Nucleic Acids Research, 2005, Vol. 33, No. 2 ***693–703***
doi:10.1093/nar/gki218

Structural characterization of the highly conserved 98-base sequence at the 3′ end of HCV RNA genome and the complementary sequence located at the 5′ end of the replicative viral strand

Mariola Dutkiewicz and Jerzy Ciesiołka*

Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznań, Poland

Received September 28, 2004; Revised December 20, 2004; Accepted January 11, 2005

ABSTRACT

Oligoribonucleotides that corresponded to the X regions of the (+) and (−) polarity strands of HCV RNA, as well as several shorter oligomers comprising defined stem-loop motifs of their predicted secondary structure models, were analyzed by Pb^{2+}-induced cleavage, partial digestion with specific nucleases and chemical modification. Patterns characteristic of the motifs were compared with those obtained for the full-length molecules and on the basis of such 'structural fingerprinting' conclusions concerning folding of regions X were formulated. It turned out that the secondary structure model of X(+) RNA proposed earlier, the three-stem-loop model composed of hairpins *SL1*, *SL2* and *SL3*, was only partially consistent with our experimental data. We confirmed the presence of *SL1* and *SL3* motifs and showed that the single-stranded stretch adjacent to the earlier proposed hairpin *SL2* contributed to the folding of that region. It seemed to be arranged into two hairpins, which might form a hypothetical pseudoknot by changing their base-pairing systems. These data were discussed in terms of their possible biological significance. On the other hand, analysis of the X(−) RNA and its sub-fragments supported a three-stem-loop secondary structure model for this RNA.

INTRODUCTION

Infections with hepatitis C virus, HCV, are one of the major health problems that concerns 170 million people worldwide, which accounts for over 3% of the population (1). HCV is a small (40–60 nm in diameter), enveloped virus with a single-stranded RNA genome of positive-sense polarity, which acts as a template in replication and translation processes (2–5). The virus has been classified to the *Flaviviridae* family together with pesti- and flaviviruses as a new genus hepaciviruses. Similarly to the other members of the family, HCV has a single open reading frame (ca. 9000 bases) encoding a precursor polyprotein, which is processed into at least 10 viral proteins. During replication of the virus its genomic RNA is transcribed into a complementary RNA strand, which subsequently constitutes a template for synthesis of new viral genomes. In infected cells genomic (positive-sense) and replicative (negative-sense) RNA strands are present in the ratio of 10:1 (6).

At the 5′ and 3′ end of HCV RNA there are two untranslated regions, 5′-UTR and 3′-UTR, which are required for virus translation and replication. The 5′-UTR possesses 341 nt, which are highly conserved in six different types of the virus (more than 90% conservation) (4,7). Most nucleotides of this region serve as an internal ribosomal entry site, IRES, which directs cap-independent translation of viral genes (8). The complementary RNA stretch, located at the 3′ end of viral replicative strand, contains the initiation site for replication by the viral RNA-dependent RNA polymerase (9,10).

The 3′-UTR consists of three distinct stretches: (i) short, nonconserved region that is virus type specific, (ii) variable-length polypyrimidine tract, and (iii) almost completely conserved stretch of 98 nt at the 3′ terminus. The last-named region is termed region X, (+) for genomic RNA and (−) for complementary RNA strand generated during virus replication. The X region was identified in 1995 (11,12), 6 years after the discovery of the virus. No homologous sequence has been found either in relative viruses or viruses phylogenetically distinct or in known cellular mRNAs (13). The 3′-UTR is required for virus *in vivo* infectivity and replication, and it probably contributes to genome stability and translation efficiency (4,5).

*To whom correspondence should be addressed. Tel: +48 61 8528503; Fax: +48 61 8520532; Email: ciesiolk@ibch.poznan.pl

Since the 5′ and 3′ terminal regions of both viral strands, plus and minus polarity, seem to be extremely important for virus life cycle, their folding has been extensively studied. The structure of IRES element of plus polarity viral RNA and its selected sub-domains has been analyzed by biochemical and NMR methods (3,5). Recently, interesting data have also been gathered on folding of the complementary RNA fragment, which is located at the 3′ end of viral strand of minus polarity (14,15). Folding of peculiar region X(+) and its complement X(−), present at the opposite ends of the viral strands, is of particular interest. Experimental data have suggested that region X(+) consists of a very stable long hairpin *SL1* and a much less ordered region of 52 nt (16,17). The latter region has been arranged into two shorter hairpin motifs *SL2* and *SL3* based on computer predictions. However, this arrangement has not been confirmed by experimental probing data and it has been suggested that the 52 nt long sequence may exist in multiple conformations (16,17). For region X(−), two different secondary structure models have been proposed based on computer predictions (18). Very recently, the structure of this region has been investigated by enzymatic mapping and thermodynamic modeling (19) and the data have been consistent with the presence of a large terminal stem loop, which constitutes 'a mirror image' of the *SL1* motif found in the X(+) region.

Here, we describe the results of the structural analysis of oligonucleotides that correspond to the X regions of the (+) and (−) polarity strands of HCV RNA by the Pb^{2+}-induced cleavage method, partial digestion with specific nucleases and chemical probing. Oligonucleotides that represent selected structural elements of the predicted secondary structure models of X(+) and X(−) RNAs were also synthesized, their structures were analyzed in solution and the results were compared to those obtained with the full-length molecules. On the basis of such 'structural fingerprinting' conclusions concerning folding of these peculiar regions of viral RNA were drawn.

MATERIALS AND METHODS

Materials

The materials used in this study were from the following sources: [γ-^{32}P]ATP (4600 Ci/mmol) was from ICN and all the chemicals were from Sigma or Serva. Enzymes: *Taq* polymerase, nuclease S1, ribonuclease T1, polynucleotide kinase, T7 RNA polymerase, ribonuclease inhibitor and T4 RNA ligase were purchased from MBI Fermentas. Chemically synthesized oligoribonucleotides: SL2ab and SL2(−) were purchased from Dharmacon (Lafayette).

DNA template constructs and RNA synthesis

All oligodeoxynucleotides used in the construction of the DNA templates were deprotected after synthesis and purified on 12% polyacrylamide gels. DNA bands were excised, eluted with 0.3 M sodium acetate, pH 5.2, 1 mM EDTA, precipitated with 3 vol of ethanol and DNA was recovered by centrifugation and dissolved in TE buffer. The DNA templates for the X(+) RNA and fragments SL3, SL2 and SL3(−) were prepared as follows (20,21): two DNA oligomers were used for synthesis of each molecule (see Table 1). Equimolar amounts of both oligomers were annealed and a double-stranded DNA template was generated by PCR, except for SL3(−), for which hybridized oligomers gave the full-length dsDNA template. The reaction mixture contained: 1 μM of both DNA oligomers, 10 mM Tris–HCl, pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl, 200 μM each dNTP and 1.25 U of *Taq* polymerase. The reaction was performed on a Biometra UNOII thermocycler for 7 cycles of 30 s at 94°C, 30 s at 50°C and 2 min at 72°C. The reaction products were purified by phenol/chloroform (1:1) extraction, precipitated with ethanol and the DNA template was finally dissolved in TE buffer. The DNA template for X(+) RNA was diluted to a concentration of 0.3 μM and used in the subsequent PCR reactions for preparation of the DNA template for X(−) RNA. Primers used in this reaction are shown in Table 1. The PCR conditions were as described above, with 3 nM template, 0.25 μM concentration of the primers and 30 cycles used. The obtained DNA template was purified, dissolved in TE buffer and used in transcription reactions.

Table 1. Oligonucleotides used in this study

Oligomers and their length (nt)	Method of synthesis
X(+) (98)	*In vitro* transcription; primers: D1 (taatacgact cactataggt ggctccatct tagccctagt cacggctagc tgtgaaaggt ccgtgagccg c) and D2 (acatgatctg cagagaggcc agtatcagca ctctctgcag tcatgcggct cacggacc)
SL3 (20)	*In vitro* transcription; primers: T7 (taatacgact cactatag) and D4 (agggctaaga tggagccacc tatagtgagt cgtatta)
SL2 (31)	*In vitro* transcription; primers: T7 (taatacgact cactatag) and D5 (gctcacggac ctttcacagc tagccgtgac ctatagtgag tcgtatta)
SL2ab (33)	Chemical synthesis
X(−) (99)	*In vitro* transcription; primers: D6 (taatacgact cactataggc atgatctgca gagaggc) and D7 (ggtggctcca tcttagc); template: DX (taatacgact cactataggt ggctccatct tagccctagt cacggctagc tgtgaaaggt ccgtgagccg catgactgca gagagtgctg atactggcct ctctgcagat catgt)
SL3(−) (24)	*In vitro* transcription; primers: D10 (taatacgact cactatagga gggctaagat ggagccactc c) and D11 (ggagtggctc catcttagcc ctcctatagt gagtcgtatt a)
SL2(−) (27)	Chemical synthesis

The transcription reactions contained: 0.4 μM DNA template, 40 mM Tris–HCl, pH 8.0, 10 mM $MgCl_2$, 2 mM spermidine, 5 mM DTT, 1 mM each NTP, 2.5 mM guanosine, 750 U/ml ribonuclase inhibitor and 2000 U/ml T7 RNA polymerase. Following 4 h incubation at 37°C the RNA transcripts were purified on denaturing 8% polyacrylamide gels and labeled with ^{32}P at their 5′ or 3′ ends with polynucleotide kinase or RNA ligase according to standard procedures.

Metal ion-induced cleavage

Prior to cleavage reaction with Pb^{2+} ions the ^{32}P-labeled RNAs were renatured in the buffer: 40 mM NaCl, 10 mM Tris–HCl, pH 7.2, 10 mM $MgCl_2$ by heating at 65°C for 5 min and slow cooling to 37°C. The samples were then supplemented with tRNA carrier (Boehringer) to a final RNA concentration of 8 μM. Subsequently, lead acetate solution was added and the

reactions proceeded at 37°C for 10 min. The reactions were terminated by mixing their aliquots with 8 M urea/dyes/20 mM EDTA solution and samples were loaded on 12% polyacrylamide, 0.75% bis-acrylamide and 7 M urea gels. Electrophoresis was performed at 2000 V for 2–3 h, followed by autoradiography at −70°C with an intensifying screen or phosphorimaging with Typhoon 8600 analyzer (Molecular Dynamics).

Nuclease mapping

The ^{32}P-labeled RNA transcripts were renatured and supplemented with tRNA carrier as described above. Limited digestions with nuclease S1 and ribonuclease T1 were carried out in the buffer: 10 mM Tris–HCl, pH 7.2, 40 mM NaCl and 10 mM $MgCl_2$. In reactions with nuclease S1, $ZnCl_2$ was also present at 1 mM concentration. Reactions were performed at 37°C for different time intervals with 300 U/ml concentration of nuclease S1, and 50 U/ml concentration of ribonuclease T1. The reactions were terminated by adding 8 M urea/dyes/20 mM EDTA solution and freezing samples on dry ice. The reaction products were analyzed by polyacrylamide gel electrophoresis and visualized by autoradiography or phosphorimaging.

Chemical modification

Chemical modification of cytosine residues with DMS and adenine residues with DEPC was performed essentially as described (22). Briefly, 3′-end-^{32}P-labeled RNA in the buffer: 80 mM HEPES-NaOH, pH 8.0, 40 mM NaCl, 10 mM $MgCl_2$ (200 µl) was renatured and incubated with 1/20 vol of DMS in ethanol (1:12 v/v) at 37°C for 10 min or with 10 µl of DEPC at 30°C for 60 min. The reactions were terminated by ethanol precipitation. In the reaction with DMS, the pellet was dissolved in a 10 µl ice-cold solution of hydrazine in water (1:1, v/v), incubated for 10 min at 0°C and precipitated with ethanol. Subsequently, the pellets from DMS and DEPC reactions were dissolved in a 10 µl 1 M aniline-acetate, pH 4.5, buffer and incubated at 60°C for 15 min in the dark. After precipitation with ethanol the RNA was dissolved in 7 M urea and loaded on a 12% polyacrylamide gel.

Identification of cleavage sites

In order to assign the cleavage sites, products of metal-ion-induced cleavage or nuclease digestion were run along with the products of alkaline RNA hydrolysis and limited T1 nuclease digestion of the same RNA. An alkaline hydrolysis ladder was generated by incubation of 5′- or 3′-end-^{32}P-labeled RNA with 5 vol of formamide at 100°C for 10 min. Partial T1 nuclease digestion was performed in 50 mM sodium citrate, pH 5.3, 7 M urea with 0.2 U of the enzyme at 55°C for 10 min.

RESULTS AND DISCUSSION

Experimental approach

Two complementary oligonucleotides that correspond to region X(+) of genomic HCV RNA and region X(−) of replicative viral strand are shown in Figure 1. The X(+) RNA is arranged into a secondary structure model, which has been proposed earlier in the literature (16,17). The X(−) RNA is folded in an analogous model and the structure represents one

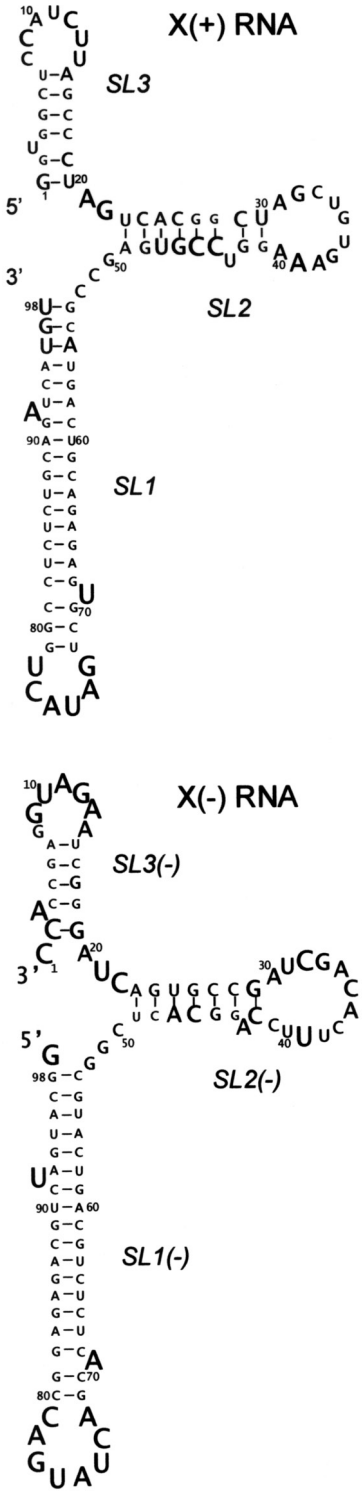

Figure 1. Secondary structure models of the RNA X(+) and RNA X(−). The sequences correspond to the highly conserved X regions of the (+) and (−) polarity strands of HCV RNA. The RNA X(+) forms a three-stem-loop structure proposed earlier in the literature, the RNA X(−) is arranged into an analogous, hypothetical model. In the figure, the propensity of a given base to be single-stranded (ss-count value) was also shown. The ss-count values were calculated using *mfold 3.1* computer program and all folds within 10% of the optimal free energy. The nucleotide symbols have different sizes: the largest symbols indicate that particular nucleotides are single-stranded in all folds (ss-count values of 1), while the smallest symbols show that in all folds nucleotides are paired (ss-count values of 0).

of two models that have been suggested based on *in silico* predictions (18). In order to verify the proposed arrangements of X(+) and X(−) RNAs we combined computer analysis of the preferred secondary structures with experimental probing of the full-length molecules as well as their selected sub-fragments.

We considered all the secondary structures of X(+) and X(−) RNAs predicted *in silico* within 10% of free energy of the most stable variants. The *mfold 3.1*. version of Zuker's program was used (23). Since numbers of predicted structures could vary depending on the folding parameters, default values proposed by the program were always used. The summarized results of computer predictions are illustrated in Figure 1, which shows the calculated propensity of given nucleotides to be single-stranded, ss-count values (24), expressed as nucleotide symbols of different sizes. The largest symbols indicate nucleotides with ss-count values of 1, which are single-stranded in all predicted folds, while the smallest symbols correspond to nucleotides with ss-count values of zero, which are always base-paired. The Figure shows that there are substantial discrepancies between the earlier proposed structural models of X(+) and X(−) RNAs and predicted *in silico* tendency of particular nucleotides to occur in single- or double-stranded regions. Furthermore, some additional discrepancies might be hidden if certain nucleotides with low ss-count values, i.e. occurring preferentially in double-stranded regions, would interact with nucleotides from different regions in distinct folds. Hence, the results of ss-count analysis give only a rough illustration of the tendency of particular regions to form secondary interactions. More precise analysis requires that all individual structures predicted *in silico* should be considered separately.

Since chemical modification and enzymatic digestion experiments have not allowed unambiguous determination of structure of X(+) RNA (16,17), in our probing experiments in addition to the full-length X(+) and X(−) RNAs we also used their sub-fragments. The fragments corresponded to selected structural motifs of secondary structures proposed for the full-length molecules. Experimental approaches involved Pb^{2+}-induced cleavage, limited digestion with specific nucleases and chemical modification. The Pb^{2+} cleavage data were particularly informative. It has been earlier shown that Pb^{2+} ions could detect not only substantial differences in the Watson–Crick base pairing but even subtle differences in the flexibility of polynucleotide chains resulting from changes in stacking or tertiary interactions (25–30). Similar comparative analysis of full-length RNAs and their sub-fragments with the use of Pb^{2+}-induced cleavage method has been also employed in our studies on the structure of 5S rRNAs (31) and on sequential folding of the ribozyme region of the hepatitis delta viral RNA (20).

Secondary structures of X(+) and X(−) RNAs predicted *in silico*

Seven secondary structure models of X(+) RNA were generated *in silico* within 10% energy of the most stable variant with $\Delta G = -46.4$ kcal/mol. The calculated ss-count values strongly support the presence of hairpin *SL1*, to a lower extent hairpin *SL3*, and suggest that region *SL2* may be folded into a structure other than a hairpin (Figure 1). The *SL1* hairpin, with two single-nucleotide bulges in the stem and a 6 nt apical loop, appears in all the seven predicted folds. However, in some folds three 3′ terminal nucleotides: U96, G97 and U98 are unpaired. In those cases, G53 and C54 present in the opposite strand of *SL1* interact with G2 and U3 while A55 is a single-nucleotide hinge between *SL1* and the remaining part of X(+) RNA. Visually, the *SL3* hairpin seems to be also supported by ss-count values (Figure 1). However, such a hairpin appears only in two out of seven models of X(+) RNA predicted *in silico*. As far as the other five cases are concerned, in one model this region is rearranged into a shorter hairpin which involves 12 instead of 20 nt from the RNA 5′ terminus while in the four other models this region does not fold into a hairpin. Most unexpectedly, the ss-count values determined for *SL2* region suggest that these nucleotides do not fold into a hairpin in almost all predicted folds. This observation contradicts the three-SL model of X(+) RNA proposed in the literature. The expected *SL2* hairpin shows up only in one out of seven structures generated *in silico*. The C44-U47 region remains single-stranded in the majority of folds while nucleotides C33-G37 from the large, apical loop are involved in intra-loop base pairing or interact with other regions of the molecule. Interestingly, the ss-count values for nucleotides from single-stranded junction *SL2*/*SL1* are unexpectedly low. In five out of seven predicted folds they interact with nucleotides of regions *SL3* or *SL2*.

The ss-count values determined for the X(−) RNA show that the molecule has fewer possibilities to assume alternative secondary structures than the X(+) RNA (Figure 1). Only three models [seven models for X(+) RNA] are generated *in silico* within the range of 10% of the optimal energy of −44.5 kcal/mol using the same folding parameters. The *SL1*(−) hairpin appears in all three models but *SL3*(−) is present only once although visual inspection of the ss-count values suggests the presence of both these hairpins. Also, only in one model the *SL2*(−) region is folded into a hairpin type motif. In the other two cases nucleotides of that region interact with *SL3*(−). The single-stranded stretch A20-C22, which joins *SL2*(−) and *SL3*(−), is present in one out of three predicted structures. Finally, nucleotides C50-G52 are involved in interactions with region *SL3*(−) in two out of three predicted folds. Thus again, it is essential to examine every thermodynamically possible secondary structure generated *in silico* in order to verify the results of the ss-count analysis.

Structural probing of X(+) RNA and fragments thereof

We synthesized the X(+) RNA and its shorter derivatives: SL2, SL2ab and SL3, which were 98, 31, 33 and 20 nt in length, respectively (Figure 2). The oligomers were subjected to a denaturation–renaturation procedure in the buffer: 40 mM NaCl, 10 mM Tris–HCl, pH 7.2 and 10 mM $MgCl_2$. Although the concentration of Mg^{2+} ions was higher than 1–5 mM, which is thought to occur *in vivo* (32), such conditions are believed to stabilize better spatial RNA folds and have been often used in RNA structural and functional studies. Electrophoresis of the renatured X(+) RNA sample in a non-denaturing polyacrylamide gel showed a strong, major band and a very weak, additional band, which migrated much slower and accounted for <10% of total RNA (data not shown).

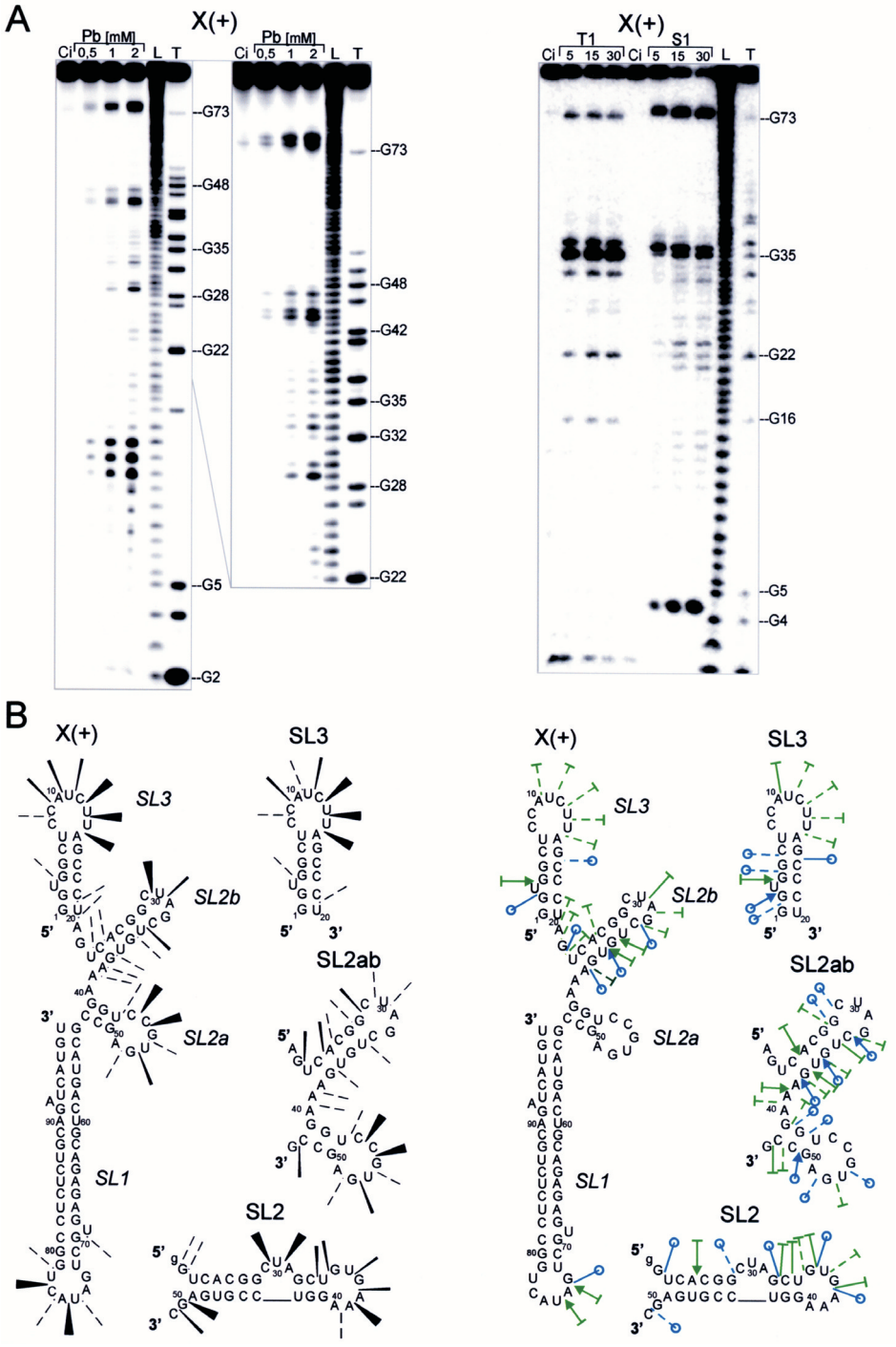

Figure 2. Probing of the structure of RNA X(+) and oligomers: SL3, SL2 and SL2ab by Pb^{2+}-induced cleavage and enzymatic digestion methods. (**A**) Autoradiograms of cleavage of the RNA X(+). The reactions were carried out with [5′-^{32}P]labeled RNA at 37°C with 0.5, 1 and 2 mM Pb^{2+} ions for 10 min or for 5, 15 and 30 min with ribonuclease T1 and nuclease S1. Lanes: Ci, reaction control; L, formamide ladder; T, limited hydrolysis by RNase T1. Guanine residues are labeled on the right. For probing with Pb^{2+}, the short and long run of the gel is shown. (**B**) Cleavages induced in the RNA X(+) and oligomers: SL3, SL2 and SL2ab by Pb^{2+} ions (left panel), ribonuclease T1 and nuclease S1 (right panel). Cleavages are displayed on the secondary structure models, which are most consistent with experimental data. Relative intensities of Pb^{2+} cleavages are marked as follows: dotted lines, weak; lines, strong; black triangles, very strong cleavages. Enzymatic cleavages are denoted by dotted lines, lines or arrows, corresponding to weak, strong and very strong cleavages, respectively. The symbols are additionally marked with open circles or short, perpendicular lines, for RNase T1 and nuclease S1 cleavages, respectively.

The full-length X(+) RNA and its sub-fragments were subjected to structural probing with Pb^{2+} ions and single-strand specific nuclease S1 and ribonuclease T1. Figure 2 shows the Pb^{2+}-induced cleavage and enzymatic probing data obtained for the X(+) RNA. In the Figure the probing data for RNA sub-fragments are also summarized while the corresponding autoradiograms are presented in Figure A in Supplementary Material. The secondary structures shown in Figure 2B are

most consistent with our experimental data, which are discussed below in detail.

The Pb^{2+}-induced cleavage pattern of X(+) RNA is fully consistent with the presence of hairpins *SL1* and *SL3* taking into account our knowledge of Pb^{2+} cleavage specificity (25–30). Phosphodiester linkages of base-paired nucleotides are not cleaved. Weak cleavages appear only at single-nucleotide bulges U3 and U69 as well as at C19 and U20, in the region that closes hairpin *SL3*. This hairpin is destabilized by the presence of terminal GU non-standard base pair and a single-nucleotide bulge. Strong and very strong Pb^{2+}-induced cleavages occur exclusively in the apical loops of *SL1* and *SL3*. In the RNA stretch between U23 and A49, which forms the *SL2* hairpin in the previously proposed structure of X(+) RNA (Figure 1), numerous cleavages are observed. Unexpectedly, cleavages occur in presumably double-stranded stem regions, e.g. at C29, C44, C45, G46, U47 and G48. On the other hand, no cleavages are observed at nucleotides A31G32 in the apical loop, and, most unexpectedly, in single-stranded region G50-C52 immediately adjacent to the earlier predicted hairpin *SL2* (Figures 1 and 2B). Clearly, the U23-A49 stretch folds into a structure other than a hairpin structure. The arrangement of this region proposed in Figure 2B explains most of the observed Pb^{2+} cleavage data (see also Discussion). The results of enzymatic digestions are less clear (Figure 2). Cleavages occurring in the apical loops of hairpins *SL1* and *SL3* fully support the presence of such motifs in the full-length molecule. However, other cleavages, observed predominantly in two regions: U20-G22 and U30-G37, are difficult to explain using any structural model proposed by *mfold* program.

In order to find out whether hairpin motifs *SL2* or *SL3* proposed in the three-SL model of X(+) RNA really exist we synthesized two model oligonucleotides SL2 and SL3, and characterized their structures (Figure 2B and Figure A in Supplementary Material). In oligomer SL2 the first nucleotide from the 5′ end, A21, was changed to G21 in order to facilitate the formation of a stable hairpin stem. As expected, hairpin type motifs were the most stable folds of SL2 and SL3 predicted *in silico* with $\Delta G = -9.2$ kcal/mol and -4 kcal/mol, respectively. However, the *SL2* hairpin motif differs in some details from that present in the three-SL model of X(+) RNA. The apical loop has 6 instead of 10 nt and the C29-A31 region is bulged out of the stem. The computer predictions are well supported by the results of Pb^{2+} probing (Figure 2B and Figure A in Supplementary Material). The patterns of enzymatic cleavages are also consistent with these structures although some presumably single-stranded RNA regions are not cleaved. For example, no cleavages occur in the apical loops of SL3 and SL2, at 5′ and 3′ side, respectively.

The cleavage patterns characteristic of oligomer SL3 is found in the corresponding region of patterns obtained for the full-length X(+) RNA (Figure 2B). It suggests that the corresponding sequences fold into closely related structures. On the other hand, the cleavage patterns of region *SL2* in X(+) RNA and oligomer SL2 are clearly distinct. This indicates that some other nucleotides must be involved in the folding of that region within the full-length molecule.

It was conceivable that the single-stranded stretch which joins *SL2* and *SL1* in the three-SL model of X(+) RNA might be involved in folding of the adjacent *SL2* region (Figure 1). This assumption was verified by structural analysis of oligomer SL2ab, a derivative of SL2 extended at its 3′ terminus with 2 nt: C52 and G53. If our assumption were correct the cleavage pattern of SL2ab should resemble that observed for the corresponding region of the X(+) RNA. Indeed, the respective Pb^{2+}-induced patterns are very similar (Figure 2 and Figure A in Supplementary Material). Most diagnostic cleavages occur within regions U43–U47 of both molecules. The corresponding sequence is not cleaved in oligomer SL2. Moreover, chemical modification of the N3 positions of cytosine residues in SL2ab with DMS shows that C29, C33, C44 and C45 are reactive (Figure 3 and Figure B in Supplementary Material). DEPC modifies the non-Watson–Crick N7 positions of adenine residues A31, A38, A39 and A40. The patterns are similar to those observed in the corresponding region of X(+) RNA. The only differences concern C29, modified very weakly in the fragment but inaccessible in the full-length molecule, and A25, which is unmodified in the fragment but a weak band at the corresponding position occurs in the pattern of X(+) RNA. Altogether, the SL2ab oligomer and the corresponding sequence of X(+) RNA seem to fold similarly. This definitely is not a hairpin type motif since in the SL2 oligomer, which forms a hairpin, only one cytosine residue, C29, is accessible to modification with DMS (Figure 3C). It has to be emphasized that oligomers SL2 and SL2ab fold into very distinct spatial structures despite only a 2 nt difference in their length. Thus, designing shorter minimal RNA constructs, which is a frequently used approach in investigations of large RNA molecules, is not straightforward and the relevance of such constructs to the full-length, parental molecules needs always to be tested.

We think that the SL2ab oligomer is, most likely, arranged into two small hairpins. One hairpin has a six-base pair stem and a 4 nt loop, and in the other, the three-base pair stem is closed with a 6 nt loop. Such structure for this RNA fragment is also predicted by *mfold* program as the most thermodynamically stable with $\Delta G = -10$ kcal/mol. The corresponding sequence of X(+) RNA is folded in the same way in one out of seven structural models predicted *in silico*. This is the fifth structure ($\Delta G = -42.5$ kcal/mol), generated within 10% of energy of the most stable variant ($\Delta G = -46.4$ kcal/mol).

The secondary structure model of X(+) RNA shown in Figure 2B, although consistent with most experimental data, does not explain all observations. In particular, different S1 and T1 digestion patterns are observed for oligomer SL2ab and the corresponding region of full-length molecule. It may be a consequence of large disproportion in the size of the enzymatic probes and the analyzed hairpin motifs *SL2a* and *SL2b*, which additionally are embedded into surrounding sequences in the X(+) RNA. It is also possible that the assemblage of the hairpins is other than simple, which is illustrated in Figure 3D. These two motifs are located one upon the other and after rearrangement of base pairing systems they form a pseudoknot. Although other bases are paired, almost the same nucleotides are involved in the Watson–Crick type interactions. Hence both structures, two hairpins and pseudoknot, could not be simply distinguished experimentally. In particular, it seems not possible with probes mapping unpaired nucleotides, which are essentially the same in both forms. Their stabilization energies might be also similar as the number of base pairs is roughly the same. More precise estimation is not possible since spatial arrangement of individual helices is not known

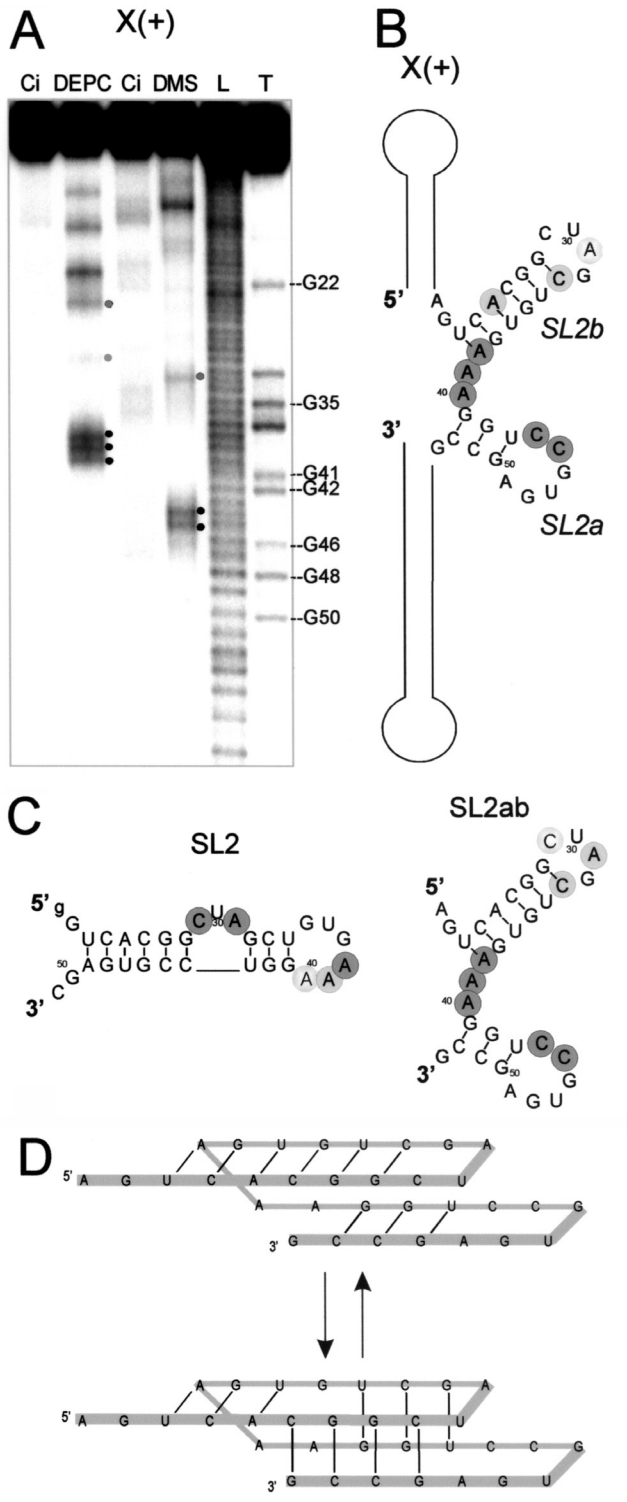

Figure 3. Chemical probing of the *SL2* region of RNA X(+) and oligomers SL2 and SL2ab. (**A**) Probing of adenine residues with DEPC and cytosine residues with DMS; [3′-^{32}P]end-labeled RNA was used. Lanes: Ci, reaction control; L, formamide ladder; T, limited hydrolysis by RNase T1. In the autoradiogram the modified residues are marked with dots and guanine residues are labeled on the right. (**B** and **C**) Correlation of the predicted secondary structures with the experimental data. Relative intensities of strong, weak and very weak modification of adenine residues with DEPC and cytosine residues with DMS are marked by circles with decreasing shadowing. (**D**) Hypothetical structural rearrangement of the *SL2ab* region into a pseudoknot structure (see text for details).

and the contribution of stacking interactions to structure stabilization cannot be taken into account. Interestingly, the existence of a conserved pseudoknot structure has also been proposed at the 3′ terminus of the flavivirus genomic RNA, West Nile virus (33). The pseudoknot base pairings involved nucleotides in the loop of a short hairpin and those in an internal loop of a long stem-loop motif. These interactions were supported by the results of UV-melting experiments of two RNAs with mutations that disrupted the proposed pseudoknot base pairings. Similar mutagenesis analysis could also help to verify the pseudoknot interactions proposed in this work, and appropriate experiments are currently being planned in our laboratory.

Structural probing of X(−) RNA and fragments thereof

The X(−) RNA with additional G residue at the 5′ end, which facilitated *in vitro* transcription, was mapped by Pb^{2+}-induced cleavage method and enzymatic digestion with nuclease S1 and ribonuclease T1. In parallel, shorter oligomers: SL3(−) and SL2(−), corresponding to defined motifs of X(−) RNA secondary structure, were analyzed. The Pb^{2+}-induced cleavage and enzymatic probing data obtained for the full-length RNA are shown in Figure 4. In the Figure the probing data for the shorter oligomers are also summarized while the corresponding autoradiograms are presented in Figure C in Supplementary Material. The cleavage patterns obtained for the full-length molecule appeared to be consistent with both structural models of X(−) RNA: the 'extended model' composed of two long hairpins and the 'three-stem-loop model' (18,19). Hence, these alternative models could be only verified by analyzing subfragments of X(−) RNA: SL3(−) and SL2(−).

The cleavage patterns characteristic of oligomer SL3(−) are present in the corresponding regions of patterns obtained for X(−) RNA (Figure 4B). Thus these RNA stretches have to be arranged into very similar structural motifs. A hairpin type structure is generated for SL3(−) *in silico* with $\Delta G = -5.2$ kcal/mol. As expected, in the apical loops of SL3(−) and X(−) RNA several cleavages are induced by Pb^{2+} ions as well as RNase T1 and nuclease S1. Interestingly, in the SL3(−) oligomer, one strong RNase T1 cleavage and two weak Pb^{2+}-induced cuts occur at G8G9, which are not observed in the X(−) RNA. Possibly, in the latter molecule nucleotides G8G9 may be involved in some kind of higher order interactions.

The results of experimental probing as well as analysis *in silico* suggest that the SL2(−) oligomer folds into a hairpin motif with a six-membered internal loop and a five-membered apical loop (Figure 4B). Such a hairpin is generated by *mfold* as the most thermodynamically stable structure with $\Delta G = -6.6$ kcal/mol. The program also predicts another hairpin with a large, nine-membered apical loop (U31-U40) and a 2 nt bulge (C42A43). This structure, however, is less stable ($\Delta G = -5.6$ kcal/mol) and is not supported by our experimental data.

Finally, Pb^{2+} cleavages induced in the *SL*(−1) region of X(−) RNA fully agree with its arrangement into a very stable double-stranded stem. The apical loop appears to be larger than that present in the *SL*(1) region of X(+) RNA. This is a consequence of lack of interactions between A72 and C79 whereas uridine and guanosine residues present at the

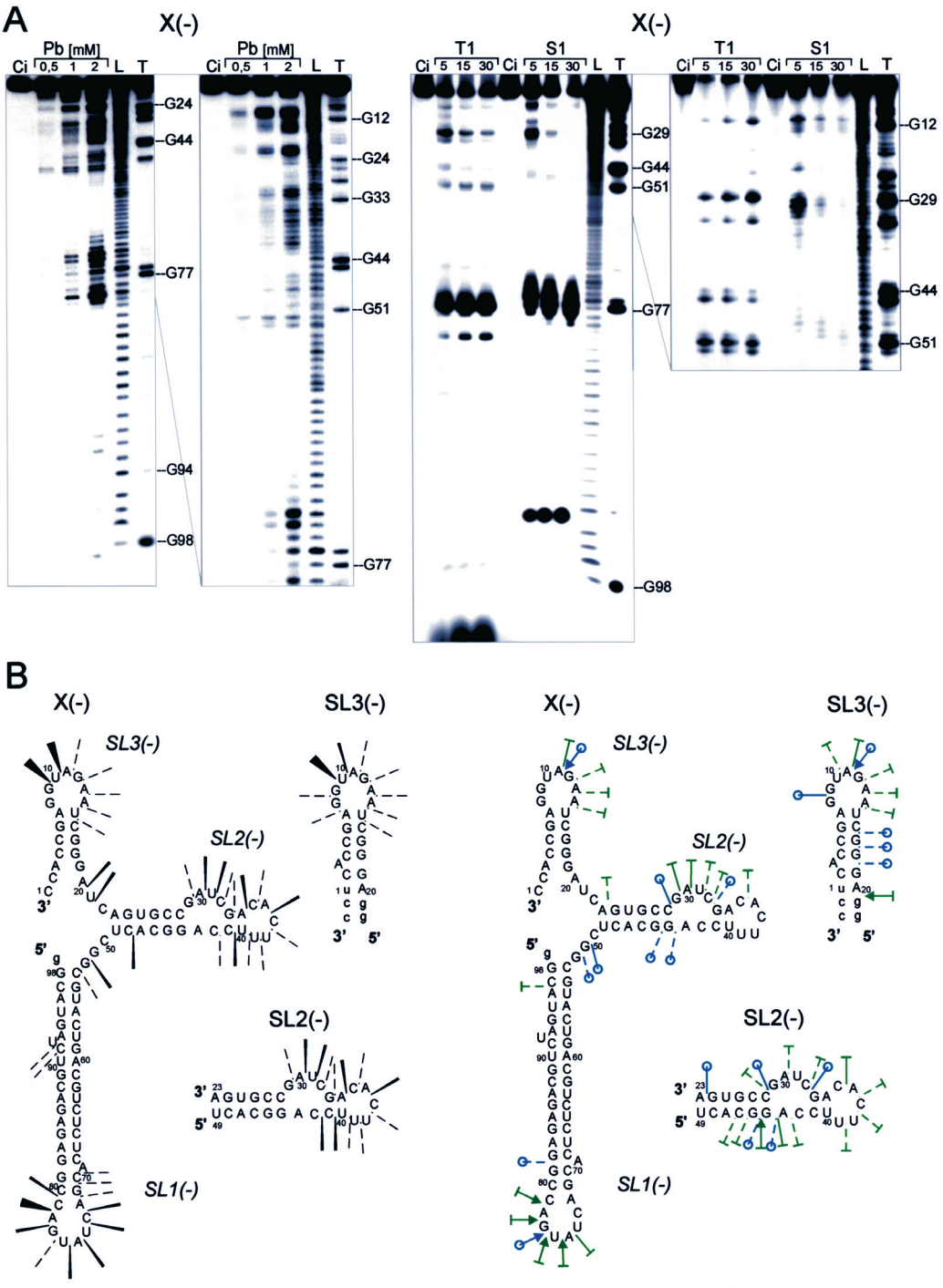

Figure 4. Probing of the structure of RNA X(−) and oligomers: SL3(−) and SL2(−) by Pb^{2+}-induced cleavage and enzymatic digestion methods. (**A**) Autoradiograms of cleavage of the RNA X(−). The reactions were carried out with [5′-^{32}P]labeled RNAs at 37°C with 0.5, 1 and 2 mM Pb^{2+} ions for 10 min or for 5, 15 and 30 min with ribonuclease T1 and nuclease S1. Lanes: Ci, reaction control; L, formamide ladder; T, limited hydrolysis by RNase T1. Guanine residues are labeled on the right. The short and long runs of the gels are shown. (**B**) Cleavages induced in the RNA X(−) and oligomers: SL3(−) and SL2(−) by Pb^{2+} ions (left panel), ribonuclease T1 and nuclease S1. Cleavages are displayed on the secondary structure models, which are most consistent with experimental data. Relative intensities of Pb^{2+} cleavages are marked as follows: dotted lines, weak; lines, strong; black triangles, very strong cleavages. Enzymatic cleavages are denoted by dotted lines, lines or arrows, corresponding to weak, strong and very strong cleavages, respectively. The symbols are additionally marked with open circles or short, perpendicular lines, for RNase T1 and nuclease S1 cleavages, respectively.

corresponding positions of X(+) RNA form a nonstandard base pair. This difference is easily detected by Pb^{2+} ions confirming the great potential of the method for the monitoring of even subtle changes in the flexibility of polynucleotide chains.

The results of mapping of hybridization accessibility of the *SL1* and *SL*(−1) hairpins with oligonucleotide libraries and RNase H digestion have also suggested different sizes of the apical loops (15,19).

Summarizing, our results support the secondary structure model of X(−) RNA consisting of three hairpin motifs. This model is also generated by *mfold* program with $\Delta G = -42.7$ kcal/mol. However, the 'extended' variant consisting of two long hairpin motifs is predicted as the most thermodynamically stable, with $\Delta G = -44.5$ kcal/mol.

Functional role of regions X(+) and X(−)

Although initiation of synthesis of negative RNA strand at the 3′ end of the X region is a crucial step in the replication of HCV genome, this process is poorly understood. It has been shown that the X region of HCV virus interacts specifically with viral protein NS5B, which possesses RNA-dependent RNA polymerase activity (34–36). Footprinting analysis of NS5B–X(+) RNA complex with RNase T1 has revealed that guanosine residues at positions 41, 42, 50 and 53 are protected by the protein (35). This region appears to be a part of the sequence involved in the formation of a hypothetical pseudoknot motif in the X(+) RNA, which is suggested in this work. Direct interactions of viral polymerases with pseudoknot structures at the 3′ ends of viral RNAs have been observed for EMCV (37) and TMV (38).

Studies performed *in vitro* with X(+) RNA and recombinant NS5B enzyme have shown that *de novo* RNA synthesis starts from a single-stranded region closest to the 3′ end of the X region, at U78 in the apical loop of *SL1* (35). Other authors localize the initiation site in the pirimidine-rich region of *SL1* at C86 and U87 (39). Several studies have suggested that only a single-stranded 3′ terminus of an RNA template can bind effectively to the active site of NS5B protein (35,40,41). Moreover, experiments using various artificial templates show that the enzyme initiates RNA synthesis preferentially from the very 3′-terminal or penultimate nucleotide (42,43). In human hepatoma cells a strong preference for GU dinucleotide present at the template 3′ end has been observed (44).

The specificity of recombinant NS5B protein toward model templates is not high and region X is not always necessary for transcription (10,41,45,46). However, the importance of region X for replication is more pronounced in human hepatoma cell lines. By using the HCV replicon system based on subgenomic HCV RNAs it has been shown that all hairpin motifs of region X: *SL1*, *SL2* and *SL3* are required for replication (44,47,48). Moreover, only few point mutations are tolerated in that region (44,47).

The X region of HCV can bind specifically several other viral and host proteins. They include the viral NS3 protein with helicase activity, which presumably is an important component of the replication complex (49). The X region also binds cellular proteins p87 and p130, whose function is not known (50), and ribosomal proteins L22, L3, S3 and mL3 (18). Interestingly, in the case of EBV and HVP1 positive influence of L22 on their translation efficiency has been observed (51). However, in the case of virus Qβ ribosomal proteins of the host are involved in the formation of viral replication complex (52). Therefore, it is unclear whether interactions of ribosomal proteins with region X of HCV are important for replication or for translation, or for both these processes.

Polypyrimidine tract-binding protein (PTB) is one of the early identified cellular proteins, which are able to interact with 3′-UTR of HCV. The binding region comprises 21 nt from the 5′ end of region X and contains a part of consensus sequence that is recognized by PTB (17). Other authors have shown that 19 nt from the 5′ end of region X and 7 nt of the upstream sequence are responsible for efficient complex formation (53). It has been shown by mutagenesis that both the nucleotide sequence of *SL2* and *SL3* domains as well as the secondary structure of this region are important for interactions with PTB (17,53). In particular, nucleotide sequence 33-CUG-35 seems to be involved in these interactions. In the three-stem-loop model of X(+) RNA these nucleotides are located in the large apical loop of hairpin *SL2*. However, in the SL2 oligomer, which represents hairpin-type arrangement of the corresponding fragment of X(+) RNA, the apical loop is smaller, the 29-CUA-31 region is bulged out of the stem, and most importantly, sequence 33-CUG-35 is partially paired (this work, Figures 2 and 3). In the newly proposed model of X(+) RNA, sequence 33-CUG-35 is located in the stem region of hairpin *SL2b*, or alternatively, it participates in the formation of a hypothetical pseudoknot structure. Different proposed arrangements of the *SL2* region could potentially be discussed in terms of interactions of X(+) RNA with PTB, however, the existing data seem to be insufficient for such considerations and require further experimentation. Interestingly, a model of translation efficiency enhancement by PTB has been proposed, in which the protein mediates interactions between the 3′ and 5′ ends of the genome with the involvement of another hypothetical protein Y (54).

It turns out that very high conservation of region X, reaching 100% for the first 52 nt, does not imply that these nucleotides are arranged into an easy to define, stable secondary structure [(16,17), this work]. On the other hand, region X folds autonomously and it does not interact with other regions of viral RNA (16). Thus the observed 'structural heterogeneity' seems to be its intrinsic feature that may be functionally important. It has been proposed earlier that multiple structures formed by the 5′ 52 nt of region X may be required for binding to different viral and/or cellular proteins, each having a distinct role in the viral life cycle (16,17). It is known that certain conservative regions of RNA can adopt alternative structures, which function as regulatory elements or molecular switches between various processes. For example, in alfamo- and ilar-viruses their 3′ ends are folded into several hairpin motifs, which suppress the first step of replication, synthesis of negative RNA strand, by strong binding of an envelope protein (55). Alternatively adopted RNA pseudoknot structure is required for efficient genome replication of AMV either *in vitro* or *in vivo*. Therefore, the virus uses alternative structures of 3′-UTR to switch between replication and translation (55). Similarly, the possibility of formation of a pseudoknot alternatively with a two-hairpin motif in region X of HCV, as suggested in this work, might be important for changing the biological function of that region between translation and replication processes.

CONCLUSIONS

The peculiar region X found at the 3′ terminus of HCV RNA has no respective homologous sequence in other viral or cellular RNA molecules. Knowledge about how this region folds into a spatial structure is crucial to understand the molecular biology of the virus and may be helpful in finding therapeutics against HCV infections. Herein we investigated

folding of two oligoribonucleotides, X(+) RNA, which corresponds to the 3′ terminal region of genomic RNA, and X(−) RNAs, the 5′ terminal region of complementary RNA strand generated during virus replication.

The secondary structure model of the X(+) RNA that has been proposed in the literature (16,17), and which consists of three hairpin motifs *SL1*, *SL2* and *SL3*, is only partially supported by our experimental probing data. Comparative analysis of the full-length molecule and fragments corresponding to the defined motifs of its predicted structure confirms the presence of hairpins *SL1* and *SL3*. The remaining RNA stretch folds, most likely, into two hairpins, which may further form a hypothetical pseudoknot after changing their base-pairing systems. On the other hand, structural analysis of the X(−) RNA and its sub-fragments supports a three-stem-loop model for this RNA.

A comparison of the secondary structures of X(+) and X(−) RNA illustrates that these RNA molecules do not fold into 'mirror images'. Such 'images' could be expected, assuming simplified understanding of folding principles. However, such expectations are met only in some special cases, if oligonucleotides form relatively stable hairpin type motifs. For longer, complementary RNA sequences asymmetry of folding is already revealed by computer programs predicting their secondary structures since the programs allow formation of GU interactions other than the standard Watson–Crick base pairs. As several other non-standard base pairs have been found in RNA molecules, which are not predictable *in silico*, really existing molecules will be undoubtedly more asymmetric. In that light, distinct folding of X(+) and X(−) RNAs is fully understandable. The opposite terminal regions of viral genomic and replicative strands, IRES(+) and IRES(−), are also folded into substantially different structures (5,14,15,56). Therefore, it seems that upon fulfillment of their distinct biological functions complementary RNA sequences may employ either the sequence or shape recognition principle.

The ability of X(+) RNA to assume alternative structures seems to be restricted to the middle part of its sequence. Most likely, this region is arranged into two hairpin motifs, which may form a hypothetical pseudoknot structure. Such structural dynamics might be important for the biological function of that region, e.g. for switching between its involvement in replication or translation (16). It resembles functioning of other RNA structural elements, so-called riboswitches, which take part in the regulation of gene expression at the transcriptional or translational level (57,58). Thus, if highly conserved RNA sequences, which, for example, are often found in viral untranslated regions, do not form obvious secondary structures, their alternative RNA foldings may be crucial for these regions' biological performance.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

We thank Professor Anne-Lise Haenni for stimulating discussion and members of our laboratory for valuable comments on the manuscript. Barbara Smolska is thanked for her excellent technical assistance. This work was supported by the Polish Committee for Scientific Research, grants 6P04B 01720 and 3P04A 02024. Funding to pay the Open Access publication charges for this article was provided by the Polish Committee for Scientific Research.

REFERENCES

1. World Health Organization (1999), Hepatitis C: global prevalence (update). *Weakly Epidemiol. Rec.*, **49**, 425–427.
2. Reed,K.E. and Rice,C.M. (2000) Overview of hepatitis C virus genome structure, polyprotein processing, and protein properties. *Curr. Top. Microbiol. Immunol.*, **242**, 55–84.
3. Branch,A.D. (2000) Hepatitis C virus RNA codes for proteins and replicates: does it also trigger the interferon response? *Semin. Liver Dis.*, **20**, 57–68.
4. Shi,S.T. and Lai,M.M. (2001) Hepatitis C viral RNA: challenges and promises. *Cell. Mol. Life Sci.*, **58**, 1276–1295.
5. Penin,F., Dubuisson,J., Rey,F.A., Moradpour,D. and Pawlotsky,J.M. (2004) Structural biology of hepatitis C virus. *Hepatology*, **39**, 5–19.
6. Gowans,E.J. (2000) Distribution of markers of hepatitis C virus infection throughout the body. *Semin. Liver Dis.*, **20**, 85–102.
7. Cohen,J. (1999) The scientific challenge of hepatitis C. *Science*, **285**, 26–30.
8. Tsukiyama-Kohara,K., Iizuka,N., Kohara,M. and Nomoto,A. (1992) Internal ribosome entry site within hepatitis C virus RNA. *J. Virol.*, **66**, 1476–1483.
9. Behrens,S.E., Tomei,L. and De Francesco,R. (1996) Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *EMBO J.*, **15**, 12–22.
10. Lohmann,V., Korner,F., Herian,U. and Bartenschlager,R. (1997) Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. *J. Virol.*, **71**, 8416–8428.
11. Tanaka,T., Kato,N., Cho,M.J. and Shimotohno,K. (1995) A novel sequence found at the 3′ terminus of hepatitis C virus genome. *Biochem. Biophys. Res. Commun.*, **215**, 744–749.
12. Kolykhalov,A.A., Feinstone,S.M. and Rice,C.M. (1996) Identification of a highly conserved sequence element at the 3′ terminus of hepatitis C virus genome RNA. *J. Virol.*, **70**, 3363–3371.
13. Tanaka,T., Kato,N., Cho,M.J., Sugiyama,K. and Shimotohno,K. (1996) Structure of the 3′ terminus of the hepatitis C virus genome. *J. Virol.*, **70**, 3307–3312.
14. Schuster,C., Isel,C., Imbert,I., Ehresmann,C., Marquet,R. and Kieny,M.P. (2002) Secondary structure of the 3′ terminus of hepatitis C virus minus-strand RNA. *J. Virol.*, **76**, 8058–8068.
15. Smith,R.M., Walton,C.M., Wu,C.H. and Wu,G.Y. (2002) Secondary structure and hybridization accessibility of hepatitis C virus 3′-terminal sequences. *J. Virol.*, **76**, 9563–9574.
16. Blight,K.J. and Rice,C.M. (1997) Secondary structure determination of the conserved 98-base sequence at the 3′ terminus of hepatitis C virus genome RNA. *J. Virol.*, **71**, 7345–7352.
17. Ito,T. and Lai,M.M. (1997) Determination of the secondary structure of and cellular protein binding to the 3′-untranslated region of the hepatitis C virus RNA genome. *J. Virol.*, **71**, 8698–8706.
18. Wood,J., Frederickson,R.M., Fields,S. and Patel,A.H. (2001) Hepatitis C virus 3′X region interacts with human ribosomal proteins. *J. Virol.*, **75**, 1348–1358.
19. Smith,R.M. and Wu,G.Y. (2004) Secondary structure and hybridization accessibility of the hepatitis C virus negative strand RNA 5′-terminus. *J. Viral Hepat.*, **11**, 115–123.
20. Matysiak,M., Wrzesinski,J. and Ciesiołka,J. (1999) Sequential folding of the genomic ribozyme of the hepatitis delta virus: structural analysis of RNA transcription intermediates. *J. Mol. Biol.*, **291**, 283–294.
21. Wrzesinski,J., Łęgiewicz,M., Smólska,B. and Ciesiołka,J. (2001) Catalytic cleavage of *cis*- and *trans*-acting antigenomic delta ribozymes in the presence of various divalent metal ions. *Nucleic Acids Res.*, **29**, 4482–4492.
22. Krol,A. and Carbon,P. (1989) A guide for probing native small nuclear RNA and ribonucleoprotein structures. *Methods Enzymol.*, **180**, 212–226.
23. Zuker,M., Mathews,D.H. and Turner,D.H. (1999) Algorithms and thermodynamics for RNA secondary structure prediction: a practical guide. In Barciszewski,J. and Clark,B.F.C. (eds), *RNA Biochemistry and*

Biotechnology. NATO ASI Series, Kluwer Academic Publishers, pp. 11–43.
24. Chadalavada,D.M., Knudsen,S.M., Nakano,S. and Bevilacqua,P.C. (2000) A role for upstream RNA structure in facilitating the catalytic fold of the genomic hepatitis delta virus ribozyme. *J. Mol. Biol.*, **301**, 349–367.
25. Górnicki,P., Baudin,F., Romby,P., Wiewiórowski,M., Krzyżosiak,W.J., Ebel,J.P., Ehresmann,C.H. and Ehresmann,B. (1989) Use of lead (II) to probe the structure of large RNA's. Conformation of the 3′ terminal domain of *E.coli* 16S rRNA and its involvement in building the tRNA binding sites. *J. Biomol. Struct. Dyn.*, **6**, 971–984.
26. Brunel,C.H., Romby,P., Westhof,E., Ehresmann,C. and Ehresmann,B. (1991) Three-dimensional model of *Escherichia coli* ribosomal 5S RNA as deduced from structure probing in solution and computer modeling. *J. Mol. Biol.*, **221**, 293–308.
27. Ciesiołka,J., Lorenz,S. and Erdmann,V.A. (1992) Structural analysis of three prokaryotic 5S rRNA species and selected 5S rRNA-ribosomal-protein complexes by means of Pb(II)-induced hydrolysis. *Eur. J. Biochem.*, **204**, 575–581.
28. Ciesiołka,J., Lorenz,S. and Erdmann,V.A. (1992) Different conformational forms of *Escherichia coli* and rat liver 5S rRNA revealed by Pb(II)-induced hydrolysis. *Eur. J. Biochem.*, **204**, 583–589.
29. Ciesiołka,J., Michałowski,D., Wrzesinski,J., Krajewski,J. and Krzyżosiak,W.J. (1998) Patterns of cleavages induced by lead ions in defined RNA secondary structure motifs. *J. Mol. Biol.*, **275**, 211–220.
30. Ciesiołka,J. (1999) Metal ion-induced cleavages in probing of RNA structure. In Barciszewski,J. and Clark,B.F.C. (eds), *RNA Biochemistry and Biotechnology*. NATO ASI Series, Kluwer Academic Publishers, pp. 111–121.
31. Ciesiołka,J. and Krzyżosiak,W.J. (1996) Structural analysis of two plant 5S rRNA species and fragments thereof by lead-induced hydrolysis. *Biochem. Mol. Biol. Int.*, **39**, 319–328.
32. Traut,T.W. (1994) Physiological concentrations of purines and pyrimidines. *Mol. Cell. Biochem.*, **140**, 1–22.
33. Shi,P.-Y., Brinton,M.A., Veal,J.M., Zhong,Y.Y. and Wilson,W.D. (1996) Evidence for the existence of a pseudoknot structure at the 3′ terminus of the flavivirus genomic RNA. *Biochemistry*, **35**, 4222–4230.
34. Cheng,J.C., Chang,M.F. and Chang,S.C. (1999) Specific interaction between the hepatitis C virus NS5B RNA polymerase and the 3′ end of the viral RNA. *J. Virol.*, **73**, 7044–7049.
35. Oh,J.W., Sheu,G.T. and Lai,M.M. (2000) Template requirement and initiation site selection by hepatitis C virus polymerase on a minimal viral RNA template. *J. Biol. Chem.*, **275**, 17710–17717.
36. Kolykhalov,A.A., Mihalik,K., Feinstone,S.M. and Rice,C.M. (2000) Hepatitis C virus-encoded enzymatic activities and conserved RNA elements in the 3′ nontranslated region are essential for virus replication in vivo. *J. Virol.*, **74**, 2046–2051.
37. Cui,T. and Porter,A.G. (1995) Localization of binding site for encephalomyocarditis virus RNA polymerase in the 3′-noncoding region of the viral RNA. *Nucleic Acids Res.*, **23**, 377–382.
38. Osman,T.A., Hemenway,C.L. and Buck,K.W. (2000) Role of the 3′ tRNA-like structure in tobacco mosaic virus minus-strand RNA synthesis by the viral RNA-dependent RNA polymerase *in vitro*. *J. Virol.*, **74**, 11671–11680.
39. Kim,M., Kim,H., Cho,S.P. and Min,M.K. (2002) Template requirements for *de novo* RNA synthesis by hepatitis C virus nonstructural protein 5B polymerase on the viral X RNA. *J. Virol.*, **76**, 6944–6956.
40. Hong,Z., Cameron,C.E., Walker,M.P., Castro,C., Yao,N., Lau,J.Y. and Zhong,W. (2001) A novel mechanism to ensure terminal initiation by hepatitis C virus NS5B polymerase. *Virology*, **285**, 6–11.
41. Kao,C.C., Yang,X., Kline,A., Wang,Q.M., Barket,D. and Heinz,B.A. (2000) Template requirements for RNA synthesis by a recombinant hepatitis C virus RNA-dependent RNA polymerase. *J. Virol.*, **74**, 11121–11128.
42. Shim,J.H., Larson,G., Wu,J.Z. and Hong,Z. (2002) Selection of 3′-template bases and initiating nucleotides by hepatitis C virus NS5B RNA-dependent RNA polymerase. *J. Virol.*, **76**, 7030–7039.
43. Butcher,S.J., Grimes,J.M., Makeyev,E.V., Bamford,D.H. and Stuart,D.I. (2001) A mechanism for initiating RNA-dependent RNA polymerization. *Nature*, **410**, 235–240.
44. Yi,M. and Lemon,S.M. (2003) Structure-function analysis of the 3′ stem-loop of hepatitis C virus genomic RNA and its role in viral RNA replication. *RNA*, **9**, 331–345.
45. Oh,J.W., Ito,T. and Lai,M.M. (1999) A recombinant hepatitis C virus RNA-dependent RNA polymerase capable of copying the full-length viral RNA. *J. Virol.*, **73**, 7694–7702.
46. Reigadas,S., Ventura,M., Sarih-Cottin,L., Castroviejo,M., Litvak,S. and Astier-Gin,T. (2001) HCV RNA-dependent RNA polymerase replicates *in vitro* the 3′ terminal region of the minus-strand viral RNA more efficiently than the 3′ terminal region of the plus RNA. *Eur. J. Biochem.*, **268**, 5857–5867.
47. Yi,M. and Lemon,S.M. (2003) 3′ nontranslated RNA signals required for replication of hepatitis C virus RNA. *J. Virol.*, **77**, 3557–3568.
48. Friebe,P. and Bartenschlager,R. (2002) Genetic analysis of sequences in the 3′ nontranslated region of hepatitis C virus that are important for RNA replication. *J. Virol.*, **76**, 5326–5338.
49. Banerjee,R. and Dasgupta,A. (2001) Specific interaction of hepatitis C virus protease/helicase NS3 with the 3′-terminal sequences of viral positive- and negative-strand RNA. *J. Virol.*, **75**, 1708–1721.
50. Inoue,Y., Miyazaki,M., Ohashi,R., Tsuji,T., Fukaya,K., Kouchi,H., Uemura,T., Mihara,K. and Namba,M. (1998) Ubiquitous presence of cellular proteins that specifically bind to the 3′ terminal region of hepatitis C virus. *Biochem. Biophys. Res. Commun.*, **245**, 198–203.
51. Toczyski,D.P. and Steitz,J.A. (1991) EAP, a highly conserved cellular protein associated with Epstein–Barr virus small RNAs (EBERs). *EMBO J.*, **10**, 459–466.
52. Senear,A.W. and Steitz,J.A. (1976) Site-specific interaction of Qbeta host factor and ribosomal protein S1 with Qbeta and R17 bacteriophage RNAs. *J. Biol. Chem.*, **251**, 1902–1912.
53. Tsuchihara,K., Tanaka,T., Hijikata,M., Kuge,S., Toyoda,H., Nomoto,A., Yamamoto,N. and Shimotohno,K. (1997) Specific interaction of polypyrimidine tract-binding protein with the extreme 3′-terminal structure of the hepatitis C virus genome, the 3′X. *J. Virol.*, **71**, 6720–6726.
54. Ito,T. and Lai,M.M. (1999) An internal polypyrimidine-tract-binding protein-binding site in the hepatitis C virus RNA attenuates translation, which is relieved by the 3′-untranslated sequence. *Virology*, **254**, 288–296.
55. Olsthoorn,R.C., Mertens,S., Brederode,F.T. and Bol,J.F. (1999) A conformational switch at the 3′ end of a plant virus RNA regulates viral replication. *EMBO J.*, **18**, 4856–4864.
56. Gallego,J. and Varani,G. (2002) The hepatitis C virus internal ribosome-entry site: a new target for antiviral research. *Biochem. Soc. Trans.*, **30**, 140–145.
57. Mandal,M. and Breaker,R.R. (2004) Gene regulation by riboswitches. *Nat. Rev. Mol. Cell. Biol.*, **5**, 451–463.
58. Nudler,E. and Mironov,A.S. (2004) The riboswitch control of bacterial metabolism. *Trends Biochem. Sci.*, **29**, 11–17.