Alignment of capsid protein VP1 sequences of all human rhinovirus prototype strains: conserved motifs and functional domains

Pia Laine,1 Soile Blomqvist,1 Carita Savolainen,1 Koen Andries2 and Tapani Hovi1

1Enterovirus Laboratory, Department of Viral Diseases and Immunology, National Public Health Institute (KTL), Mannerheimintie 166, 00300 Helsinki, Finland
2Johnson & Johnson Pharmaceutical Research and Development, Beerse, Belgium

An alignment was made of the deduced amino acid sequences of the entire capsid protein VP1 of all human rhinovirus (HRV) prototype strains to examine conserved motifs in the primary structure. A set of previously proposed crucially important amino acids in the footprints of the two known receptor molecules was not conserved in a receptor group-specific way. In contrast, VP1 and VP3 amino acids in the minor receptor-group strains corresponding to most of the predicted ICAM-1 footprint definitely differed from those of the ICAM-1-using major receptor-group strains. Previous antiviral-sensitivity classification showed an almost-complete agreement with the species classification and a fair correlation with amino acids aligning in the antiviral pocket. It was concluded that systematic alignment of sequences of related virus strains can be used to test hypotheses derived from molecular studies of individual model viruses and to generate ideas for future studies on virus structure and replication.

INTRODUCTION

Human rhinoviruses (HRVs), like their very close relatives the human enteroviruses (HEVs), belong to the large family Picornaviridae, which comprises nine genera altogether (King et al., 2000). The genus Rhinovirus is composed of 100 designated serotypes, including two antigenically distinct subtypes of one serotype (HRV1A and HRV1B). According to the HRV prototype sequence data on the VP4/VP2-coding genomic region, 74 serotypes (including HRV1A and HRV1B) belong to the species Human rhinovirus A (HRV-A) and 25 to Human rhinovirus B (HRV-B) (Horsnell et al., 1995; Savolainen et al., 2002). One serotype, HRV87, was found to be related closely to enterovirus 68 and is thus a member of the species Human enterovirus D (Blomqvist et al., 2002; Savolainen et al., 2002; Oberste et al., 2004). Recent phylogenetic analyses of the complete VP1 (Ledford et al., 2004; Laine et al., 2005)- and partial non-structural 2A (Laine et al., 2005)- and 3D (Savolainen et al., 2004)-coding regions of the viral genome have shown the same clustering pattern of the serotypes.

The HRV genome consists of a single-stranded mRNA of about 7200 nt, which is translated in the host-cell cytoplasm to a precursor polypeptide of about 2200 aa (Rueckert, 1996). Like the other picornaviruses, the small (∼30 nm), icosahedral rhinovirus capsid is composed of 60 structural subunits, protomers, that each contain one copy of the four capsid proteins VP1–VP4. In HRV14, VP1 contains 290 aa, VP2 262 aa, VP3 236 aa and VP4 69 aa (Stanway et al., 1984). The external capsid surface is composed of parts of the three larger proteins VP1–VP3, whilst the smaller VP4 is located at the inside of the capsid shell. High-resolution three-dimensional structures of many HEVs, such as poliovirus 1 strain Mahoney (Hogle et al., 1985), and several HRVs (Kim et al., 1989; Zhao et al., 1996; Hadfield et al., 1997; Kolatkar et al., 1999; Verduguer et al., 2000) have been determined by X-ray diffraction and/or cryo-electron microscopy, and they have indicated a remarkable structural similarity. VP1 is the most external and dominant of all picornaviral surface proteins and therefore contains most of the motifs known to interact with cellular receptors and neutralizing mAbs (reviewed by Rueckert, 1996).

In addition to the division of HRVs into the two species HRV-A and HRV-B, they have also been classified into major and minor groups according to receptor usage. Members of the major-receptor group of HRVs, initially reported to comprise 90 serotypes, use the intercellular adhesion molecule 1 (ICAM-1) as receptor (Abraham & Colonno, 1984; Uncapher et al., 1991; Olson et al., 1993), whilst HRV1A, HRV1B and serotypes 2, 29, 30, 31, 44, 47, 49 and 62, comprising the classical major-minor receptor group, use members of the low-density lipoprotein receptor (LDL-R) family (Mischak et al., 1988; Hofer et al., 1994; Marlovits et al., 1998; Hewat et al., 2000). HRV serotype 87, now known to belong to the species HEV-D (Blomqvist et al., 2002;
Savolainen et al., 2002; Oberste et al., 2004), does not use either of these receptors (Uncapher et al., 1991). ICAM-1 and poliovirus receptors are members of the immunoglobulin superfamily (IgSF) and binding of these receptors to the virion surface destabilizes the virus and initiates the uncoating process, whilst non-IgSF receptors, like LDL-R, do not induce such actions (Prchla et al., 1994; Rossmann et al., 2002). Interestingly, two HRV serotypes classified in the major-receptor group, HRV23 and HRV25, were not inhibited by a soluble form of ICAM-1 that interfered with virus–receptor interaction of all of the other ICAM-1-using HRV serotypes (Crump et al., 1993). The interactions of some HRV serotypes, including HRV2 (Verdaguer et al., 2000, 2004), HRV14 (Kolatkar et al., 1999) and HRV16 (Hadfield et al., 1997; Kolatkar et al., 1999), with fragments of the corresponding receptor have been studied in atomic detail. These studies support the earlier hypothesis (Rossmann & Palmenberg, 1988) that the long, thin ICAM-1 molecule binds to a narrow canyon surrounding the icosahedral fivefold axis of the HRV surface, involving interaction with one or two neighbouring protomers (Olson et al., 1993; Kolatkar et al., 1999). Amino acids at the canyon floor of HRVs have been reported to be more conserved than those at the rim of the canyon and other viral surfaces, which are more prone to substitutions (Rossmann et al., 1985; Rossmann & Palmenberg, 1988; Bella & Rossmann, 1999).

In contrast to ICAM-1, the minor-group receptor binds to the star-shaped dome on the fivefold axis of the virion (Hewat et al., 2000). The BC and HI loops in the north side of the canyon of the VP1 protein cover the majority of the footprint of the minor-group receptor molecule (Hewat et al., 2000; Neumann et al., 2003; Vlasak et al., 2003). A lysine in the HI loop (K224 in HRV2) was found to be fully conserved among the minor-group representatives, but was not specific for the minor group, and the overall surface potential of the surrounding receptor footprint was considered to be important for binding (Vlasak et al., 2003).

On the basis of inhibition of replication of all 100 HRV serotypes by several chemically different compounds, the genus Rhinovirus was further divided into antiviral groups A and B, which also appeared to be genetically different (Andries et al., 1990). Two-thirds of HRV serotypes, including all of the LDL-R-utilizing serotypes, were found to belong to antiviral group B, and the remaining serotypes to antiviral group A. The binding site of an antiviral compound to the viral capsid of HRV14 has been shown to correspond to a hydrophobic pocket lying beneath the canyon floor in the centre of the VP1 β-barrel (Andries et al., 1989, 1990). This pocket is either empty or occupied by a ‘pocket factor’, and binding of an antiviral agent to the pocket is considered to stabilize the virus against conformational changes induced by, for example, receptor binding (Chapman et al., 1991; Rossmann et al., 2002). Ledford et al. (2004) recently discussed the variation of amino acid residues lining the pocket, with special reference to sensitivity to one antiviral compound, pleconaril.

In the process of generating the complete VP1 sequences of all HRV prototypes (Laine et al., 2005), we also amplified and sequenced the 3' end of the VP3-coding part of the genome, which is included, together with the entire VP1 sequences, in the current analysis of conservation of deduced amino acid sequences in previously proposed functional domains of the proteins.

**METHODS**

**Deduced amino acid sequence accession numbers and secondary-structure PDB codes.** GenBank accession numbers for the previously submitted (Laine et al., 2005) partial VP3 and complete VP1 sequences of 96 serotypes are AY450462–AY450549 and AY458604–AY458611, and those of the entire genomes of HRV1B, HRV2, HRV14, HRV16 and HRV89 are D00239, X02316, L05355, L24917 and M16248, respectively. Protein Database (PDB) identification codes for the obtained sequence and secondary structures were HRV1A, 1RI; HRV2, 1FPN; HRV3, 1RH; HRV14, 4RH; HRV16, 1AYM.

The 96 recently sequenced HRV prototype strains were originally obtained from different sources (the National Institute for Public Health and the Environment, the Netherlands; Janssen Pharmaceuticals, Beerse, Belgium; ATCC, Rockville, USA; and the Haartman Institute, University of Helsinki, Finland) and were passaged in HeLa-Ohio cell culture at 33 °C as described previously (Savolainen et al., 2002; Laine et al., 2005). Viral RNA had been extracted from infected cell cultures by using an RNAasy Total RNA kit (Qiagen). RT-PCR followed by nucleotide sequencing were as described previously (Laine et al., 2005).

**Sequence analysis.** Deduced VP1 amino acid sequences were aligned by using the CLUSTAL_X (version 1.83) multiple sequence-alignment program (Thompson et al., 1997) and shading of alignments, indicating different degrees of similarity, was generated with the GeneDoc program (version 2.6.002) (Nicholas et al., 1997). ‘Conserved motifs’ within and between both HRV species were defined arbitrarily as three or more consecutive amino acid containing either fully conserved (100 % identity) and/or similar amino acid residues (Thompson et al., 1997). The latter were defined by observing that variation of amino acids at a given column in the alignment was limited to one of the following groups of amino acids, referred to by their single-letter codes: STA, NEQK, NHQK, NDEQ, QHRK, ILMV, MILF, HY or FYW.

For analysis of amino acid variation in the assumed ICAM-1 footprint on the virion surface, receptor-interacting canyon amino acid residues in the carboxy-terminal VP3 and the entire VP1 region of HRV serotypes 16 and 14 (Kolatkar et al., 1999) were taken as ‘templates’ for HRV-A and HRV-B strains, respectively. Corresponding amino acid residues of all HRV serotypes in the alignment were included in the analysis, comprising examination of conservation of individual positions in the alignment and a neighbour-joining phylogenetic analysis of the selected amino acid strings by using the CLUSTAL_X program.

Co-segregation of phylogenetic grouping and antiviral sensitivity of HRV strains was re-examined by using previously obtained data on antiviral sensitivity (Andries et al., 1990). MIC data obtained for 15 capsid-binding antivirals against all HRV prototype strains were used to compute interactions between and relative positions of each strain and antiviral compound in a three-dimensional space, applying hierarchical and nearest-neighbour cluster analysis (Lewi, 1980).
two-dimensional plot of this Spectral Map Analysis was used for graphical demonstration of the observed relative differences in sensitivities of rhinovirus serotypes (Andries et al., 1990).

In the current study, the above three-dimensional dataset was visualized by using the Scatter3D program (http://www.st-andrews.ac.uk/~wjh/scatter3d/). This program allows viewing of a three-dimensional plot from any direction.

**RESULTS**

Alignment of deduced VP1 amino acid sequences

In order to study conserved motifs in the deduced VP1 amino acid sequences of HRV prototype strains, the generated VP1 sequences were aligned with CLUSTAL_X followed by manual editing. As reported before (Ledford et al., 2004; Laine et al., 2005), HRV-A showed more within-species sequence variation than HRV-B (Fig. 1). Within both species, the major amino acid sequence and domain-length differences were accommodated in the BC and DE loops and in the carboxy termini of VP1 which, according to X-ray analysis of the crystallized strains, are located at the surface of the virion. In general, the loop sequences were more conserved between HRV-B than between HRV-A strains (Fig. 1). The long CD loop was relatively conserved within both species. Alignment of all HRV serotypes revealed species-specific amino acid insertions, one in each species: a 2 aa insertion in the DE loop of the HRV-B serotypes and a 1 aa insertion in the beginning of the βG strand of the HRV-A serotypes. Sixteen HRV-A serotypes had 1 or 2 aa-long deletions in the GH loop.

<table>
<thead>
<tr>
<th>HRV1A</th>
<th>HRV16</th>
<th>HRV2</th>
<th>HRV3</th>
<th>HRV1AAS</th>
<th>HRV16SS</th>
<th>HRV2SS</th>
<th>HRV3SS</th>
<th>HRV1ASS</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPVWYTVNVEVTVLPFPFKEYTETGQYTVTQITRFPLFLGCVTVIRTRKT---VTYNMGTQIN</td>
<td>HEPVWYTVNVEVTVLPFPFKEYTETGQYTVTQITRFPLFLGCVTVIRTRKT---VTYNMGTQIN</td>
<td>HEPVWYTVNVEVTVLPFPFKEYTETGQYTVTQITRFPLFLGCVTVIRTRKT---VTYNMGTQIN</td>
<td>GLEDSIVYTVNVEVTVLPFPFKEYTETGQYTVTQITRFPLFLGCVTVIRTRKT---VTYNMGTQIN</td>
<td>HEPVWYTVNVEVTVLPFPFKEYTETGQYTVTQITRFPLFLGCVTVIRTRKT---VTYNMGTQIN</td>
<td>HEPVWYTVNVEVTVLPFPFKEYTETGQYTVTQITRFPLFLGCVTVIRTRKT---VTYNMGTQIN</td>
<td>HEPVWYTVNVEVTVLPFPFKEYTETGQYTVTQITRFPLFLGCVTVIRTRKT---VTYNMGTQIN</td>
<td>HEPVWYTVNVEVTVLPFPFKEYTETGQYTVTQITRFPLFLGCVTVIRTRKT---VTYNMGTQIN</td>
<td>HEPVWYTVNVEVTVLPFPFKEYTETGQYTVTQITRFPLFLGCVTVIRTRKT---VTYNMGTQIN</td>
</tr>
</tbody>
</table>

**Fig. 1.** Locations of designated conserved motifs in deduced amino acid sequences (five upper lines) and secondary-structure elements (five lower lines) in human rhinovirus capsid protein VP1. Sequences and secondary-structure elements shown here were obtained from the Protein Database and they represent the five HRV serotypes for which X-ray crystallographic analysis of virion structure has been published. Alignment of VP1 sequences of all 101 HRV prototype strains was used to identify conserved motifs (for definition, see text); motifs specific for the species HRV-A are shaded grey and those for the species HRV-B are shaded black. Motifs present in both HRV species are denoted with asterisks above the alignment. Dots represent individual amino acid columns that are either fully conserved or similar within all HRV serotypes. Secondary structure for each residue is calculated according to Kabsch & Sander (1983) and the assignments in the secondary-structure alignment are: H, helix; B, residue in isolated β-bridge; E, extended β-strand; G, 310 helix; T, hydrogen-bonded turn; S, bend; -, no secondary structure; X, a gap generated by CLUSTAL_X. The marked borders of the loops and the β-strands are selected based on the determined secondary structures. Grey colouring in secondary-structure alignment represents similar structure in >80% of studied strains.
The length of the pre βB amino-terminal part of the VP1 protein was the same (74 aa) in all HRV serotypes. Six of the ten minor-group prototype strains and two proposed major-group serotypes, HRV23 and HRV25, had the shortest VP1 sequences (281–283 aa). Compared with the major receptor-group serotypes, the relative deletions in the minor-group sequences were accommodated in the BC, DE and GH loops.

Conserved motifs

We have reported before that, in accordance with the relatively greater coherence of the nucleotide sequences of the VP1 gene within the species HRV-B, serotypes belonging to the HRV-B species shared >68% identity of VP1 amino acid sequences, compared with 58% identity within the HRV-A species (Laine et al., 2005). Alignments of deduced amino acids generated for all HRV prototype strains as described above were examined for conserved protein motifs with the criterion of three or more consecutive fully conserved or similar amino acid residues (Thompson et al., 1997). Accordingly, 24 conserved motifs were observed within the HRV-B species (black shading in amino acid sequences of HRV3 and HRV14 in Fig. 1), compared with HRV-A (grey shading in Fig. 1) with 18 conserved motifs. The motifs conserved within species A or B were often in the same locations in the alignment, but those of HRV-B were typically longer.

Altogether, 10 conserved amino acid motifs were shared by both HRV species. Half of them were located in or near the β-strands (Table 1). According to a crystal-structure model, the first two amino-terminal motifs are likely to lie under the amino terminus of VP3 (Rossmann et al., 1985) in potential contact with RNA, whereas the third is located near the βB strand. The rest of the protein motifs (iv–x) were plotted on pentamer subunits derived from the X-ray crystallographic structures of HRV14 (Arnold & Rossmann, 1988) and HRV16 (Hadfield et al., 1997) by using the RasMol (version 2.5) program. According to the results, motifs (v), (vi) and (x) (Table 1) are likely to be exposed at the virion surface. The fully conserved FYDGY motif was seen on the surface of serotype 14, but not on HRV16. Within-species conserved protein motifs were distributed throughout the entire VP1 protein, except for the BC loop and the carboxy-terminal region. Strikingly, β-strands B, C and F of HRV-A did not contain conserved motifs, although corresponding motifs were identified in HRV-B.

Table 1. VP1 amino acid motifs conserved within the genus Rhinovirus

<table>
<thead>
<tr>
<th>Motif</th>
<th>Location (amino acid position in Fig. 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>ETG</td>
</tr>
<tr>
<td>(ii)</td>
<td>[ILMV][ETR]</td>
</tr>
<tr>
<td>(iii)</td>
<td>FLGR[AS]</td>
</tr>
<tr>
<td>(iv)</td>
<td>RF[D]NJE</td>
</tr>
<tr>
<td>(v)</td>
<td>[IV]PPG</td>
</tr>
<tr>
<td>(vi)</td>
<td>W[EQ][ST]</td>
</tr>
<tr>
<td>(vii)</td>
<td>[KR][F][ST][ILV][P][FY]</td>
</tr>
<tr>
<td>(viii)</td>
<td>[ILV][AT][S]AY</td>
</tr>
<tr>
<td>(ix)</td>
<td>FYDGY</td>
</tr>
<tr>
<td>(x)</td>
<td>R[ILMV][IV]</td>
</tr>
</tbody>
</table>

Amino acids inside brackets equate a single column with similar amino acids listed.

Fig. 1 also shows a correlation of the conserved amino acid motifs with the known secondary structures revealed by the published X-ray analysis of crystallized prototype strains. For the latter, the corresponding amino acid sequences and the published secondary structures of the five strains were first aligned separately, followed by the profile–profile alignment in CLUSTAL_X. The generated alignment was then edited manually. Whilst the amino acid sequences could be divided clearly into two groups according to the HRV species, the secondary-structure elements did not reveal any obvious species-specific pattern. Structurally, these five strains representing the two HRV species were very similar, yet minor strain-specific characteristics could be observed in the amino and carboxy termini, as well as in the BC, DE, EF and GH loops, mainly due to the presence or absence of α-helices.

Amino acid residues differentiating ICAM-1- and LDL-R-using HRV serotypes

During HRV binding to ICAM-1, the D1 domain of ICAM-1 penetrates into the receptor canyon of HRV and interacts with parts of VP1 and VP3, whilst the CD loop of ICAM-1 lies against the ‘south’ rim of the canyon, including VP2 of HRV16 (Bella & Rossmann, 1999). It has been demonstrated previously that ICAM-1 recognizes different amino acid residues on the canyon floors of HRV14 and HRV16, representing the two HRV species (Kolatkar et al., 1999). These data were used as a basis for selecting species-specific sets of potentially ICAM-1-interacting amino acid residues from the alignment containing deduced sequences of the carboxy terminus of VP3 and the entire VP1 (HRV16/HRV-A: Thr3179, Pro3180, Asp3181, Thr3182, Ser3185, Gly3186, Ile3187, Arg1205, Val1209 and Asp1213; HRV14/HRV-B: Asp3177, Pro3178, Asp3179, Thr3180, Gly3181, Ile3182, Asp3183, Pro3184, Asp3185, Arg1205, Asn1206, Asp1209 and Gly1212). The resulting non-linear amino acid strings were then aligned in a species-specific way and found to be highly conserved within the HRV-B species, whilst considerable variation was observed among the HRV-A serotypes (Fig. 2). The major-group serotypes belonging to the HRV-A species shared four fully conserved amino acids (Thr, Pro, Asp, Gly), whereas only the glycine was also shared with the minor receptor-group strains. All ten HRV-A strains using LDL-R, as well as serotypes HRV23 and HRV25, showed a non-major group-like pattern, especially concerning the VP3 residues. In a ‘dendrogram’
generated from the selected non-linear amino acid strings of all HRV-A serotypes, the minor-group strains and serotypes 23 and 25 formed three clusters that were clearly separate from the major-group serotypes (Fig. 3).

Colonna et al. (1988) previously reported that, based on site-directed mutagenesis studies on HRV14, the VP1 amino acid residues Lys103, Pro155, His220 and Ser223 may have an impact in the virus–receptor interaction. All of these amino acid residues also lie within the footprint of truncated ICAM-1 molecule. In the alignment of all HRV VP1 sequences, Pro155 was fully conserved, whereas the position corresponding to Lys103 was fairly variable (Lys, Arg, His, Asp, Asn, Gln, Ala, Gly or Pro). The position of Ser223 was again more uniform, presenting with aliphatic amino acids [mostly serine or threonine (95/101)] and the remaining six strains had an alanine (HRV61, HRV25, HRV62, HRV31 and HRV47) or a valine (HRV84) residue in the corresponding position (not shown). As regards amino acids at the positions of Lys103, Pro155 and Ser223, no clear distinction between the two receptor groups could be observed in the amino acid sequence alignment. Variation of amino acids at the fourth position, His220 (corresponding to the Asp213 in HRV16), showed different trends in the two receptor groups (His, Ala and Asp in the major group; Asp, Asn and Arg in the minor group). However, HRV1A and HRV1B of the minor group shared aspartic acid with 49 major-group serotypes at this position. Again, HRV23 and HRV25 clustered among the majority of the minor-group serotypes.

### Antiviral-sensitivity groups

After sorting all prototype strains (except HRV87) into the two established HRV species on the basis of VP4/VP2-coding sequences, we previously reported a fair co-segregation of the two designated antiviral groups with the phylogenetic classification (Savolainen et al., 2002). Antiviral group A contained all HRV-B strains (red dots in Fig. 4) and most HRV-A strains belonged to antiviral group B (black dots in Fig. 4a). However, the co-segregation was not complete, as seven HRV-A serotypes (HRV8, HRV13, HRV32, HRV43, HRV45, HRV54 and HRV95) appeared to belong to antiviral group A or were among the few strains with an ambiguous location in the published two-dimensional projection of the coordinates. We now had access to three-dimensional coordinates of the sensitivity database and used the Scatter3D program to rotate the coordinates, enabling visualization from different directions. Fig. 4 shows the previously published two-dimensional projection of the coordinates (Fig. 4a) and a new one (Fig. 4b), presenting a better separation of HRV-A and HRV-B strains. It was generated by an approximately 145° rotation of the three-dimensional coordinates around the y axis as compared to...
with the projection seen in Fig. 4(a). Of the ‘wrongly located’ strains, HRV serotypes 13, 32, 43, 45 and 54 are now in antiviral group B, together with all but two of the HRV-A strains. HRV8 and HRV95 still appear to remain among the HRV-B strains in antiviral group A and HRV27, an HRV-B serotype, is now translocated to antiviral group B (Fig. 4b).

We made an attempt to clarify the remaining discrepancy by creating an alignment of the non-linear set of amino acids corresponding to the 25 residues reported to line the drug-binding hydrophobic pocket within the β-barrel of VP1 in HRV1A and HRV14, respectively (Kim et al., 1993). However, this attempt did not give a simple answer to the ‘aberrant’ localization of the above three serotypes. We found that the two species showed definite differences in the patterns of amino acids, as seen in the alignment published by Ledford et al. (2004), with HRV8 and HRV95 showing typical HRV-A-like patterns and HRV27 a typical HRV-B pattern. Interestingly, Ile215 was fully conserved in HRV-B, whilst the corresponding amino acid (position 208) was variable in HRV-A (usually Thr). In HRV8 and HRV95, it was Ile, as among the HRV-B strains, whereas the genetically closely related HRV45 had Val at this position, like several other HRV-A strains.

**DISCUSSION**

Our aim in this study was to use multiple alignment of deduced amino acid sequences of capsid protein VP1 of all HRV prototype strains to examine conservation of selected motifs, suggested previously to have certain functional roles in HRV biology, and to identify new conserved motifs for future functional studies. Most conserved motifs, as
designated by a linear stretch of three or more conserved amino acids, were located, not surprisingly, in the structurally highly conserved \( \beta \)-barrel core of the folded protein, but some were also identified in the loops between the \( \beta \)-strands and in the amino-terminal region. Individual VP1 amino acids suggested previously to be crucial for HRV–ICAM-1 interactions, based on mutagenesis experiments and structural studies in model strains, were not conserved in either HRV species. On the other hand, an alignment of a predicted non-linear ICAM-1 footprint sequence of VP3/VP1 amino acids clearly outclassified the minor receptor-group HRV-A strains, together with HRV23 and HRV25, from all of the ICAM-1–using major receptor-group strains. An improved processing of previously generated three-dimensional coordinates of a multivariate analysis of antiviral sensitivity of all HRV serotypes (Andries et al., 1989) revealed almost-complete co-segregation of HRV species and the designated antiviral class.

The overall alignment of the deduced amino acid sequences of all HRV prototype strains revealed, as expected, conserved motifs in the \( \beta \)-barrel coinciding with the documented secondary-structure motifs and extensive sequence variation in most of the intervening loops. The BC loop, harbouring the major neutralization antigen in studied HRV serotypes (Rossmann et al., 1985; Sherry & Rueckert, 1985; Sherry et al., 1986), was also responsible for most length variation of the entire VP1 protein. The carboxy-terminal end of the protein likewise showed considerable variation in both sequence and length. This is understandable, as this part of VP1 is close to the surface of the virion, allowing more relative freedom for the resulting conformational variation than the capsid wall. In contrast, the amino-terminal part was relatively conserved in both species and did not vary in length at all before the BC loop. This is in contrast to the closely related enteroviruses, which appear to show both ‘species-specific’ and ‘serotype-dependent’ length variation (Palmenberg, 1989; Hovi et al., 2000) in the corresponding region of VP1. According to the GenBank-submitted sequences, the ‘pre-BC loop’ length of VP1 of different enterovirus strains ranges from 64 to 88 aa (data not shown).

We defined an arbitrary 3 aa linear stretch as a designated conserved motif, knowing well that functional motifs are not necessarily linear in sequence. Indeed, as regards, for instance, the identified neutralization antigenic sites, the opposite is a more frequent observation (Rossmann et al.,...
Species-specific conserved motifs were more frequent among the HRV-B serotypes than among the HRV-A serotypes and they were often longer than the corresponding motifs of HRV-A serotypes. These observations may be partly due to the smaller number of serotypes in the HRV-B species. However, the opposite was true in the amino-terminal region, where HRV-A showed relatively more conservation. One region, conserved here for both HRV species, is the motif 1032-plLDAETGHT/paLTAnETGAT (Fig. 1), which is also highly conserved among enteroviruses (Hovi & Roivainen, 1993). This motif appears to be translocated reversibly from the inner capsid surface to the virion surface in free poliovirus particles under physiological conditions (Roivainen et al., 1993), reminiscent of the stable translocation of the amino terminus of VP1 following receptor interaction (Fricks & Hogle, 1990). For coxsackievirus A9, Airaksinen et al. (2001) found in their saturation mutagenesis studies that this motif appears to have a role in regulating the stability/flexibility of the capsid wall. In the X-ray crystallographic studies of virion structures of all studied enterovirus and HRV strains, conserved residues in this motif interacted with conserved residues in other capsid proteins. Whilst it is not known whether this specific motif or the amino-terminal region of VP1 in general has a similar functional role in HRV biology, this kind of strategy might be useful in exploring potential functional roles of the other conserved motifs.

The VP1 protein is involved centrally in the virion–receptor interactions of both HRV receptor groups. In the minor-receptor group, LDL-R or a related receptor molecule interacts with the VP1 protein only (Hewat et al., 2000), whilst the ICAM-1 footprint in the receptor canyon of the major-group viruses comprises residues of VP1, VP2 and VP3 (Bella & Rossmann, 1999; Kolatkar et al., 1999). In the case of the minor-receptor group, the amino acid sequence in the receptor footprint is not literally decisive in the specificity, as the only conserved amino acid residue, a lysine in the H1 loop, is shared by not only the two ambiguously grouped serotypes HRV23 and HRV25 (Crump et al., 1993), but also by eight other major-group serotypes (not shown). The pattern of the surface charge in the footprint, in addition to the lysine, was considered crucial for the receptor specificity of the minor-receptor group (Vlasak et al., 2003). In order to study potential conservation and specificity of the ICAM-1 footprint, we made alignments separately in the two HRV species for amino acid residues in the carboxy terminus of VP3 and various parts of VP1, selected according to the previously reported corresponding residues in HRV16 and HRV14, respectively (Bella et al., 1998; Colonno et al., 1988). In spite of the fact that the sequences available did not cover all involved residues, the analysed predicted ICAM-1 footprint residues differed strikingly between the major- and the minor-group serotypes. Within the species, again, HRV-B serotypes showed a great conservation, whilst amino acids of the HRV-A serotypes were more variable. The ICAM-1-using serotypes belonging to the HRV-A species were, on average, very similar and the remaining 12 strains, including all of the ones known to use LDL-R, as well as HRV23 and HRV25, could be distinguished into three different groups according to the amino acid ‘sequence’ of the presumed footprint. These 12 strains also form three noticeable clusters with high bootstrapping values in phylogenetic analysis based on the entire VP1 sequences (Ledford et al., 2004; Laine et al., 2005), and there is recent experimental evidence that demonstrates unequivocally that HRV23 and HRV25 indeed use VLD-R and not ICAM-1 as the receptor in HeLa cells (Vlasak et al., 2005). In contrast to the traditional classification, HRV23 and HRV25 are thus true minor receptor-group serotypes.

Certain individual amino acid residues in the ICAM-1 footprint were reported to be involved directly in the receptor interactions of HRV14 and HRV16, respectively (Kolatkar et al., 1999). In our alignments, none of them was fully conserved or specific for the major-receptor group. It is possible that, again, the overall pattern of residues in the footprint may be more important for the interaction than the individual amino acids themselves. As regards the minor-group serotypes in this context, the observed divergence of the footprint residues from those in the major group probably has nothing to do with the LDL-R interaction, but these aberrant residues might interfere with the binding to ICAM-1 or with subsequent steps in the ICAM-1–virus interaction. We made no attempts to prove this hypothesis.

According to an earlier analysis, the co-segregation of HRV species and the two previously determined antiviral-sensitivity classes (Andries et al., 1989) was fairly good, but left several HRV-A serotypes in the ‘wrong’ antiviral group (Savolainen et al., 2002). However, rotating the three-dimensional coordinates of the antiviral data now revealed a new view on the distribution, which solved most of the previous co-segregation problems, but still left space for additional explanations concerning three serotypes. Ledford et al. (2004) reported that the non-linear set of VP1 amino acid residues aligning the ‘antiviral’ hydrophobic pocket varies according to the relative sensitivity to pleconaril. In both HRV species, the greater the difference in these amino acids between any two strains, the larger the difference in effective concentrations of pleconaril against the two strains. Whilst a definite HRV-A-like amino acid pattern was seen for HRV8 and HRV95 in this alignment, they shared one amino acid with all HRV-B strains and differed at this position from the genetically closely related HRV45 (Laine et al., 2005), as well as from all other HRV-A serotypes. In contrast, HRV27 – an HRV-B serotype translocated from the ‘correct’ antiviral group A to the ‘wrong’ group B by the described rotation – had an amino acid pattern typical of HRV-B serotypes. A ‘phylogenetic’ analysis of the assumed pocket-aligning non-linear sets of amino acids in different prototype strains, carried out as described above for the ICAM-1 footprints, did not change these conclusions (data not shown). Data from several laboratories have suggested...
that some stocks of HRV prototype strains are contaminated with poliovirus type 1 (Davies et al., 2003; Savolainen & Hovi, 2003; Blomqvist, 2004). However, this is not a likely explanation for the ‘aberrant’ location of HRV27, because poliovirus type 1 is no longer in the same antiviral group. In hindsight, it is surprising that a relatively straightforward phenotypic feature, like antiviral sensitivity, was able to predict the genetic clustering of all real HRV serotypes almost perfectly.

In conclusion, by aligning deduced amino acid sequences of the capsid protein VP1 of all HRV prototype strains, we have identified a number of both species- and genus-specific conserved motifs throughout the primary structure of the protein. Further studies are needed to assess whether any of them has a specific function in HRV biology. Individual ICAM-1 footprint amino acids, suggested previously to play a key role in virion–receptor interactions, were not conserved according to the alignment. On the other hand, amino acid residues corresponding to the documented ICAM-1 footprint were shown to classify the two receptor groups unequivocally. Amino acid residues aligning in the hydrophobic pocket in VP1 showed a varying range of divergence and species-specific similarity. This was in agreement with the observation that the antiviral-sensitivity grouping of HRV strains based on a previous multivariate analysis showed almost-complete agreement with the species classification.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Academy of Finland.

REFERENCES


Oberste, M. S., Mahier, K., Schnurr, D. & 9 other authors (2004). Enterovirus 68 is associated with respiratory illness and shares biological features with both the enteroviruses and the rhinoviruses. J Gen Virol 85, 2577–2584.


