Equine rhinovirus serotypes 1 and 2: relationship to each other and to aphthoviruses and cardioviruses

G. Wutz,1 H. Auer,1† N. Nowotny,2 B. Grosse,1 T. Skern1 and E. Kuechler1

1 Institute of Biochemistry, Medical Faculty, University of Vienna, Dr Bohr-Gasse 9/3, A-1030 Vienna, Austria
2 Institute of Virology, Veterinary University of Vienna, Linke Bahngasse 11, A-1030 Vienna, Austria

Equine rhinoviruses (ERVs) are picornaviruses which cause a mild respiratory infection in horses. The illness resembles the common cold brought about by rhinoviruses in humans; however, the presence of a viraemia during ERV-1 infection, the occurrence of persistent infections and the physical properties are all more reminiscent of foot-and-mouth disease virus (FMDV). cDNA cloning and sequencing of the genomes of ERV-1 and ERV-2 between the poly(C) and poly(A) tracts showed that the serotypes are heterogeneous. Nevertheless, the genomic architecture of both serotypes is most similar to that of FMDV. Indeed, a comparison of the derived protein sequences of ERV-1 shows that their identity is greatest to FMDV. In contrast, most ERV-2 proteins are more related to encephalomyocarditis virus (EMCV) proteins than they are to FMDV or ERV-1. These results place ERV-1 alongside FMDV in the aphthovirus genus of the picornavirus family and indicate that this virus may serve as a model system for examining the biology of FMDV.

Introduction

Picornaviruses are small animal viruses with a single-stranded RNA genome of positive polarity. The family includes a number of important human and animal pathogens, such as poliovirus, hepatitis A virus (HAV) and foot-and-mouth disease virus. Based on physico-chemical properties and host pathogenicity, the members are classified into five genera (Newman et al., 1973; Rueckert, 1990). These are the human rhinoviruses (HRVs), of which there are over 100 antigenically distinct serotypes, the murine cardioviruses [encephalomyocarditis virus (EMCV), mengovirus and Theiler’s murine encephalomyelitis virus (TMEV)], the human enteroviruses (which include polioviruses, coxsackie A and B viruses and echoviruses), hepatoviruses (HAV) and aphthoviruses (FMDV), of which there are seven known serotypes.

Author for correspondence: Tim Skern.
Fax +43 1 79 86 224. e-mail timothy.skern@univie.ac.at
† Present address: Department of Occupational Medicine, University of Vienna, Währinger Gürtel 18–20, A-1090 Vienna, Austria.

The nucleotide sequence data presented in this paper have been deposited in the EMBL database under accession numbers X96870 (ERV-1) and X96871 (ERV-2).
Fig. 1. For legend see opposite.
Fig. 1. Nucleotide sequence of ERV-1 cDNA between the poly(C) and poly(A) tracts. The derived amino acid sequence of the polyprotein and the proposed positions of the mature proteins are indicated. The four indirect repeats following the poly(C) tract are underlined. Viral proteins are named according to the nomenclature of Ruecke & Wimmer (1984).
Fig. 2. For legend see page 1724.
Fig. 2. For legend see page 1724.
Methods

ERV-1 (strain PERV) (Pluratten, 1962) and ERV-2 (strain P1436/71) (Steck et al., 1978), as well as reference antisera, were obtained from M. Weiss (Institute of Virology, University of Berne, Switzerland). Both ERV serotypes were plaque-purified, propagated in rabbit kidney cells (RK-13) and centrifuged on sucrose or cesium chloride gradients (Brown & Cartwright, 1963; Newman et al., 1973). The identity of the viruses was checked by immunoprecipitation with the reference antisera. DNA manipulations were carried out using standard techniques (Sambrook et al., 1989). The isolation of viral RNA and synthesis and cloning of cDNA was checked by immunoprecipitation with the reference antisera. DNA synthesis kit. Following ligation of ERV serotypes were plaque-purified, propagated in rabbit kidney cells (RK-13) and centrifuged on sucrose or cesium chloride gradients (Brown & Cartwright, 1963; Newman et al., 1973). The occurrence of repeated sequences in the ERV-1 genome was verified by direct sequencing of the genomic viral RNA (Fichot & Girard, 1990). Primer extension on the ERV-1 genomic RNA was performed with oligonucleotides complementary to nt 5'--3' and nt 77--93 using Superscript reverse transcriptase (Life Technologies). Positions of the primer extension primers were determined by comparison with the derived amino acid sequences of other picornaviruses using the Staden programs (Dear & Staden, 1991) and the GAP program of the Genetics Computer Group, Wisconsin. The EMBL and SwissProt or PIR identifiers of the picornavirus sequences used are: FMDV, pifmdv2.em_vi, polg_fmdv2.swiss; EMCV, pimecmv.em_vi, polg_emcv.swiss; Echovirus 1 Mahoney, pol_em_vi, polg_pollm.swiss; Echovirus 2, piemcvpolyp.em_vi, polg_emcvpolyp.swiss; TMEV, tmepp.em_vi, polg_tmevb.swiss; Poliovirus 1, poll.em_vi, polg_poll1.swiss; Poliovirus 2, polg_hrvl4.swiss; Poliovirus 3, polg_hpav2.swiss; Echovirus 22, polg_hrv22.swiss; Echovirus 9, polg_hrv9.swiss; Echovirus 14, polg_hrv14.swiss; HRV-14, pihrv14.em_vi, echpicorn.em_vi, a46182.pir.

Results and Discussion

Genome organization in equine rhinoviruses

A total of 7734 and 8828 non-homopolymeric nucleotides was cloned and sequenced for ERV-1 and ERV-2, respectively (Figs 1 and 2). A poly(A) tract was present at the 3' end of the cloned sequence. Clones at the 5' end of both ERV-1 and ERV-2 commenced with a run of C residues (up to 17), implying the presence of a homopolymeric C region. Such sequences have been previously reported in aphtho- and cardioviruses (Duke et al., 1990; Rowlands et al., 1978); they separate the main body of the genome (the large (L) fragment) from a smaller one (the S fragment). All attempts to derive sequence information on the approximately 300 nt (as estimated by primer extension in ERV-1) 5' to the poly(C) region have as yet been unsuccessful. The sequences reported here therefore correspond to the L fragments. The start of protein synthesis was predicted by searching for polyuridylic stretches preceding an AUG codon which allowed translation into a large open reading frame. Such AUG codons were found in ERV-1 (nt 881) and in ERV-2 (nt 895), although in each case an unusual aspect was noted. For ERV-1, AUG-881 was immediately followed by a second AUG codon and by a pair of AUG codons 63 nt downstream. This situation is similar to that in FMDV, in which two single AUG codons are also present within 84 nt of each other at the start of the open reading frame; both are used (Sangar et al., 1987). For ERV-1, the second AUG of each pair possessed the better Kozak (AXXAGUGG) sequence.

For ERV-2, two further in-phase codons follow AUG-895 at 904 and 916, with the second having the best Kozak sequence. Again, however, in the absence of any protein sequencing data, we assume that protein synthesis initiates at AUG-895.

The 5' and 3' untranslated regions

The poly(C) tract and the initiating AUG are separated in ERV-1 and ERV-2 by almost 200 nt more than in FMDV (880 nt).
and 894, compared to 714 nt). Immediately following the poly(C) tract in all FMDV serotypes are three to four indirect repeats of 30–31 nt separated by a spacer of 12 nt. Four similar indirect repeats (marked R1–R4 in Fig. 1) are present in the genome of ERV-1 but not in ERV-2. The identity between the repeats of ERV-1 and FMDV serotype O1k is restricted to the hexanucleotide GTAAAA. The function of these sequences is not clear. In FMDV, they have been proposed to be involved in pseudoknot structure formation (Clarke et al., 1987).

The 400 nt of the 5'UTR of picornaviruses preceding the initiating AUG codon are known to fold into a higher ordered RNA structure known as the internal ribosome entry site (IRES) to which 40S ribosomal subunits bind prior to initiating protein synthesis. Three types of RNA folding patterns in IRES structures have been determined, those in aphthoviruses and cardioviruses being classed as type I. IRESs of human rhinoviruses and enteroviruses are classed as type II and that of HAV as type III (Jackson et al., 1990). Analysis of the two ERV IRESs showed that their RNA folding patterns are of type I (data not shown). The two sets of initiating AUGs in ERV-1 are 23 and 86 nt distant from the start of the essential cis-acting oligopyrimidine tract. In ERV-2, this distance is 24 nt.

The 3'UTRs of ERVs are noteworthy for the unusual length in ERV-2 (667 compared to 110 and 92 nt in ERV-1 and FMDV, respectively) and the high U content (55 %) in ERV-1. The highest level of identical nucleotides with other picornaviruses was 43 (one gap was necessary) for ERV-1 with ERV-2 and 51 (again one gap was required) for ERV-2 with EMCV.

### Analysis of the protein coding regions

The ERV serotypes 1 and 2 code for polyproteins of 2246 and 2589 amino acids. The identification of the mature viral proteins (marked in Fig. 1 and Fig. 2) and the prediction of cleavage sites were based on the homology with other picornaviruses. The requirement for a tyrosine residue as the third amino acid helped to locate protein 3B (VPg, virus protein genome-linked). The cleavage sites appear to be heterogeneous, although a majority possess either a glutamate or glutamine at P1 (N-terminal to the cleavage site) and a glycine, serine, threonine, valine or asparagine at P1' (C-terminal to the cleavage site). In general, the cleavage sites are reminiscent of those of aphtho- and cardioviruses. To investigate the relationship of the two ERV serotypes to each other and to other picornaviruses, the derived protein sequences were compared with each other and with those of representatives of other picornavirus genera (Table 1a, b).

### Equine rhinovirus 1

For ERV-1, the highest identity was obtained with the proteins of FMDV and EMCV. Values with poliovirus, HRV-14 and HAV were much lower. Significantly, there is 32% identity with the leader (L) protein of FMDV, but none with the L protein of EMCV. As the L protein is the first to be synthesized, the presence of two possible AUG initiation codons in FMDV results in two forms of the L proteinase in the infected cell, known as Lab and Lb (Sangar et al., 1987). The

### Table 1. Percentage identity between derived amino acid sequences of (a) ERV-1 and (b) ERV-2 proteins and those of selected picornaviruses

Values were calculated using the GCG GAP program.

<table>
<thead>
<tr>
<th>Virus</th>
<th>L</th>
<th>VP4</th>
<th>VP2</th>
<th>VP3</th>
<th>VP1</th>
<th>2A</th>
<th>2B</th>
<th>2C</th>
<th>3A</th>
<th>VPg*</th>
<th>3C</th>
<th>3D</th>
</tr>
</thead>
<tbody>
<tr>
<td>FMDV</td>
<td>32</td>
<td>59</td>
<td>35</td>
<td>39</td>
<td>33</td>
<td>88</td>
<td>36</td>
<td>47</td>
<td>24</td>
<td>33</td>
<td>40</td>
<td>51</td>
</tr>
<tr>
<td>EMCV</td>
<td>–</td>
<td>45</td>
<td>33</td>
<td>40</td>
<td>24</td>
<td>–</td>
<td>18</td>
<td>38</td>
<td>22</td>
<td>25</td>
<td>27</td>
<td>47</td>
</tr>
<tr>
<td>Poliovirus</td>
<td>–</td>
<td>26</td>
<td>32</td>
<td>29</td>
<td>19</td>
<td>–</td>
<td>15</td>
<td>34</td>
<td>20</td>
<td>35</td>
<td>19</td>
<td>32</td>
</tr>
<tr>
<td>HRV-14</td>
<td>–</td>
<td>33</td>
<td>31</td>
<td>26</td>
<td>22</td>
<td>–</td>
<td>27</td>
<td>33</td>
<td>17</td>
<td>22</td>
<td>25</td>
<td>33</td>
</tr>
<tr>
<td>HAV</td>
<td>–</td>
<td>22</td>
<td>28</td>
<td>18</td>
<td>–</td>
<td>20</td>
<td>20</td>
<td>17</td>
<td>38</td>
<td>23</td>
<td>29</td>
<td>–</td>
</tr>
</tbody>
</table>

(a) ERV-1 proteins

<table>
<thead>
<tr>
<th>Virus</th>
<th>L</th>
<th>VP4</th>
<th>VP2</th>
<th>VP3</th>
<th>VP1</th>
<th>2A</th>
<th>2B</th>
<th>2C</th>
<th>3A</th>
<th>VPg*</th>
<th>3C</th>
<th>3D</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERV-1</td>
<td>23</td>
<td>47</td>
<td>38</td>
<td>40</td>
<td>27</td>
<td>81</td>
<td>24</td>
<td>38</td>
<td>22</td>
<td>33</td>
<td>29</td>
<td>42</td>
</tr>
<tr>
<td>FMDV</td>
<td>18</td>
<td>41</td>
<td>37</td>
<td>33</td>
<td>26</td>
<td>88</td>
<td>19</td>
<td>37</td>
<td>18</td>
<td>25</td>
<td>21</td>
<td>40</td>
</tr>
<tr>
<td>EMCV</td>
<td>–</td>
<td>40</td>
<td>40</td>
<td>42</td>
<td>25</td>
<td>–</td>
<td>23</td>
<td>40</td>
<td>19</td>
<td>16</td>
<td>21</td>
<td>41</td>
</tr>
<tr>
<td>Poliovirus</td>
<td>–</td>
<td>21</td>
<td>27</td>
<td>32</td>
<td>26</td>
<td>–</td>
<td>19</td>
<td>32</td>
<td>15</td>
<td>19</td>
<td>23</td>
<td>33</td>
</tr>
<tr>
<td>HRV-14</td>
<td>–</td>
<td>18</td>
<td>32</td>
<td>32</td>
<td>25</td>
<td>–</td>
<td>17</td>
<td>32</td>
<td>19</td>
<td>15</td>
<td>23</td>
<td>35</td>
</tr>
<tr>
<td>HAV</td>
<td>–</td>
<td>–</td>
<td>28</td>
<td>24</td>
<td>22</td>
<td>–</td>
<td>15</td>
<td>29</td>
<td>21</td>
<td>19</td>
<td>23</td>
<td>26</td>
</tr>
<tr>
<td>Echovirus 22</td>
<td>–</td>
<td>15</td>
<td>26</td>
<td>22</td>
<td>21</td>
<td>–</td>
<td>21</td>
<td>30</td>
<td>23</td>
<td>25</td>
<td>21</td>
<td>28</td>
</tr>
</tbody>
</table>

(b) ERV-2 proteins

* The value for FMDV protein Vpg is for Vpg1; the values for Vpg2 and Vpg3 are 17% and 34% (compared with ERV-1) and 44% and 25% (compared with ERV-2), respectively.
Fig. 3. Alignment of the derived amino acid sequences of VP1 from ERV-I, ERV-2, FMDV, HRV-14 and mengovirus. Alignments are based on the identification of structurally equivalent residues as determined for FMDV, mengovirus and HRV-14 by Acharya et al. (1989). Secondary structure elements are indicated; positions of three or more identical residues are shaded. Key: (#) indicates that the first amino acid of HRV-14 shown is 33; (*) Indicates the presence of a further 33 C-terminal amino acids for ERV-2.
presence of the two pairs of initiating codons in ERV-1 thus implies Lαb and Lβb forms. The L protein of FMDV has been shown to be a proteinase, cleaving between its own C terminus and the N terminus of VP4 to free itself from the growing peptide chain (Strebel & Beck, 1986). The L proteinase also cleaves the cellular protein eIF4G (eukaryotic initiation factor 4G, formerly referred to as p220 or eIF-4G) (Devaney et al., 1988; Kirchweger et al., 1994). A conserved cysteine–tryptophan amino acid pair and a histidine residue are essential for activity (Gorbalenya et al., 1991; Piccione et al., 1995; Roberts & Belsham, 1995). These residues are conserved within the sequence of the L protein of ERV-1, implying that it possesses proteolytic activity. Support for this is provided by cleavage of eIF4G during ERV-1 replication in RK-13 cells (data not shown).

The nature of the 2A polypeptide region is also diagnostic for picomavirus genera; in HRVs, human enteroviruses and cardioviruses, this polypeptide comprises 140–160 amino acids. In aphthoviruses, however, it has only 16 amino acids and cleaves autocatalytically at its own C terminus between the glycine and proline of the sequence NPGP (Ryan et al., 1991). The ERV-1 2A polypeptide also has 16 amino acids. 14 of which (including the NPG motif) are identical to those of FMDV.

ERV-1 does not, however, possess two other aphthovirus characteristics; these are the presence of three coding sequences for VPg and the amino acid triplet RGD (arginine–glycine–aspartic acid) in VP1. This sequence mediates the binding to the cellular receptor of the virus, which has been recently identified as a member of the integrin family (Berinstein et al., 1985; Fox et al., 1989; Mason et al., 1995). None of the ERV-1 capsid proteins contain an RGD triplet and only one VPg sequence can be identified (Fig. 1).

Analysis of the capsid proteins

Analysis of the three-dimensional structures of picomavirus capsids has shown that each capsid protein has an eight-stranded β-barrel structure; the different biological properties are conferred by the sequences (known as loops) linking the β-strands. To investigate the relationship between the secondary structure of the capsid proteins of ERV-1 and other picornaviruses, alignments based on structurally equivalent residues were made. Using the three-dimensional structure of the virus capsids of FMDV, mengovirus and HRV-14, Acharya et al. (1989) produced structural alignments of three of the capsid proteins. These served as templates to produce an accurate alignment of ERV-1 capsid proteins; in Fig. 3, the putative positions of eight β-strands and two α-helices of VP1 are indicated. Their sequence and length is closest to those of VP1 and FMDV, indicating a similar overall architecture of the capsid protein. However, the amino acid sequences joining these secondary structure elements are generally longer in ERV-1 than in FMDV; a notable exception is, however, the βG–βH loop of FMDV which, in addition to containing the RGD sequence for recognizing the cellular receptor, represents the main antigenic determinant of FMDV (Rowlands et al., 1983). This loop in ERV-1 is 11 amino acids shorter and does not contain the RGD sequence as mentioned above. Despite a similar basic scaffold, therefore, antigenicity and receptor recognition in the two viruses must differ. Secondary structure elements for ERV-1 VP2 and VP3 could also be identified in this way. Only one difference was noted between ERV-1 and FMDV; loop βE–αB in VP2 of ERV-1 is extended by 11 amino acids compared to FMDV (data not shown).

Equine rhinovirus 2

Analysis of the derived amino acid sequences of the ERV-2 proteins also revealed the presence of a leader protein and a 2A sequence as found in FMDV and ERV-1 (Table 1b). However, the identity of the L protein of ERV-2 with those of FMDV and ERV-2 was low. Nevertheless, the positions of the cysteine–tryptophan pair and of the characteristic histidine residue are conserved, indicating that it presumably exhibits proteinase activity (Fig. 4). Once again, the cleavage of eIF4G during ERV-1 replication strengthens this idea (data not shown). The 2A region of ERV-2 also comprises 16 amino acids, 14 of which (including the NPG sequence) are identical to FMDV.

However, apart from the 2A proteinase, VP3, VP4 and 3D, the identity between ERV-2 proteins and those of FMDV or ERV-1 was less than 40%. In addition, the ERV-2 proteins 2B, 3A and 3C comprise considerably more amino acids than their ERV-1 counterparts (147, 20 and 47, respectively). As the number of amino acids in 2B seemed exceptional, PCR amplification was used to eliminate a possible cDNA cloning artifact. PCR fragments were amplified from the 2B region directly from the ERV-2 cDNA and from a cDNA clone; the fragments were identical (data not shown), indicating that the cDNA clones were a true reflection of the RNA genome.

The alignment of ERV-2 VP1 (based on structurally equivalent residues) is similar to that of FMDV and ERV-1 (Fig. 3), although the βH–βI loop and the C terminus are unusually long in ERV-2. Once again, the βG–βH loop is smaller than that of FMDV and no RGD sequence is present. The secondary structure elements in VP2 and VP3 could also be readily identified (data not shown). As with ERV-1, the βE–αB loop [known as the ‘puff’ in HRV-14 (Rossmann et al., 1985)] is extended in ERV-2 to comprise 39 amino acids and is thus similar to that in mengovirus.

Further examination of the data presented in Table 1(b) revealed that the identity of several ERV-2 proteins with those of EMCV was as great as or actually greater than with ERV-1 or FMDV. This is the case for the capsid proteins VP2 and VP3 and the non-structural proteins 2B, 2C and 3D. This suggested a closer relationship between ERV-2 and EMCV than between ERV-2 and FMDV, despite the presence of the above-mentioned characteristic L and 2A proteins. To analyse this
Fig. 4. Alignment of the amino acid sequences of the leader (1) proteins of FMDV O1k, ERV-1 and ERV-Z. Initial alignments were generated with the PILEUP program and were then optimized by editing with LINEUP. Also shown are the first five amino acids of the respective VP4. As this sequence (GAGXS) is the myristylation site (Chow et al., 1987), it is highly conserved and was used to define the C terminus of the L proteins.
point further, dendrograms were constructed (Fig. 5) using the amino acid sequence of the 2C (Fig. 5a) and VP2 (Fig. 5b) polypeptides to illustrate picornaviral relationships (Hyypiä et al., 1992). Fig. 5 shows that ERV-1 is clearly located in the aphthovirus branch; in contrast, ERV-2 leaves the aphthovirus/cardiovirus branch before these two separate from each other. Very similar dendrograms were also generated with amino acid sequences of the 3CD polypeptides and with other phylogeny programs (data not shown). These data show that ERV-1 and ERV-2 differ considerably from each other and that ERV-1 is more closely related to FMDV than ERV-2.

In this regard, it is noteworthy that ERV-2 appears to contain a pseudo-VPg sequence of 18 amino acids just preceding the true VPg sequence (nt 6385–6438). Identity to the true VPg sequence is limited to the sequence around the tyrosine residue (VYNA compared to AYNI) and is low compared to other VPg sequences (data not shown), but such a sequence is not present in ERV-1 or in any other picornaviruses except FMDV. In the absence of protein data, given the low levels of identity and the lack of a proper cleavage site, we have considered these amino acids as part of 3A, although the existence of a second VPg cannot be excluded based on present data.

Concluding remarks

The work reported here clearly shows a low level of identity between the corresponding proteins of the two ERV serotypes. This is in contrast to the relationships of serotypes from other picornaviruses within a genus; for example, poliovirus serotypes are quite closely related, as are those of FMDV. Even amongst the more variable HRVs, values of 70–80% identity between HRV-89 and HRV-2 and 34–60% between HRV-2 and HRV-14 have been found (Duechler et al., 1987; Skern et al., 1985). Indeed, identities between the proteins of HRVs and poliovirus (HRV-2 and poliovirus, 30–56%; HRV-14 and poliovirus, 41–65%) are higher than between the two ERV serotypes (see Fig. 5).

The above comparisons stress the importance of including sequence and structural data in classifying viruses. The type of proteins encoded in the genetic information of the two ERV serotypes demonstrate that they are more closely related to FMDV, although they are more akin to human rhinoviruses with regard to the disease symptoms. However, the identity of the ERV proteins with those of FMDV was only around 40% and a significant identity with EMCV was registered for ERV-2.

ERV-1 represents the first low pathogenic member of the aphthovirus family and may therefore serve as a model for FMDV which can be manipulated in laboratories not equipped with high-level containment facilities. It will be of interest to examine which of the different molecular characteristics are responsible for the biological properties of ERV and FMDV.

This work was supported by the Austrian Ministries of Science, Research and the Arts and of Agriculture and Forestry. We thank M. Weiss (Berne) for virus and antisera, J. Kolodziejek, S. Kühtreiber-Poech, H. Lussy and K. Walk for technical assistance, N. Knowles, R. Jackson and D. Blaas for helpful discussions.

References


Berinstein, A., Roivainen, M., Hovi, T., Mason, P. W. & Baxt, B. (1995). Antibodies to the vitronectin receptor (integrin αvβ3) inhibit binding and


Received 12 February 1996; Accepted 19 April 1996