Nucleic Acids Research, Vol. 20, No. 7 ***1739–1745***

Towards identification of *cis*-acting elements involved in the replication of enterovirus and rhinovirus RNAs: a proposal for the existence of tRNA-like terminal structures

Evgeny V.Pilipenko, Svetlana V.Maslova, Alexander N.Sinyakov and Vadim I.Agol*
Institute of Poliomyelitis and Viral Encephalitides, The USSR Academy of Medical Sciences, Moscow Region 142782 and Moscow State University, Moscow 119899, Russia

Received November 8, 1991; Revised and Accepted March 3, 1992

ABSTRACT

On the basis of a comparative analysis of published sequences, models for the secondary structure of the 3′-terminal [poly(A)-preceding] untranslated region of the entero- and rhinovirus RNAs were worked out. The models for all these viruses share a common core element, but there are an extra enterovirus-specific element and still an additional element characteristic of a subset of enterovirus RNAs. The two latter models were verified for poliovirus and coxsackievirus B genomes by testing with single-strand and double-strand specific enzymatic and chemical probes. A tRNA-like tertiary structure model for the 3′-terminal folding of enterovirus RNAs was proposed. A similar folding was proposed for the 3′ termini of the negative RNA strands as well as for the 5′ termini of the positive strand of all entero- and rhinovirus RNAs. Implications of these data for template recognition during negative and positive RNA strands synthesis and for the evolution of the picornavirus genomes are discussed.

INTRODUCTION

The 7.5–8.5 kb-long single-strand (SS) RNA genome of picornaviruses (small naked icosahedral animal viruses) is replicated by the virus-coded multi-component machinery using mechanisms that are presently poorly understood (for reviews, see refs. 1, 2). Nevertheless, there is little doubt that these mechanisms include recognition of the template [positive (+) and negative (−) strands] 3′ termini during initiation of the respective complementary strand synthesis; initiation of (+) strands may in principle involve also recognition of the 'left' end of the double-stranded (DS) replicative form. It could be anticipated that the recognizable signals on the (+) and (−) templates have something in common, since only a single RNA polymerase species appears to be encoded in the picornaviral genome. Nevertheless, the very 3′-termini of the (+) and (−) strands are quite different, a poly(A) tail and a heteropolymeric sequence, respectively. It seems highly unlikely that such an unspecific segment as a poly(A) tail is solely recognized in the (+) strand template; rather, certain poly(A)-preceding sequences could be supposed to contain appropriate *cis*-acting signals. As far as we are aware, no kinship has previously been noted, however, between the primary structures of the poly(A)-adjoining segment of the (+) strand, on the one hand, and the deduced structure of the 3′ end of the (−) strand, on the other. We decided therefore to look for a possible similarity between these genomic segments at higher structural levels, those of secondary and tertiary structures.

Since reliable secondary structure models for a ca. 100 nucleotides-long segment, forming a separate domain at the 5′ end of the viral RNAs, have been available (3, 4), we decided first to define the secondary structure of the appropriate segment at the 3′ end. Representatives of only two picornavirus genera, *Enterovirus* (polioviruses, coxsackieviruses, echoviruses, and other enteroviruses) and *Rhinovirus* (causative agents of the majority of common cold cases), were included in this study, because all of them exhibit a very similar structural organization of the entire 5′ untranslated region, distinct from that in other picornaviruses (3, 5–8). The previously described approach (5) was employed. First, models for the secondary structure of the 3′ ends of the entero- and rhinovirus (+) RNAs were derived on the basis of a comparative analysis of published sequences, and then these models were verified by using SS- and DS-specific probes. A further analysis allowed us to suggest that the 3′ termini of enterovirus (+) RNAs could acquire a tRNA-like L-shaped tertiary folding. A similar folding appeared to be characteristic also of the 3′ termini of the (−) strands and of the 5′ termini of the (+) strands of all entero- and rhinovirus RNAs. Common and distinct structural features of the termini of the genomes of different representatives of entero- and rhinoviruses allowed us to speculate about their functional significance as well as evolution.

MATERIALS AND METHODS

The conditions used for the full-length viral RNA treatment with dimethyl sulfate (DMS) and with S1 and cobra venom (CV1) nucleases as well as for location of the modified nucleotides or cleavage sites by the primer extension technique were described previously (5). The conditions for the *B. cereus* and *Phy M* nucleases treatment (both enzymes from Pharmacia) were as

* To whom correspondence should be addressed

follows. One μg of RNA was preincubated in 30 mM tris-HCl, pH 7.5, 50 mM KCl, 75 mM K acetate, 2 mM $MgCl_2$, and 2 mM dithiothreitol at 30° for 6 min, followed by the incubation, in a total volume of 20 μl, with *B. cereus* (1 unit) or *Phy M* (1.5 units) nucleases for 15 min at the same temperature. After two phenol deproteinizations followed by two ethanol precipitations, the modified RNA species were dried *in vacuo* and stored until sequencing.

Oligodeoxynucleotide primers T_{14}CTCCG and T_{14} CGCAC were used for the reverse transcription of the poliovirus (PV), type 1 and type 3, and coxsackievirus (Cox) B3 RNAs, respectively.

Nucleotide sequences of the genomes of PV1−3, CoxB1, B3, B4, human rhinoviruses (HRV) 1B, 2, 14, and 89 were taken from the references listed in (5), but a few positions of the CoxB3 3-UTR were corrected on the basis of our own determinations. The other sequences were from the following sources: CoxA9 (9), CoxA21 (10), bovine enterovirus (BEV) (11), swine vesicular disease virus (SVDV) (12), enterovirus 70 (EV70) (13), ECHO6 (14), and ECHO11 (15).

RESULTS

Secondary structure models for the 3′-termini of entero- and rhinovirus RNAs

As stated in Introduction, our first goal was to derive a consensus secondary structure model for the 3′-terminal portion of entero- and rhinovirus RNAs involved in the interaction with the viral

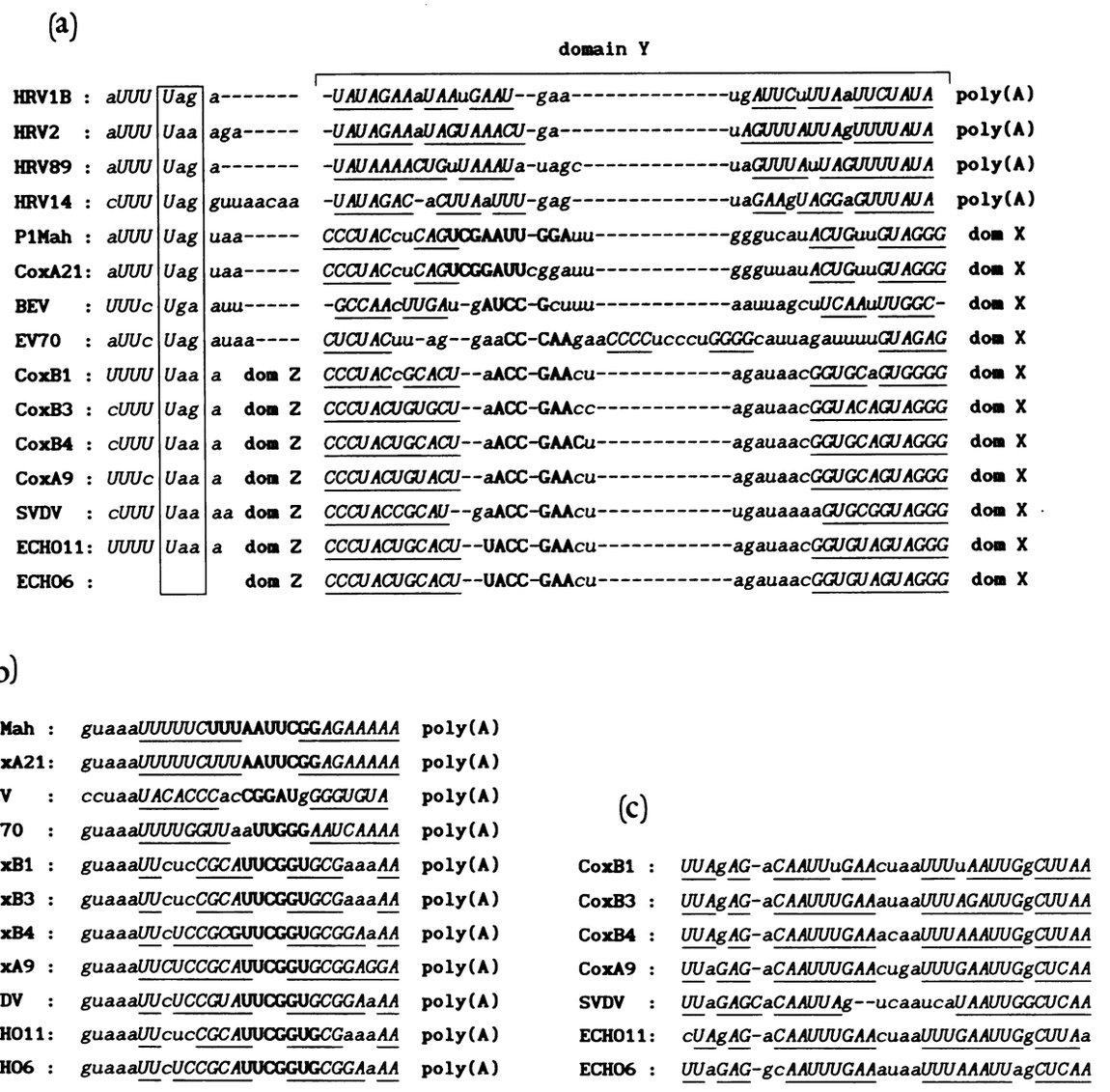

Fig. 1. Structural elements in the 3-UTRs of entero- and rhinoviruses. General organization of the 3-UTRs and the primary structure of domains Y (a), X (b) and Z (c). Terminator codons of the polyprotein are boxed (ECHO6 sequence around the terminator is not available). Oligonucleotides with a potential for base pairing are printed in upper case letters; those of them assumed to be involved in the intra-domain secondary or inter-domain tertiary interactions are underlined and printed in bold characters, respectively; the Us from the neighborhood of the terminator codons assumed to pair with poly(A) to form element S are also printed in upper case letters. Presumedly unpaired nucleotides are shown in smaller case letters. Dashes are used for the alignment purposes. RNA segments separating domain Y from the terminator codons and poly(A) are designated, in panel (a), domains (dom) Z and X, respectively. A few positions in the previously reported CoxB3 RNA sequence were changed on the basis of the evidence obtained in this study.

RNA polymerase. We thought it appropriate to perform such an analysis with the 3'-untranslated regions (3-UTRs) of the genome, where the recognizable *cis*-acting signals were expected to be primarily located. The problem, however, was that the 3-UTR lengths varied considerably, from about 40 nt in rhinoviruses up to about 100 nt in some enteroviruses. We argued that the signals common for all the entero- and rhinoviruses (if there are such common elements at all) should already be present within the shortest subset of the 3-UTRs, that of rhinoviruses. As shown in Fig. 1a, these segments could be reasonably well aligned with each other, despite a significant divergence of the nucleotide sequences. Importantly, all of them exhibited a potential to form a similar hairpin-loop structure (the respective complementary sequences are underlined). The similarly-sized segments able to generate a hairpin-loop structure (though with somewhat shorter stems) were found among the enterovirus 3-UTRs as well (Fig. 1a; see also ref. 16); the presence of a unique decanucleotide insert in the EV70 3-UTR should perhaps be mentioned. Although the intergroup divergence of the primary structures was considerable, several positions were highly conserved among the whole set of sequences. In general, however, the maintenance of the secondary structure could be ascribed to compensating mutations. This core element shared by all the entero- and rhinovirus 3-UTRs was designated domain Y.

All the enteroviruses, as opposed to rhinoviruses, had at least an additional segment, ~25 nt in length, between domain Y and the poly(A) tail (Figs. 1a and 1b). This segment proved to be relatively well conserved and exhibited a potential to form a separate hairpin-loop, sometimes with the aid of a few nt from the poly(A) tail. This element was named domain X. The enterovirus 3-UTRs could further be subdivided, using as a criterion the distance between the termination codon (UAG, UAA, or UGA) and domain Y. The spacer was only a few nt in length in the PV, CoxA21, BEV, and EV70 genomes. This subset of enteroviruses will be referred to as polio-like. On the other hand, there were inserts ca. 35 nt in length between the terminator and domain Y in CoxB1, CoxB3, CoxB4, SVDV, Cox A9, ECHO6 and ECHO11 RNAs (Figs. 1a and 1c). These highly conserved inserts, domain Z, could be unambiguously aligned; they also exhibited a potential to form a hairpin-loop structure. The enteroviruses having domain Z will be referred to as CoxB-like.

A characteristic feature of domains X, Y, and Z is that they invariably have either a perfect stem or, if internal loops in the stem are present, these lops are symmetric and hence exerting a relatively mild destabilizing effect (17).

An oligo(U) tract (4–5 nt in length, including U-residues of the terminator codons) was located just upstream of the terminator codons. This element could be assumed to pair with A-residues from the poly(A) tail; a putative helix thus generated was designated S. In the rhinovirus RNAs, element S could simply serve to extend the helix of domain Y; on the other hand, it might ensure the formation of a 'closed' structure consisting of two or three hairpin-loops in the polio-like and CoxB-like 3-UTRs, respectively. In the latter case, the 3-UTR could be described as having a cloverleaf-like structure (Fig. 2).

The experimental verification of the model was performed using the PV1, PV3, and CoxB3 RNAs as examples. Nucleases

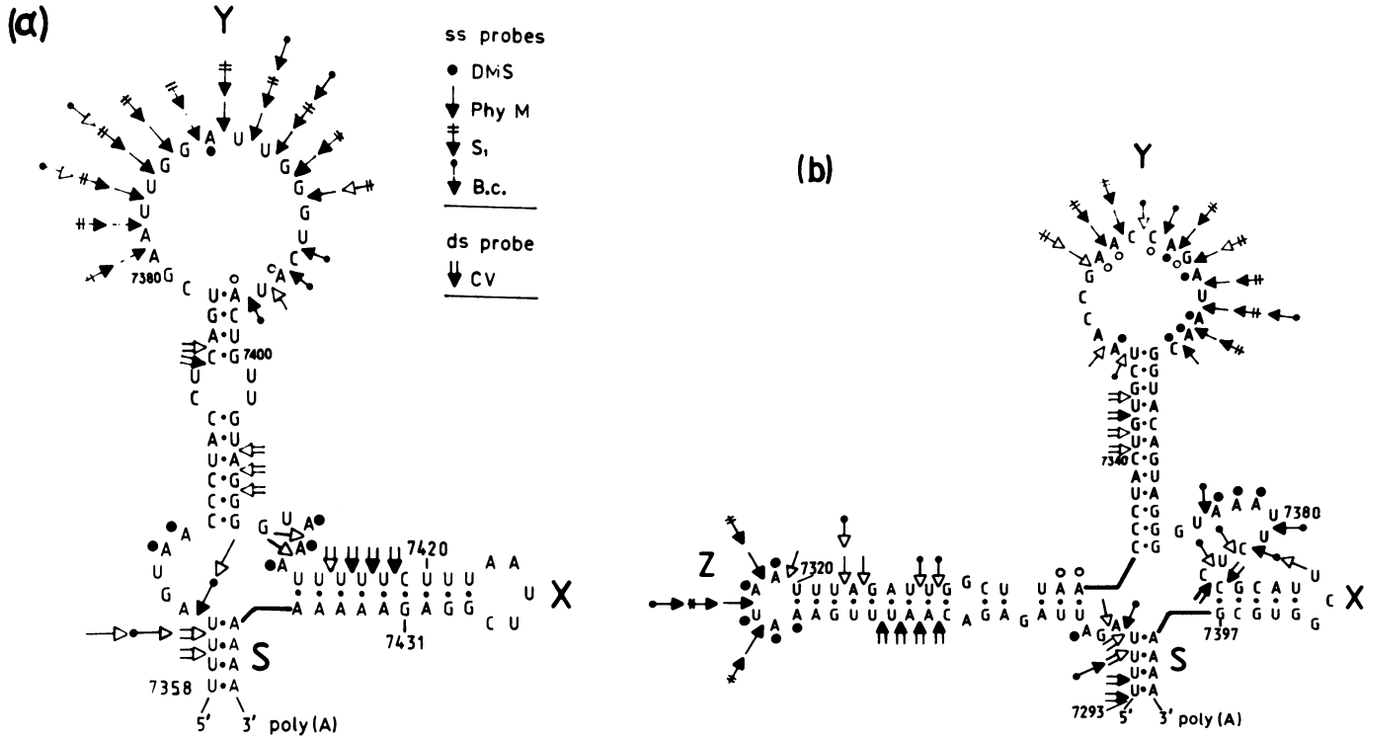

Fig. 2. Secondary structure models for the 3-UTRs of PV1 (a) and CoxB3 (b) RNAs, and their experimental verification. The positions of cleavages or modifications induced by SS- and DS-specific probes, as determined by the primer extension technique, are indicated using symbols given in panel a [Phy M, S, B.c. (*Bacillus cereus*), and CV correspond to the respective nucleases]. Relatively weak signals are represented by open symbols. The pattern of cleavages and modifications for the PV3 3-UTR was very similar to that presented in panel a (not shown).

S1 (attacking SS regions), *Phy M* (cleaving after A and U), and from *B. cereus* (cleaving after U and C) as well as dimethyl sulfate served as probes for single-strandedness, while nuclease CV1 was a DS-specific reagent. Mapping of SS- and DS-specific signals gave results that were in reasonable accord with the predicted structures (Fig 2). CV1 nuclease-generated cleavages in the vicinity of the termination codons might be taken as evidence for the reality of helical element S. The scarcity of signals obtained from domain X was probably due to its close proximity to the 3′ end of the primer used for the analysis. The only apparent inconsistency between the experimental data and the anticipated structure concerned domain Z in the Cox-B3 3-UTR (Fig. 2b), on the *3′* branch of which several weak SS-specific signals were observed; due to the generation by this domain of strong DS-specific signals as well, we are inclined to consider the former as fortuitous ones, probably resulting from the destabilization of A-U- and G-U-rich hairpin Z following the cleavage of its loop.

Tertiary structure modeling

As shown in Fig 1, a highly conserved potential for inter-domain interactions (between domains X and Y) was observed (boldface letters). According to the secondary structure models proposed (Fig. 2), the potentially pairing nucleotides should be confined to the loops of the respective domains. Taking into account that the oligo(A) moieties of domains X and S are direct continuations of one another, these domains should very likely form a common stacked helical element. Combining these two considerations (the

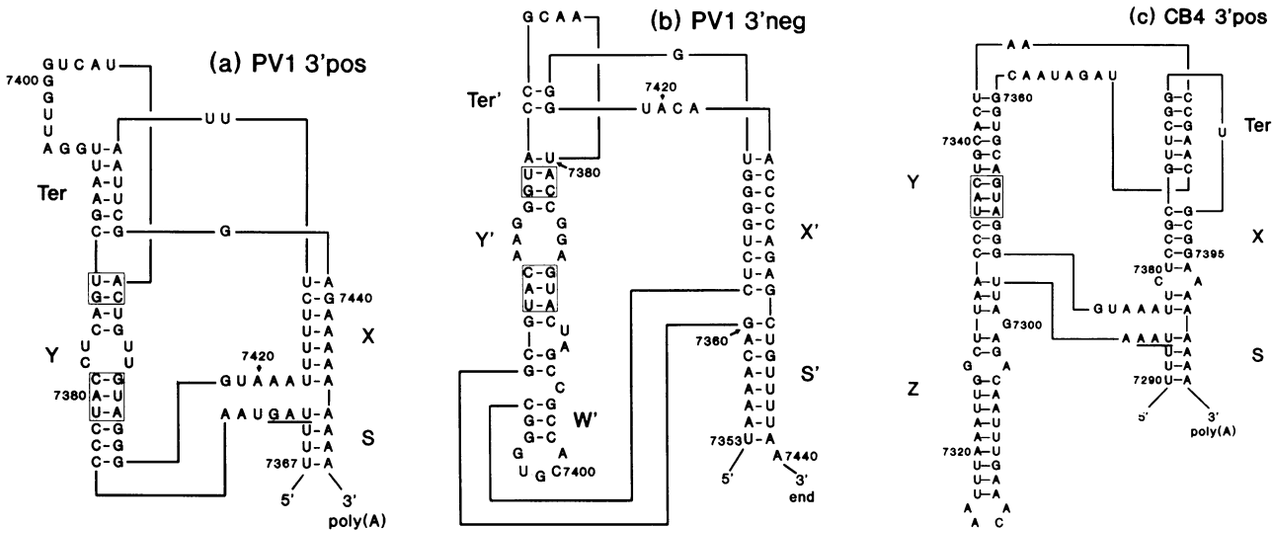

Fig. 3. Tertiary structure models for the 3-UTRs of the PV1 (+) (panel a) and (−) (panel b) RNA strands as well as of CoxB4 (+) RNA strands (panel c). The termination codons are underlined. The elements shared by the domains Y and Y′ of the 3-UTRs of the both (+) and (−) strands are boxed. For other details, see the text.

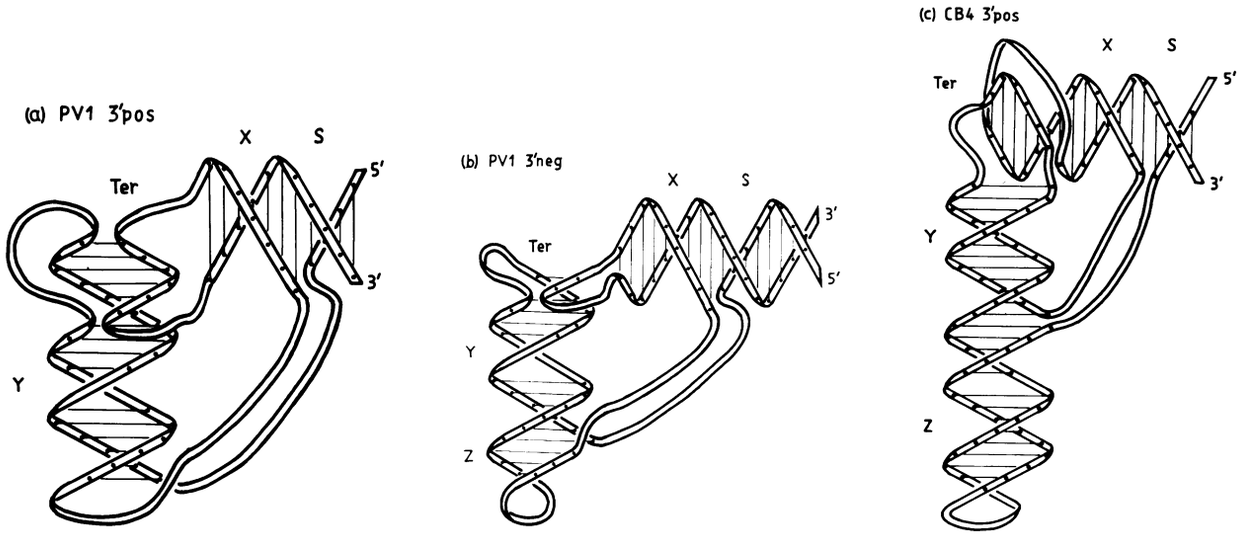

Fig. 4. Schematic representation of the spatial organization of the 3-UTRs of the PV1 (+) (panel a) and (−) (panel b) RNA strands as well as of CoxB4 (+) RNA strands (panel c). The proposed structures are based on those shown in Figs. 3 and 6.

interaction between loops of domains X and Y, on the one hand, and the stacking between domains X and S, on the other) we arrived at the model of the tertiary 3-UTR structure of the polio-like subset of enteroviruses shown in Fig. 3a. Since there is no intervening nucleotides between domains Y and Z in the CoxB-like subset of 3-UTRs, the stacking between the helical elements of the respective domains could also be envisioned (Fig. 3c). Furthermore, the nucleotides paired as a result of tertiary interactions (element Ter) could extend, by stacking, the stems of either domain Y or domain X. Since we had no reasons to prefer one variant of stacking to the other, we arbitrarily ascribed, solely for illustrative purposes, the former and the latter variants to the polio-like and CoxB-like structures, respectively (Fig. 3).

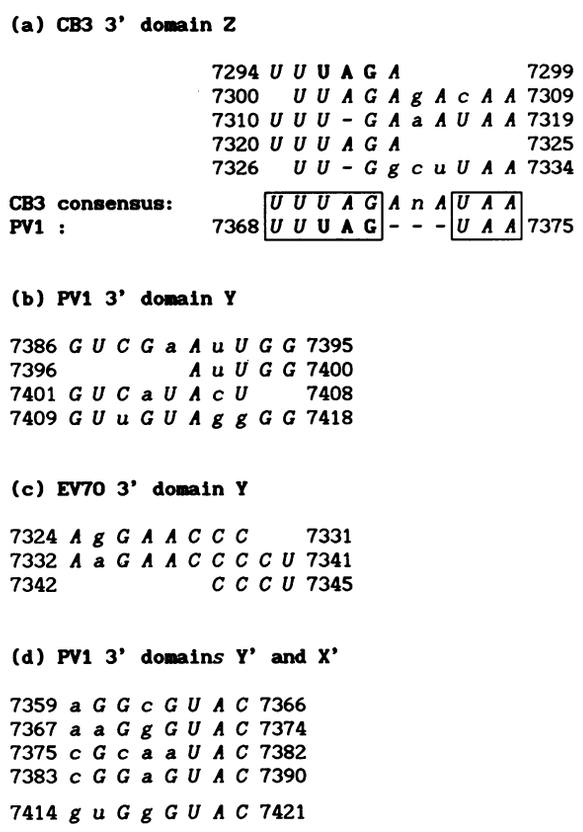

Fig. 5. Repeating motifs within (a) domain Z, as exemplified by the CoxB3 3-UTR; (b) domain Y, as exemplified by the PV1 3-UTR; (c) the insert found in domain Y of the EV70 3-UTR; and (d) domains X' (the bottom line) and Y', as exemplified by the PV1 (−) RNA strand. Nucleotides found at the given positions in at least 66% of the repeating units are printed in upper case letters. In panel (a), the stop codons are prined in bold, and elements shared by the repeating units of CoxB3 domain Z and the PV1 sequence around the terminator codon are boxed.

The structures we arrived at were strikingly similar, in their overall organization, to tRNA species (Fig. 4). This point will be discussed in more detail below.

Oligonucleotide repeats in the 3-UTRs

It has been shown previously that some cases of considerable length variations among the 5'-untranslated regions (5-UTRs) of the picornaviral genomes are due to duplications of nucleotide segments (18). We wondered whether a similar explanation did hold true for the 3-UTR length variability too. No evidence was found to suggest that the acquisition of novel domains was accomplished through duplications of extended segments, but short repetitions turned out to be quite abundant. Thus, entire domain Z in the CoxB-like RNAs could be regarded as composed of 3 direct incomplete repeats, in turn comprising 5 degenerate copies of the basic AU-rich repeating motif; this motif matched quite well a sigle copy element in the domain Z-lacking PV sequence around the terminator UAG (Fig. 5a). These observations hinted that domain Z in the CoxB-like genomes might originate from a burst of short repetitions. Likewise, a significant portion of domain Y could be viewed as made up of tandemly repeated oligonucleotides, even though they displayed a marked divergence (Fig. 5b). One may argue that the unique insert within domain Y of EV70 RNA (Fig. 1) is also a vestige of the multiplication of a short sequence (Fig. 5c). No clear-cut evidence for a similar origin of relatively small domain X could be found; nevertheless, it appeared to share a motif, AAUUxGG-A, with domain Y.

DISCUSSION

The structure of 3-UTRs in entero- and rhinovirus RNAs

Using a combined comparative and experimental approach we derived three consensus secondary structure models, for rhinoviruses, for a group of enteroviruses (PV, CoxA21, BEV, and EV70; polio-like), and for the rest of enteroviruses (Cox A9, CoxB1, CoxB3, CoxB4, SVDV, ECHO6 and ECHO11; CoxB-like). These models are primarily composed of one, two, and three hairpin-loop domains, respectively. Although our models share some features with certain individual computer-derived foldings reported by Auvinen and Hyypia for 5 viruses (15), the whole grouping and some specific examples are quite different.

Upon an attempt to define the tertiary structure of the respective 3-UTRs we arrived at an unexpected conclusion that these regions, in the enterovirus RNAs, appear to have a potential to acquire an L-shaped, or tRNA-like, overall organization. It seems appropriate to discuss the 3-UTR/tRNA similarity in greater detail.

We found it convenient to describe the enterovirus 3'-terminal structures as the following arrays of consecutive stretches, $S_5Y_5Ter_5Y_3X_5Ter_3X_3S_3$ and $S_5Z_5Z_3Y_5Ter_5Y_3X_5Ter_3X_3S_3$ for the

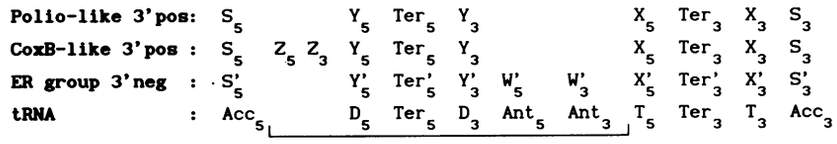

Fig. 6. Schematic representation of the structural organization of the 3-UTRs of polio-like and CoxB-like (+) RNA strands and of entero- and rhinovirus (ER) (−) RNA strands, in comparison with tRNA. The bracketed and non-bracketed stretches form two 'legs' of the L-shaped tRNA molecule, respectively; the same is true of their structural analogs at the termini of viral RNAs. See the text for explanations.

polio-like and CoxB-like RNAs, respectively (subscripts 5 and 3 correspond to the 5′-proximal and 3′-proximal branches of the appropriate DS helical element). Using the same convention, we could write the structure of a tRNA molecule as $Acc_5D_5Ter_5D_3Ant_5Ant_3T_5Ter_3T_3Ac_3$, where Acc, D, Ant and T stand for the acceptor, D, anticodon and T stems, respectively, and Ter designates the tertiary Watson-Crick interaction between the D and T loops. Then, the viral and tRNA segments could be aligned using the Ter elements as marking points (Fig. 6). Such an alignment allowed us to consider S, Y, and X elements in the viral 3-UTRs as structural analogs of the Acc, D, and T stems, respectively. In the framework of this reasoning, the 3-UTRs of the both polio-like and CoxB-like viral RNA species lacked an analog of the anticodon stem, the latter RNAs having an additional stem, Z, instead. It is noteworthy that these deviations of the viral structures from that of their tRNA counterpart do not affect essential features of the L-shaped organization. The reasons for maintaining the overall design are obvious, (i) extra-domain Z serves, by stacking on Y, merely to elongate one 'leg' of L (Fig. 3); and (ii) the absence of the anti-codon stem merely shortens its other 'leg'.

It should be admitted that the above tertiary structure models, as opposed to the secondary structure ones, are based on theoretical considerations only, and even though the evolutionary approach proved to be a powerful tool, more direct verification of the models is needed.

Similarity between the 3′-terminal structures of (+) and (−) strands

Designing the present study, we wanted to define common structural elements in the 3′-terminal regions of the (+) and (−) strands of picornaviral RNAs, as candidate *cis*-acting replication elements. Obviously, the structure of the 3′ end of the (−) strand should mirror that of the 5′ end of the (+) strand. Consensus cloverleaf secondary structure models for a ca. 100 nt-long region at the 5′ end of the (+) strand and, by implication, at the 3′ end of the (−) strand of entero- and rhinovirus RNAs have already been proposed on the basis of both comparative considerations (3) and experimental testing (4). Upon inspection of the appropriate structures, a conserved potential for tertiary interactions between two hairpin loops of the 3′-terminal cloverleaf on the (−) strand was found (illustrated for PV1 in Fig. 3b); a similar interaction at the 5′ end of the (+) strand is also possible (not shown). As a result of such interactions and of stacking of the domains in pairs, these termini could acquire an L-shaped conformation (Fig. 4b), strikingly similar to that proposed for the 3′ end of the (+) strand of both polio-like (Fig. 4a) and CoxB-like RNAs (Fig. 4c). Using the convention described above, we designated the duplex that joins the very 5′- and 3′-terminal segments of the cloverleaf on the 3′ end of the (−) strand [or the 5′ end of the (+) strand] as stem S′ (stem '*a*' in ref. 4) and the element formed by the hairpin-hairpin tertiary interaction as Ter′. The hairpin that generates a common stacked stem with S was termed X′ (stem-loop 'b'), and the hairpin that enters into the tertiary interaction with X′ was termed Y′ (stem-loop 'd'); the remaining small domain was termed W′ (stem-loop 'c') (it may be noted that domains Y and Y′ share common elements at the primary structure level; see boxed oligonucleotides in Fig. 3). Then, the structure of the 3′-terminal region of the entero- and rhinovirus RNA (−) strand could be written as follows 5′ $S'_5Y'_5Ter'_5Y'_3W'_5W'_3X'_5Ter'_3X'_3S'_3$ 3′ (Fig. 6). Accepting this, we could regard element W′ as an analog of the tRNA anticodon stem-loop. Again, no evidence is available to choose between the two possible variants of stacking of Ter′, either on X′ or on Y′. In either case, however, the stacking would be expected to stabilize the Ter′ element composed of two G-C pairs (further stabilization could in principle be achieved at the expense of partial unwinding of Y′).

Functional implications

A major aim of the present study was to get insights into the nature of the *cis*-acting elements serving as recognizable signals for the replication machinery. Two interrelated points deserve some considerations: they concern the extent of similarity (or diversity) of putative signals at the (i) 3′ ends of the (+) strands of different viruses; and (ii) 3′ ends of the (+) and (−) RNA strands of a given virus.

The extent of the variability of the 3-UTR structures among the entero- and rhinovirus RNA genomes is unexpected and intriguing. Since domain Y is shared by all the members of this virus group one may assume that it plays a major part in the template recognition, other domains accomplishing some auxiliary, e. g., stabilizing, function; it could be envisioned that such stabilization may be more important under certain conditions than under the others (due, for example, a higher optimal temperature for the reproduction of enteroviruses than of rhinoviruses). On the other hand, structural variability of the 3-UTRs of entero- and rhinoviruses may reflect peculiarities of initiation of the (−) strand synthesis in each viral subgroup; these peculiarities may or may not be related to differences in the natural host cells.

It is interesting to note that insertions of oligonucleotide linkers just between Y_5 and Ter_5 of poliovirus type 1 RNA (cf., Figs. 3a and 6) resulted in viable mutants with either wild-type or 3ts phenotypes (19). Remarkably, such insertions would not abolish the potential for the both secondary and tertiary structure formation, but the overall stability of the mutant *cis*-element may be diminished, especially in the case of the *ts* mutant (unpublished observations).

There is little doubt that the same enzyme, viral RNA polymerase, performs a key role in the synthesis of both (+) and (−) strands, though the possibility that the initiation of the two strands is accomplished by different mechanisms could not be ruled out (cf., 4). One might expect that the signals recognized by the enzyme on the 3′ ends of the both kinds of templates should resemble each other, at least to some extent. The results presented above demonstrate that this is very likely to be the case for the CoxB-like subgroup of enteroviruses; here, these signals appear to be contained within very similar L-shaped structures composed of three hairpin-loop elements in each case. In polio-like enteroviruses, the 3′-terminal structures of the (+) and (−) strands, while differing from one another in the number of the constituent hairpin-loop elements, may nevertheless also share a common L-shaped configuration. The 3′-terminal structures appear to be more divergent, however, in the (+) and (−) RNA strands of rhinoviruses, being represented by a solitary stem-loop and an L-shaped element, respectively. Although the functional significance of this divergence is uncertain, it should be noted that the overall structure of the 3′ end of the (−) strand is highly conserved among all the entero- and rhinoviruses, suggesting that all of them share a common mechanism for the (+) strand initiation, disregarding differences, if any, in the modes of the initiation of the (−) strand. While considering possible differences between utilization of (+) and (−) strands as templates, one

perhaps should take into account that the newly synthesized RNA polymerase molecule (polypeptide 3D) finds itself just in the close proximity to the 3′ terminal *cis*-element of the (+) strand; this circumstance may facilitate their mutual recognition.

An additional puzzle was raised by the results reported by Andino et al. (4) demonstrating that the synthesis of PV (+) RNA strands depends primarily on the secondary structure of the 5′ end of the (+) strand rather than on that of the 3′ end of the (−) strand. Due to the fact that all the proposed base pairs at the 3′ end of the PV (−) RNA strands are represented by standard A-Us and G-Cs (without participation of G-Us) (Fig. 3b), the secondary and tertiary structures of the 5′ end of the (+) strand should mirror those shown in Figs. 3b and 4b. Therefore all the considerations about the significance of structural peculiarities of the 3′ end of the (−) strands could safely be extrapolated to the 5′ end of the (+) strands. The only possible distinction may concern a stretch of mismatches within domain Y′: the mismatched bases are represented solely by purines at the 3′ ends of entero- and rhinovirus (−) RNA strands (Fig. 3b; positions 7367−7369 and 7384−7386), whereas pyrimidines should occupy this place at the (+) RNA 5′ (as well as 3′) ends . A conceivable consequence of this difference is that the stem of domain Y should be more stable in the latter cases compared to the former because pyrimidines could be more readily than purines incorporated into this stem by stacking interactions. This circumstance, in addition to the difference in the primary structure, may contribute to the preferential interaction of viral replication proteins with the 5′ ends of the (+) strands (4).

A tRNA-like structure of the template termini raises an interesting possibility that they could interact with tRNA-recognizing host proteins as is the case with some RNA phages (20) and plant viruses (21).

Evolutionary considerations

The 3-UTRs of entero- and rhinoviruses exhibit a marked variability. In rhinoviruses, this segment is considerably shorter than in enteroviruses, possessing only a single secondary structure domain (Y). It should be remembered that the rhinovirus 5-UTRs are also the shortest, lacking two duplications present in their enterovirus counterparts (18). The difference in the 3-UTR lengths could arise from either the acquisition of novel domains by enteroviruses, or the loss of some enterovirus-specific domains by rhinoviruses. Although no direct evidence is available to choose between these possibilities, we prefer the former one. Indeed, the presence of repeats suggests that at least one of the enterovirus extra 3′ domains, Z, has arisen through a burst of duplications of a short oligonucleotide adjoining the terminator codon. In the framework of this reasoning, the rhinovirus genome appears to be closer to the putative common rhino- and enterovirus ancestor, as far as the structures of the both 5- and 3-UTRs are concerned (see also, ref. 18).

In turn, the enterovirus 3-UTRs could be further divided into two subclasses, using the presence or absence of domain Z as a criterion. Interestingly, each subclass contains a CoxA representative. A close relatedness of the CoxA21 3-UTR structure to that of PV is in good accord with the relatedness of the protein-coding parts of the respective genomes (10). On the other hand, capsid polypeptides (like 3-UTRs) of CoxA9 appear to be more related to the CoxB viruses (22). Thus the current classification of some picornaviruses into the CoxA group does not appear to be supported by the primary and secondary structures of their genomic RNAs (cf., also ref. 14).

It is hardly fortuitous that the 3′-terminal organization of so different replicating molecules as genomes of picornaviruses (this study), many plant RNA viruses (21), RNA bacteriophages (23) and even so called small replicating RNA species found in phage Qβ infected *E. coli* cells (24) bear more or less obvious similarity to tRNA. This similarity may merely reflect some advantages of such an organization, e. g., its relative rigidity, but, in addition, it may have a more profound reason; indeed it was suggested that tRNA species have evolved from 3′-terminal 'tag' structures of self-replicating inhabitants of the prehistoric 'RNA world' (25). Whether or not there is any evolutionary link between the similar 5′- and 3′-terminal L-shaped structures (in other words, whether they came from a common ancestor RNA element), is another intriguing and completely open question.

Interestingly, one can find short direct repeats at the 3′ ends of not only the (+) but also of the (−) strands. Thus, nearly entire domain Y′ of all the entero- and rhinovirus RNAs is composed of 4 tandem copies of an octanucleotide with a highly conserved GUAC motif; domain X′ also contains a copy of this motif (Fig. 5d). Evolutionary and/or physiological significance of this fact is unknown.

REFERENCES

1. Koch,F. and Koch,G. (1985) The Molecular Biology of Poliovirus. Springer-Verlag, Wien.
2. Richards,O. and Ehrenfeld,E. (1990) *Curr. Top. Microb. Immun.* **161,** 189−119.
3. Rivera,V., Welsh,D. and Maizel,J. (1988) *Virology* **165,** 42−50.
4. Andino, R., Rieckhof, G. E., Trono, D. and Baltimore, D. (1990) *Cell* **463,** 369−380.
5. Pilipenko,E.V., Blinov,VM., Romanova,L.I., Sinyakov,A.N., Maslova,S.V. and Agol,V.I. (1989) *Virology* **168,** 201−209.
6. Pilipenko,E.V., Blinov,V.M, Chernov,B.K., Dmitrieva,T.M. and Agol,V.I. (1989) *Nucleic Acids Res.* **17,** 5701−5711.
7. Skinner,M.A., Raccaniello,V.R., Dunn,G., Cooper,J., Minor,P.D. and Almond,J.W. (1989) *J. Mol. Biol.* **207,** 379−392.
8. Le,S-Y. and Zuker,M. (1990) *J. Mol. Biol.* **216,** 729−741.
9. Chang,K.H., Auvinen,P., Hyypia,T. and Stanway,G. (1989) *J. Gen. Virol.* **470,** 3269−3280.
10. Hughes,P.J., North,C., Minor,P.D. and Stanway,G. (1989) *J. Gen. Virol.* **470,** 2943−2952.
11. Earle,J.A.P., Skuce,R.A., Fleming,C.S., Hoey,E.M. and Martin,S.J. (1988) *J. Gen. Virol.* **69,** 253−263.
12. Inoue,T., Suzuki,T. and Sekiguchi,K. (1989) *J. Gen. Virol.* **70,** 919−934.
13. Ryan,M., Jenkins,O., Hughes,P., Brown,A., Knowles,N., Booth,D., Minor,P. and Almond,J. (1990) *J. Gen. Virol.* **71,** 2291−2299.
14. Auvinen,P., Stanway,G. and Hyypia,T. (1989) *Arch Virol.* **104,** 175−186.
15. Auvinen,P. and Hyypia,T. (1990) *J. Gen. Virol.* **71,** 2133−2139.
16. Jenkins,O., Booth,J.D., Minor,P.D. and Almond,J.W. (1987)
17. Peritz,A.E., Kierzek,R., Sugimoto,N. and Turner,D.H. (1991) *Biochemistry* **30,** 6428−6436.
18. Pilipenko,E., Blinov,V. and Agol,V. (1990) *Nucleic Acids Res.* **18,** 13371−3375.
19. Sarnow,P., Bernstein, H.D. and Baltimore,D. (1986) *Proc. Natl. Acad. Sci.* **83,** 571−575.
20. Blumenthal,T. and Carmichael,G.G. (1979) *Annu. Rev. Bochem.* **48,** 1525−548.
21. Haenni,A-L., Joshi,S. and Chapeville,F. (1982) *Progr. Nucleic Acids Res. and Mol. Biol.* **27,** 85−104.
22. Stanway,G. (1990) *J. Gen. Virol.* **71,** 2483−2501.
23. Adhin,M., Alblas,J. and van Duin,J. (1990) *Biochim. Biophys. Acta* **1050,** 1110−118.
24. Chetverin,A.B. and Voronin,L.A. (1987) In: Structure and Biosynthesis of Proteins, vol. 1. Scientific Center for Biological Research, Pushchino, pp. 43−52 (in Russian).
25. Weiner,A.M. and Maizels,N. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 17383−7387.