) and RNase U_2 at A(7194) (lanes 13, 14) further supported the existence of a single-stranded loop in this region of the 3′ NCR. Other less pronounced RNase A-sensitive nucleotides were clustered at the base of the computer-predicted 3′ NCR stem–loop and the single-stranded region immediately downstream of the stop codon [U(7171), C(7174), U(7177) and U(7179)]. An intense signal resulting from RNase U_2 cleavage at A(7189), within the RNase V_1-sensitive helical region, supported the computer prediction of a bulge A at that position (Fig. 1A). Under conditions of partial digestion, RNase A cleaves efficiently only

after the ultimate 3′-nucleotide of a pyrimidine string sequence (17). This fact was illustrated by cleavage at U(7166), which is the last pyrimidine in a string (5′-CCUCUUUU-3′, nt 7159–7166; lanes 11 and 12). RNase *Bc* (which shares sequence specificity with RNase A) also cleaves efficiently within pyrimidine tracts (17) and revealed the predominantly single-stranded nature of the pyrimidine tract in the proximal computer-predicted stem–loop structure (Fig. 2A, lanes 15 and 16; refer to Fig. 5, below).

In order to demonstrate that the RNA secondary structure observed for the 3′ NCR contained within the 7076A+ RNA was not an artifact of the particular 7076A+ RNA sequence, 3′ NCR-containing RNAs of various lengths were used for secondary structure probing. Figure 2B shows the use of RNAs 7168A+ (lanes 1–4), 7136A+ (lanes 5–8), 7076A+ (lanes 9–12) and 6338A+ (lanes 13–16), which contain 0, 36, 98 or 831 nt of 5′-proximal $3D^{pol}$ coding sequence, respectively, for RNA secondary structure probing using RNases V_1, T_1 and A. The same overall secondary structure was observed downstream of the stop codon (nt 7166–7168) regardless of the length of 5′-proximal sequence present in the RNA. The RNase V_1-sensitive nucleotides (7184–7192) in the helical region (lanes 2, 6, 10 and 14) and RNase T_1-hypersensitive nucleotides [G(7193), G(7195) and G(7198)] in the unpaired loop region (lanes 3, 7, 11 and 15) were readily apparent. These results indicated that formation of the 3′ NCR stem–loop is not dependent upon upstream RNA sequences and justify the use of various length RNA templates for the experiments described below.

Direct RNA secondary structure probing and primer extension analysis of HRV14 3′ NCR-containing RNAs lacking poly(A)$_n$ tracts

Direct RNA secondary structure probing using γ-^{32}P-5′-end-labeled HRV14 3′ NCR-specific RNA was employed to confirm the results obtained using the primer extension method as well as to assess the role of the 3′-terminal 9 nt of the 3′ NCR and the poly(A)$_n$ tract in the formation of RNA secondary structure. Since the relative distance of the RNA sequence of interest from the radiolabeled 5′-terminal nucleotide of the template RNA determines the number of nucleotides to be resolved by polyacrylamide gel electrophoresis, shorter RNAs were used for direct structure probing than those employed in the primer extension experiments shown in Figure 2. Figure 3A shows the results of such probing using 5′-end-labeled 7136A+ and 7136A– RNAs. The 7136A– RNA lacks a poly(A) tract but contains the intact HRV14 3′ NCR and terminates with two non-viral nucleotides (-CG-3′) acquired from the *Cla*I restriction site used to linearize the pT7RV7136A– transcription vector.

The susceptibilities of specific nucleotide positions to cleavage by RNases V_1, T_1, A, U_2 and *Bc* are fundamentally consistent with results obtained using the primer extension method. The direct probing method confirms the existence of a helical region from nt 7184 to 7192 (Fig. 3A, lanes 2 and 9, boxed). The G residues at 7193 and 7195 are RNase T_1 hypersensitive (also boxed), while G(7198) and G(7201) are less susceptible to T_1 cleavage (lanes 3, 4, 10 and 11). In some experiments (see Fig. 5, lanes 3 and 4, below), G(7201) is more susceptible to RNase T_1 cleavage than G(7198) and G(7204/5), consistent with a secondary structure model in which G(7201) is bulged opposite A(7189). RNase U_2 treatment of 5′-end-labeled 7136A+ clearly shows bulged A(7189) (lane 6 and 13), as was also shown by primer extension (Fig. 2A, lanes 13 and 14). Neither A(7184) nor A(7206) are RNase U_2 hypersensitive, although no obvious possibility exists for base pairing either of these nucleotides within the stem structure.

There were no remarkable differences in the pattern of digestion for the 7136A– RNA compared with the 7136A+ RNA, suggesting that the poly(A)$_n$ tract is not involved in the RNA secondary structure with the 3′ NCR. The direct probing method also reveals the existence of a RNase V_1-sensitive region at positions 7208–7210, a string of three urydylate residues separated by two nucleotides (AU) from the poly(A)$_n$ tract (see boxed region of Fig. 3A, lanes 2 and 9). As this V_1-sensitive region also exists in the 7136A– RNA, it cannot represent a snap-back structure involving the poly(A)$_n$ tract. Direct secondary probing of 5′-end-labeled 7168A+ and 7168A– RNA (which contain no 5′-proximal $3D^{pol}$ coding sequence) revealed similar patterns of RNase sensitivity, further indicating that nucleotides 7208–7210 do not interact with sequences 5′-proximal to the stop codon (data not shown).

The results of primer extension secondary structure analysis using the 7136A+ and 7136A– RNAs with 5′-end-labeled oligonucleotide RVoligoT+9(–) are shown in Figure 3B. Only 9 nt of the RVoligoT+9(–) primer hybridize to the 3′-end of the 7136A+ RNA leaving a 2 bp mismatch (non-viral sequence from the engineerend restriction endonuclease site in the transcription vector) and overhanging $(T)_8$ tract; consequently, the exposure time for the autoradiogram showing the primer extension products of the 7136A– digestion is ~10 times longer than for the 7136A+ digestion products. These RNA secondary structure probing results are consistent with the results obtained using radiolabeled RNAs and further suggest that the poly(A)$_n$ tract does not interact with sequences within or upstream of the 3′ NCR.

Effects of salt concentration on the 3′ NCR RNA secondary structure

As described in Materials and Methods, most partial RNase digests were carried out in 200 mM KCl (final concentration) supplied by TMK buffer. To determine the effect of salt concentration on RNA secondary structure, enzymatic probing was carried out in TM(–K) buffer, which lacked KCl, either with no additional salt or supplemented with 75, 150 or 300 mM NaCl or KCl (final concentration). Increasing NaCl concentration resulted in increased helical secondary structure, as demonstrated by RNase V_1 sensitivity between nucleotides 7184 and 7192 and 7208 and 7210, inclusive (data not shown). Numerous nucleotides between positions 7166 and 7192, most notably U(7177), U(7179) and U(7187), became less susceptible to single strand-specific nuclease attack with increased salt concentration, while U(7196), in the proposed distal loop region, remained RNase A-sensitive regardless of the salt concentration. The bulged G(7181) was consistently RNase T_1 sensitive even at higher NaCl concentrations. Similar results were obtained when the reactions were supplemented with KCl instead of NaCl (data not shown). RNA secondary structure probing of 5′-end-labeled 7168A+ and 7168A– (which contain no 5′-proximal $3D^{pol}$ coding sequence) under the identical range of NaCl and KCl concentrations showed the same RNase sensitivity pattern (data not shown), indicating that 5′-proximal nucleotides were not interacting with the distal stem–loop under high salt conditions to alter the observed RNA

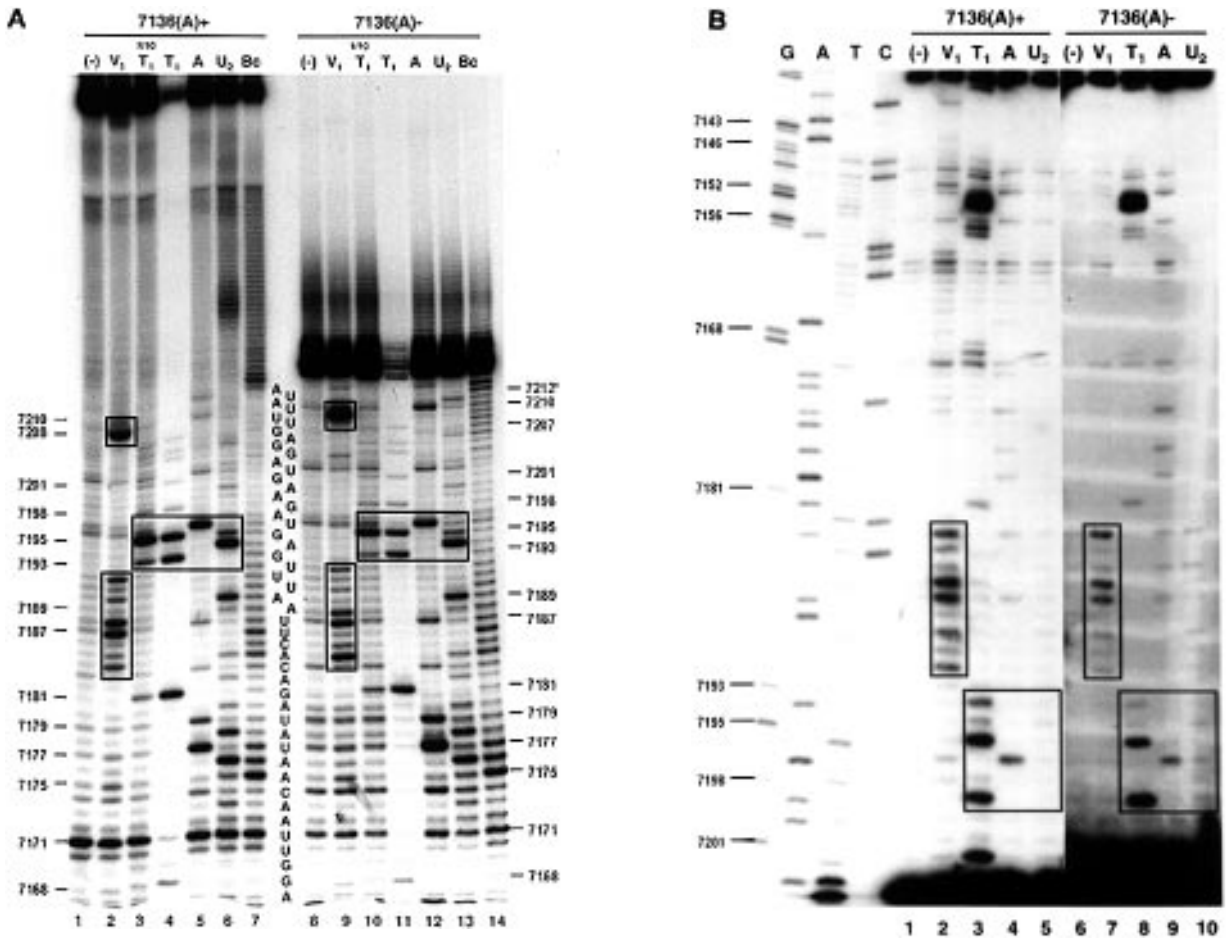

Figure 3. RNAs 7136A+ and 7136A– were used for enzymatic secondary structure probing either by primer extension or by direct structure analysis using [γ-^{32}P]ATP 5′-end-labeled RNA. Enzymatic probing of 5′-end-labeled RNAs was carried out at room temperature for 5 min essentially as described for the primer extension method (see legend to Fig. 2A) except that >10^6 c.p.m. of [γ-^{32}P]ATP 5′-end-labeled RNA was used per reaction. Following enzymatic treatment, the reaction was phenol/chloroform extracted, ethanol precipitated, resuspended in 8 M urea loading dye and analyzed on a 6% polyacrylamide–7 M urea gel. (**A**) [γ-^{32}P]ATP 5′-end-labeled 7136(A)+ RNA (lanes 1–7) or 7136(A)– RNA (lanes 8–14) was mock incubated (lanes 1 and 8) or incubated in the presence of 0.7 U RNase V_1 (lanes 2 and 9), 2 U RNase T_1 (lanes 3 and 10), 20 U RNase T_1 (lanes 4 and 11), 1 µg RNase A (lanes 5 and 12), 10 U RNase U_2 (lanes 6 and 13) or 10 U RNase *B.cereus* (lanes 7 and 14). The atserisk (*) next to position 7212 denotes that the 7136A– RNA terminates with two non-viral nucleotides (5′-CG-3′) derived from the engineered *Cla*I restriction site in the pT7RV7136A– transcription vector. (**B**) Non-radiolabeled 7136A+ RNA (lanes 1–5) or 7136A– RNA (lanes 6–10) was mock incubated (lanes 1 and 6) or incubated in the presence of 0.7 U RNase V_1 (lanes 2 and 7), 20 U RNase T_1 (lanes 3 and 8), 1 µg RNase A (lanes 4 and 9) or 10 U RNase U_2 (lanes 5 and 10). Extension reactions were performed as described in the legend to Figure 2 and the products were analyzed on an 8% polyacrylamide–7 M urea gel. The exposure time for lanes 6–10 was ~10 times that for lanes 1–5 as a result of the reduced primer complementarity of RVoligoT+9(–) for the non-polyadenylated 7136A– RNA.

secondary structure. The predominant stem–loop structure in the 3′ NCR was thus detected reproducibly under moderate to high salt conditions (bracketing physiological intracellular conditions) and was disturbed by extremely low salt conditions.

Primer extension analysis of partially ribonuclease digested 7076A+ RNA was also performed to examine long range intramolecular RNA sequence interactions under different salt conditions (data not shown). An increased helical character of RNA sequences upstream of the 3′ NCR was observed at higher salt consentrations (based upon RNase V_1 sensitivity), which was consistent with some computer-predicted RNA secondary structure models using longer RNA sequences. However, the stem–loops formed appear to be independent structures contained entirely within the $3D^{pol}$ coding region. These short helical RNA structures did not appear to interact with the 3′ NCR, as indicated by the consistent RNA secondary structure probing pattern obtained with RNAs lacking 3′ NCR flanking sequence, and presumably do not affect the architecture of the 44 nt 3′ NCR.

The diagram in Figure 4 shows the RNase susceptibility data for the wild-type HRV14 3′ NCR from the above experiments superimposed on a computer-generated (MFold) RNA secondary structure model (8). Large arrows indicate consistently strong signals on secondary structure probing autoradiograms, while small arrows indicate significant but weaker signals. This model will be described in greater detail in the Discussion.

RNA secondary structure probing of RNAs harboring deletions in the 3′ NCR

A series of deletion mutations was generated in the HRV14 3′ NCR. An 8 nt deletion (Δ8) in the 3′ NCR of HRV14 has been described previously (2). This sequence was selected for

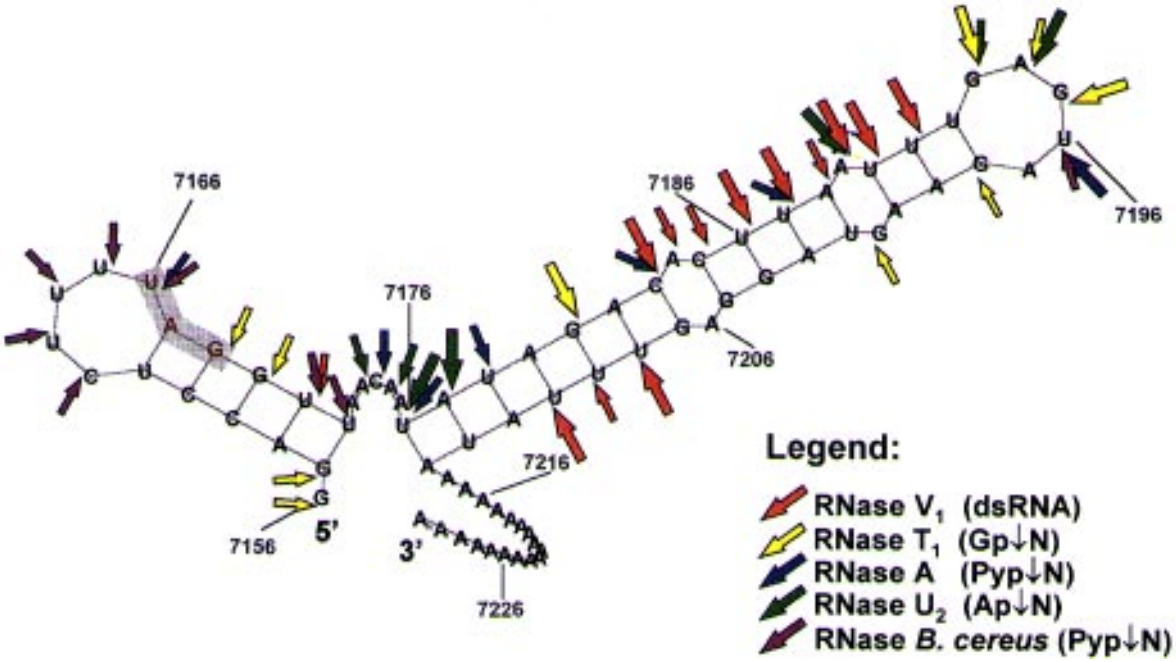

Figure 4. Summary of RNA secondary structure probing results. The results obtained from biochemical RNA secondary probing using primer extension and 5′-end-labeled RNA are shown superimposed on the MFold-generated RNA secondary model of the 7156(A)$_{20}$ RNA at 34°C with no constraints. The output was generated using PlotFold with the Squiggles representation. Large arrows denote positions which were highly sensitive to scission, while smaller arrows indicate less sensitive (but consistently noted) positions for RNase cleavage. This model is discussed in detail in the Discussion.

site-directed mutagenesis based on its conservation in the 3′ NCRs of other rhinovirus genomes (A.C.Palmenberg, personal communication) and its disruptive effect on the computer-predicted secondary structure of the 3′ NCR. The Δ8 virus showed delayed onset of RNA replication and reduced accumulation of viral RNA in infected cells, based on RNA slot blot analysis using virus derived from the pT7RV(F.L.)-based cDNA transcription vector (2). In order to further investigate the requirement for an intact stem–loop structure in the HRV14 3′ NCR to support viral replication, larger deletion mutations were engineered which were designed to abolish formation of the distal stem–loop structure. An 18 nt deletion (Δ18; nt 7175–7192) results in the deletion of half the 3′ NCR stem sequence, while a 21 nt deletion (Δ21) results in the additional deletion of most of the loop region (nt 7193–7196). A 37 nt deletion (Δ37; nt 7172–7209) results in the removal of all but 7 nt of the HRV14 3′ NCR (5′-GTTTTAT-3′), excluding the possibility that any elaborate RNA secondary structure can be assumed by the remaining RNA sequence.

Direct RNA secondary structure probing using 5′-end-labeled 7136A+ *in vitro* transcribed wild-type and Δ8 RNA clearly demonstrated the loss of the wild-type stem–loop structure in the Δ8 RNA (Fig. 5A). Probing of the wild-type RNA revealed RNase T_1 sensitivity in the loop region [G(7193) and G(7195)] and relative insensitivity in the stem region [G(7201), G(7204), G(7205) and G(7207)] (lanes 3 and 4). In contrast, the corresponding G residues were uniformly sensitive to RNase T_1 in the Δ8 RNA, demonstrating the loss or instability of the stem–loop structure in the Δ8 3′ NCR (lanes 9 and 10). Uniform digestion patterns for RNases A and U_2 were also consistent with this loss of a defined RNA secondary structure (lanes 11 and 12). The Δ8 RNA showed RNase V_1-sensitive regions between nucleotides 7185 and 7189 and after G(7207), followed by a region which was hypersensitive to RNases A and U_2 immediately upstream of the poly(A)$_n$ tract. While biochemical evidence for a stem–loop structure in the Δ8 RNA is unconvincing, it is somewhat consistent with a suboptimal secondary structure predicted using MFold (data not shown).

RNA sequences also demonstrated region-specific uniform sensitivity to single strand-specific RNases in the Δ18 and Δ21 RNAs. For example, the G residues which were protected from RNase T_1 attack in the wild-type 3′ NCR (7201, 7204, 7205 and 7207), were all susceptible to cleavage in the RNAs harboring Δ18 or Δ21 lesions (Fig. 5B, lanes 9 and 10 and lanes 15 and 16, compared with lanes 3 and 4), arguing against the existence of distinct secondary structure. The Δ21 RNA was generally more sensitive to RNase V_1 than the Δ18 RNA (lanes 8 and 14), suggesting the possible existence of some helical character in the population of RNAs, although closer examination did not identify significant base pairing potential in the Δ21 3′ NCR. Secondary structure probing on 7136A+Δ37 RNA was not performed, since no secondary structure of consequence was expected from the remaining primary sequence of the HRV14 3′ NCR.

Functional *in vivo* analysis of RNAs harboring deletions in the 3′ NCR

The Δ18, Δ21 and Δ37 mutations were constructed into the pT7-HRV14(ST) plasmid background to generate RNAs for transfection into tissue culture cells to examine their abilities to produce infectious rhinovirus. The Δ8 mutation, originally studied in a different infectious cDNA plasmid background (2), was also cloned into a pT7-HRV14(ST)-based transcription vector. Surprisingly, genome-length RNAs bearing these larger

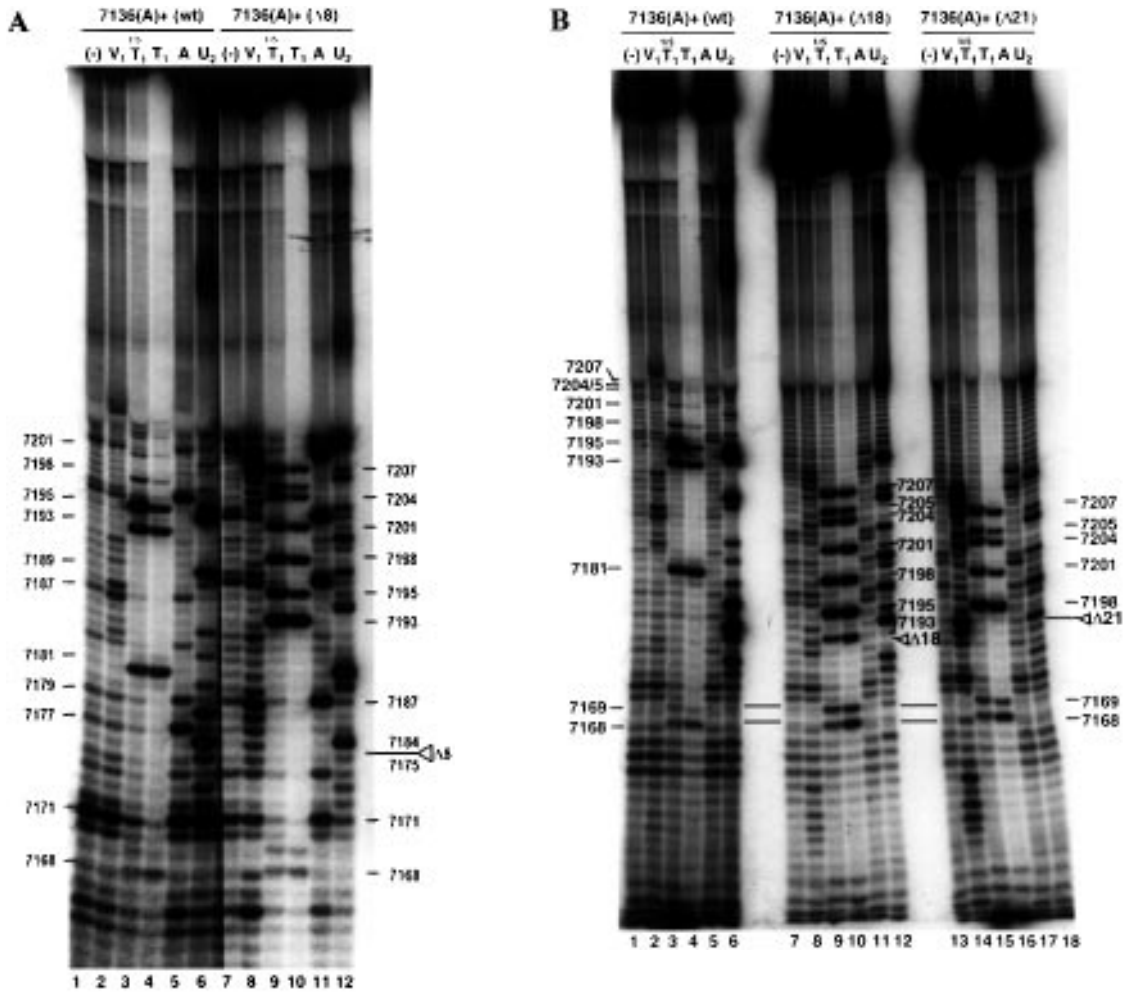

Figure 5. Secondary structure determination of RNAs harboring deletions in the 3′ NCR. Enzymatic probing of wild-type, Δ8, Δ18 and Δ37 5′-end-labeled 7136A+ RNAs was carried out using the same partial RNase digestion conditions as described in the legend to Figure 3. (**A**) Wild-type and Δ8; (**B**) wild-type, Δ18 and Δ21

deletions gave rise to viral progeny with impaired growth characteristics similar to those of the original Δ8 virus. While DEAE-mediated transfection of wild-type HRV14 RNA resulted in complete destruction of a HeLa R19 monolayer in <48 h, mutant HRV14 viruses required 10–11 days to effect complete cell lysis of a monolayer following separate transfections with Δ8, Δ18 or Δ37 RNAs. Wild-type, Δ8, Δ18 and Δ37 virus isolates obtained from liquid overlays were used to infect additional HeLa R19 cell monolayers and total cytoplasmic RNA was harvested from monolayers beginning to show cytopathic effects (CPE) following infection (8, 16, 15 and 13.5 h, respectively). The resulting RNA was then subjected to asymmetric RT-PCR sequencing as described previously (2) using the RV7035(+)/RVoligoT+9(–) or the RV7035(+)/RVoligoT+2(–) primer set. The 3′-end of RVoligoT+9(–) contains 9 nt which are complementary to the wild-type HRV14 3′ NCR and should therefore anneal to wild-type, Δ8 and Δ18 (or Δ21) RNAs but not to Δ37 RNAs, which harbor a deletion extending into this region of complementarity. RVoligoT+2(–) will amplify any RNA sequence terminating with 5′-AT$(A)_n$-3′, including the Δ37 RNAs. The identities of the wild-type, Δ8 and Δ18 viruses were confirmed using the RV7035(+)/RVoligoT+9(–) primer set, while no sequence was obtained from RNA isolated from Δ37-infected monolayers (arguing against the possibility of virus stock contamination; data not shown). Wild-type and Δ37 RNAs were then sequenced following amplification using the RV7035(+)/RVoligoT+2(–) primer set, which confirmed the existence of the Δ37 lesion in the transfection-derived virus. The demonstration of 3′ NCR deletions in virus harvested from cells showing CPE, which correspond to the deletions engineered into the pT7-HTV14(ST)-based transcription vectors, is compelling evidence that the mutated RNAs are capable of being autonomously replicated by the viral RNA replication machinery.

DISCUSSION

We have used RNA secondary structure probing techniques to examine the structure of the wild-type and mutated HRV14 RNA 3′ NCRs in solution. A diagrammatic representation of the wild-type HRV14 3′ NCR secondary structure, based on the data presented here, is shown in Figure 4. The 3′ NCR appears to fold into a single stem–loop structure which does not involve 5′-proximal $3D^{pol}$ coding sequence or the 3′ poly(A) tract. There are several bulged (i.e. non-paired) nucleotides along the length of the stem between nt ~7176 and 7192, notably G(7181) and A(7189). G(7201) was slightly more sensitive than neighboring

G(7198) and G(7204/5), supporting the model in which it is bulged opposite A(7189). The inability to detect the computer-predicted bulge A(7184) or A(7206) by biochemical probing is noteworthy. Despite the inability to detect RNase V_1 cleavage between nucleotides ~7176 and 7182, we believe this region is base paired because of the absence of cleavage by single strand-specific RNases and the series of RNase V_1-sensitive U nucleotides at the 3′-end of the 3′ NCR (nt 7208–7210) (17). Secondary structure probing using shorter RNAs, including the 3′ NCR with only two non-viral nucleotides at the 3′-end (7168A–) indicates that any secondary structure model must account for the base pairing within the stem structure without involving nucleotides outside the 3′ NCR (data not shown). The residues at positions A(7184) and/or A(7206) are not sensitive to single strand-specific RNases, although there is no obvious means of base pairing these nucleotides without generating extremely unlikely RNA secondary structure conformations (based upon MFold predictions). It is more plausible that the opposing bulged nucleotides are not susceptible to RNase U_2 as a result of their orientation within the flanking helical structure or that the nucleotides are non-canonically paired within the helix (18). DMS methylation followed by primer extension did show A(7184) to be methylation sensitive (data not shown), however, no DMS data are available for A(7206) due to the limitations of the primer extension method. The ability to detect G(7181) with RNase T_1 treatment is probably due to: (i) breathing of the weak G(7181)-U(7209) base pair flanked by A-U base pairs; (ii) the existence of a repeating dinucleotide sequence between nucleotides 7176 and 7179 (5′-AUAU-3′) which could allow an alternative stem structure to form displacing G(7181); (iii) the overall weak duplex structure predicted to exist at the base of the 3′ NCR stem–loop (for a review see 18).

The RNA secondary structure outside the 3′ NCR was less striking than in the 3′ NCR itself, although there were clearly distinct helical and single-stranded segments of RNA. We propose the existence of a short stem–loop structure immediately upstream of the stop codon with a stretch of pyrimidines (5′-UCUUUU-3′) in the loop region, however, this structure is probably a fortuitous result of the $3D^{pol}$ coding sequence. Our preliminary data suggest that the formation of this upstream stem–loop structure is not required for viral infectivity (unpublished observations). The genomic RNA sequences of HRVs 1A, 1B, 2, 9, 16, 85 and 89 are distinctly different from HRV14 in this region of the genome in that they lack the 8 nt pyrimidine tract (HRV14 nt 7159–7166) and instead have a string of four adenosines preceding a UUU codon (UUC in the case of HRV16) for phenylalanine, which is highly conserved as the C-terminal amino acid of picornavirus $3D^{pol}$ polypeptides (A.C.Palmenberg, personal communication). Although an attractive possibility and consistent with secondary structure models for the 3′ NCR of EMCV (19), we have found no experimental evidence that the four uridylate residues (nt 7163–7166) interact with the polyadenosine tract of the HRV14 genomic RNA. In addition, we have found no evidence for the existence of any long range RNA–RNA interactions or pseudoknot structures involving the HRV14 3′ NCR, as have been described for the prototypic picornavirus, poliovirus (5,6).

The data obtained from mutagenesis of the HRV14 3′ NCR suggest that while a stem–loop structure is phylogenetically conserved among the rhinoviruses, it is not absolutely required for initiation of RNA replication. The deletion of 8 nt at the base of the stem–loop structure results in a severely debilitated RNA replication phenotype *in vivo* which we previously postulated was the result of the abrogation of an RNA–protein interaction between the HRV14 3′ NCR and a 34–36 kDa host cell protein (2). Investigation of this RNA–protein interaction using the Δ18, Δ21 and Δ37 RNAs suggests that the extreme 3′-end of the 3′ NCR may be a major molecular determinant in this interaction, although an intact stem–loop may be the preferred binding site for the host proteins (unpublished data). Further deletion mutagenesis clearly demonstrates that maintenance of the 3′ NCR stem–loop structure is not absolutely essential for virus infectivity and, hence, replication complex recognition and utilization of the mutated RNA template. The infectivity of RNAs harboring the Δ37 mutation was most remarkable. Approximately 84% of the 44 nt 3′ NCR was deleted in the Δ37 3′ NCR, leaving only the primary sequence 5′-GTTTTAT-3′ between the stop codon and the $poly(A)_n$ tract. Nonetheless, virus recovered from a Δ37 RNA transfection displayed a growth phenotype similar to that of the Δ8, Δ18 and Δ21 viruses.

Several models have been proposed to explain the initiation of virus negative strand synthesis from a genomic positive strand RNA template (for a review see 1). One model suggests that a uridylylated VPg molecule (VPg-pU-pU), perhaps in the context of a larger protein precursor (i.e. 3AB) serves as a primer for the viral RNA-dependent RNA polymerase ($3D^{pol}$) (20,21). Uridylylation, polymerase priming and proteolytic maturation of viral replication proteins could occur as concerted events within a membrane bound replication complex. A second model proposes that negative strand synthesis is initiated upon the formation of a snap-back hairpin loop structure involving the 3′-end of positive strand RNAs following the addition of 3′ uridylate residues to the $poly(A)_n$ tract by a cellular enzyme such as terminal uridylyl transferase (22,23). The proposed role of the positive strand 3′ NCR in these models is to direct the replication initiation complex to the authentic template RNA through as yet unidentified RNA–protein contacts (possibly involving cellular factors). While our results do not disprove either of these models, they argue that specific sequence and/or secondary structure determinants within the 3′ NCR are not required to facilitate the basic mechanism of polymerase recognition of the positive strand template, even though these sequences may have evolved to enhance or otherwise regulate the initiation of negative strand synthesis. The correct subcellular localization of viral genomic RNAs to membrane bound replication centers (24–26) and the suggested requirement for concurrent translation of an RNA destined for use as a replication template may be additional considerations which lend efficiency and fidelity to the initiation of negative strand RNA synthesis in an infected cell (27,28).

Previous studies using poliovirus have described mutagenesis of the PV1 3′ NCR which interfered with RNA replication, presumably as the result of disruption of the proposed pseudoknot structure in this region (6,29). Characterization of large deletions in the poliovirus 3′ NCR has recently been described (30), however, these studies may not have detected poorly replicating mutant viruses due to the use of inefficient DNA transfection methodologies. Other than the previous study (2) and the results reported here, mutagenesis of the HRV14 3′ NCR has only been described in the context of a HRV14 3′ NCR/poliovirus chimeric replicon in a CAT reporter assay. While these results also suggested the tolerance of extensive primary sequence variation in this region, the mutations were not studied *in vivo* in the context

of infectious virus (31). The possible contribution of the RNA sequence or secondary structure at the 3′-terminus of the $3D^{pol}$ coding region toward directing specific template utilization has not been addressed. The potential conformational changes which are likely to occur at the 3′-end of the viral RNA template upon assembly of the protein components of the replication complex will clearly be difficult to investigate, although these studies are ongoing.

We have biochemically confirmed the existence of a single stem–loop structure in the 3′ NCR of the HRV14 genomic RNA which, based on phylogenetic primary and secondary structure comparisons, is highly conserved among the *Rhinoviridae*. The stem–loop structure is independent of 5′-flanking coding sequence or the presence of a 3′-flanking poly(A)$_n$ tract. The deletion of most of the primary sequence within the HRV14 3′ NCR, which abolishes formation of the stem–loop structure, results in a severely impaired growth phenotype but does not result in lethality, indicating that initiation of negative strand replication does not require the intact stem–loop structure. Taken together, these results offer the possibility of defining an absolute minimal sequence requirement at the 3′-end of genomic RNA to support the initiation of negative strand synthesis [herein reduced to 5′-GTTTTAT(A)$_n$-3′ following the stop codon] which will ultimately lead to the identification of the underlying molecular mechanism responsible for the process, as well as the opportunity to study the role of the 3′ NCR stem–loop structure in greatly enhancing the efficiency of negative strand synthesis through additional macromolecular contacts with the viral RNA replication machinery.

ACKNOWLEDGEMENTS

We are grateful to Holger Roehl and Louis Leong for critical reading of the manuscript. We are also indebted to Hung Nguyen and Tri Ho for assistance with cell culture. ST was a pre-doctoral trainee of a Public Health Service training grant (GM07134). This work was supported by Public Health Service grant AI22693 from the National Institutes of Health.

REFERENCES

1 Richards,O.C. and Ehrenfeld,E. (1990) *Curr. Topics Microbiol. Immunol.*, **161**, 89–119.
2 Todd,S., Nguyen,J.H. and Semler,B.L. (1995) *J. Virol.*, **69**, 3605–3614.
3 Kitamura,N., Semler,B.L., Rothberg,P.G., Larsen,G.R., Adler,C.J., Dorner,A.J., Emini,E.A., Hanecak,R., Lee,J.J., van der Werf,S. Anderson,C.W. and Wimmer,E. (1981) *Nature*, **291**, 547–553.
4 Racaniello,V.R. and Baltimore,D. (1981) *Science*, **214**, 916–919.
5 Pilipenko,E.V., Maslova,S.V., Sinyakov,A.N. and Agol,V.I. (1992) *Nucleic Acids Res.*, **20**, 1739–1745.
6 Jacobson,S.J., Konings,D.A. and Sarnow,P. (1993) *J. Virol.*, **67**, 2961–2971.
7 Callahan,P.L., Mizutani,S. and Colonno,R.J. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 732–736.
8 Zuker,M. (1989) *Science*, 244, 48–52.
9 Imai,Y., Matsushima,Y., Sugimura,T. and Terada,M. (1991) *Nucleic Acids Res.*, **19**, 2785.
10 Inouye,S. and Inouye,M. (1987) In Narang,S.A. (ed.), *Synthesis and Applications of DNA and RNA*. Academic Press, New York, NY, pp. 181–206.
11 Moazed,D., Stern,S. and Noller,H.F. (1986) *J. Mol. Biol.*, **187**, 399–416.
12 Stern,S., Moazed,D. and Noller,H.F. (1988) *Methods Enzymol.*, **164**, 481–489.
13 Eisenberg,S.P., Coen,D.M. and McKnight,S.L. (1985) *Mol. Cell Biol.*, **5**, 1940–1947.
14 Charini,W.A., Burns,C.C., Ehrenfeld,E. and Semler,B.L. (1991) *J. Virol.*, **65**, 2655–2665.
15 Campos,R. and Villarreal,L.P. (1982) *Virology*, **119**, 1–11.
16 Zuker,M. and Stiegler,P. (1981) *Nucleic Acids Res.*, **9**, 133–148.
17 Knapp,G. (1989) *Methods Enzymol.*, **180**, 192–212.
18 Gutell,R.R., Larsen,N. and Woese,C.R. (1994) *Microbiol. Rev.*, **58**, 10–26.
19 Cui,T. and Porter,A.G. (1995) *Nucleic Acids Res.*, **23**, 377–382.
20 Flanegan,J.B., Pettersson,R.F., Ambros,V., Hewlett,N.J. and Baltimore,D. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 961–965.
21 Nomoto,A., Detjen,B., Pozzatti,R. and Wimmer,E. (1977) *Nature*, **268**, 208–213.
22 Andrews,N.C., Levin,D. and Baltimore,D. (1985) *J. Biol. Chem.*, **260**, 7628–7635.
23 Tobin,G.J., Young,D.C. and Flanegan,J.B. (1989) *Cell*, **59**, 511–519.
24 Caliguiri,L.A. and Tamm,I. (1970) *Virology*, **42**, 100–111.
25 Caliguiri,L.A. and Tamm,I. (1970) *Virology*, **42**, 112–122.
26 Tershak,D.R. (1984) *J. Virol.*, **52**, 777–783.
27 Kaplan,G. and Racaniello,V.R. (1988) *J. Virol.*, **62**, 1687–1696.
28 Kuge,S., Saito,I. and Nomoto,A. (1986) *J. Mol. Biol.*, **192**, 473–487.
29 Sarnow,P., Bernstein,H.D. and Baltimore,D. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 571–575.
30 Pierangeli,A., Bucci,M., Pagnotti,P., Degener,A.M. and Perez Bercoff,R. (1995) *FEBS Lett.*, **374**, 327–332.
31 Rohll,J..B., Moon,D.H., Evans,D.J. and Almond,J.W. (1995) *J. Virol.*, **69**, 7835–7844.