CHARACTERISATION OF THE FIBER GENE
AND PARTIAL SEQUENCE OF THE EARLY REGION 4
OF BOVINE ADENOVIRUS 2
(SHORT COMMUNICATION)

R. IZADPANAH[1,4], Mária BENKŐ[2], Krisztina URSU[2,3], Á. DÁN[2,3], M. RUSVAI[1*]
and B. HARRACH[2]

[1]Department of Microbiology and Infectious Diseases, Faculty of Veterinary Science,
Szent István University, H-1581 Budapest, P.O. Box 22, Hungary; [2]Veterinary Medical
Research Institute, Hungarian Academy of Sciences, H-1581 Budapest, P.O. Box 18,
Hungary; [3]Central Veterinary Institute, H-1581 Budapest, P.O. Box 22, Hungary; [4]RAZI
Vaccine and Serum Research Institute, 21987 Tehran, P.O. Box 11365-1558, Iran

The full sequence of the fiber gene and partial sequence of the putative 17 kD protein gene of bovine adenovirus-2 (BAdV-2) were determined. The size of the fiber gene of BAdV-2 proved to be 561 amino acids, of which the amino acids 37 to 385 form a typical shaft domain of 22 repetitive motifs. On the complementary strand, a gene homologous to the 17 kD protein coded in the E4 region of several human adenoviruses was found. The sequence analysis seems to confirm the presence of an intron in the sequenced part of the E4 region.

Bovine adenovirus type 2 (BAdV-2) belongs to the genus *Mastadenovirus* (comprising adenoviruses of mammals) of the family *Adenoviridae*. It is an interesting serotype because of its ability to bypass species specificity and to cause natural infection both in cattle and sheep (Belák et al., 1983). Like other adenoviruses, BAdV-2 has a linear double-stranded DNA genome, which has been physically mapped with different restriction enzymes (Belák et al., 1986; Salmon and Haj-Ahmad, 1993). The size of the genome is approximately 32.5 kb (Benkő et al., 1988; Salmon and Haj-Ahmad, 1993). Previous investigations revealed different haemagglutination activity of BAdV-2 strains recovered from cattle or sheep (Belák and Pálfi, 1974), and BAdV-2 isolates are classified accordingly into subtypes A and B. Heteroduplexes formed between the denatured genomes of subtype A and B strains showed loops in the early region 3 (E3) and fiber genes implicating differences in these regions (Belák et al., 1983).

---

*Corresponding author; E-mail: rusvai@novell.vmri.hu; Fax: +36 (1) 251-9260

```
ATGAAACGCAGCGTGCCAGAAGACTTTGATCCTGTTTACCCTTACGGAAAAAGACCACTA 60
 M   K   R   S   V   P   E   D   F   D   P   V   Y   P   Y   G   K   R   P   L
                                                    TAIL↤↦SHAFT
AACATTCTGCCACCGTTTTATAGCTCAAATGGTTTTGTGGAGGCACCAACTGAAACTCTT 120
 N   I   L   P   P   F   Y   S   S   N   G   F   V   E   A   P   T   E   T   L
TCTTTAAAATTGGCCAATCCTGTTGATTTTATGCAAAATGGGGCTATTGGGTTAAAACTT 180
 S   L   K   L   A   N   P   V   D   F   M   Q   N   G   A   I   G   L   K   L
GGAGGTGGACTTTCAATTAACCAAGACGGCGAACTTGAATCTCAAACAATTACGTCTACA 240
 G   G   G   L   S   I   N   Q   D   G   E   L   E   S   Q   T   I   T   S   T
GTAAATCCACCTTTATATCAACAAAATGGCGGTTTAAATTTAAAATATGGAGAAGAATTT 300
 V   N   P   P   L   Y   Q   Q   N   G   G   L   N   L   K   Y   G   E   E   F
GATATTGAAAATGAAGCTCTTAAAATAAAAACCATCGCTCCAATTACTAAGACAGAAAAT 360
 D   I   E   N   E   A   L   K   I   K   T   I   A   P   I   T   K   T   E   N
GGATTAACTTTGTCTATAGGAGATGGATTAGAACTAAACTCAAATAATACACTCCAAGCT 420
 G   L   T   L   S   I   G   D   G   L   E   L   N   S   N   N   T   L   Q   A
AGATTATCAAGTGGATTAGAAATTGACAACCAAGCAATTAGACTCAGAGTTCATGAACCT 480
 R   L   S   S   G   L   E   I   D   N   Q   A   I   R   L   R   V   H   E   P
TTAAACTTAAATGCTTCTACTGGTGCTCTTCAATGCAGAATAGGAAATGGATTAACAGTA 540
 L   N   L   N   A   S   T   G   A   L   Q   C   R   I   G   N   G   L   T   V
TCTGATAATAGCTTAGTTGTCTATCCTCATGAACCTTTAAATTTAGACCAAACTTCTGGC 600
 S   D   N   S   L   V   V   Y   P   H   E   P   L   N   L   D   Q   T   S   G
AAATTACAACTAAGAGTGGGTAATGGAATTAATGTGCAAAATAGCTCGCTTGTTGCAAGA 660
 K   L   Q   L   R   V   G   N   G   I   N   V   Q   N   S   S   L   V   A   R
ATAGGGCAGGGATTGGCTTTTAACAACTCTGACATTCAAATTAATGCCGCACCTCCATTT 720
 I   G   Q   G   L   A   F   N   N   S   D   I   Q   I   N   A   A   P   P   F
ACTTTTTCTAATAATCAGCTTTCTATATCACTAGGGGATGGATTAGTAACTAATGCCAGT 780
 T   F   S   N   N   Q   L   S   I   S   L   G   D   G   L   V   T   N   A   S
CAATTAAAAGTTAATTTTGGTAAAGGACTTTTTATTAATTCTTCTGATTCATCAAAACTT 840
 Q   L   K   V   N   F   G   K   G   L   F   I   N   S   S   D   S   S   K   L
CAAGTAAACATTAGACCTCCTTTAAATTATTTGGAAATAGCAATTCTCTTACTGTAGTTA 900
 Q   V   N   I   R   P   P   L   N   Y   F   G   N   S   N   S   L   T   V   V
TCTGGTAATGGCTTAGGAGTATCTGGTACTAATCTAGGCTCTAACCTTTATGTTAAAACT 960
 T   G   N   G   L   G   V   S   G   T   N   L   G   S   N   L   Y   V   K   T
GGCAATGGATTAGAAGCTGATAGCAGTAACGTTAGAGTAAAAATTGCAAATGGGTTACAG 1020
 G   N   G   L   E   A   D   S   S   N   V   R   V   K   I   A   N   G   L   Q
TTTACTGATGGCAACATTGAAGCTAATTTAGGAAATGGGTTAACATTTTCAAACGGTCAA 1080
 F   T   D   G   N   I   E   A   N   L   G   N   G   L   T   F   S   N   G   Q
ATTACTGCAAACATTGGCGCCGGTCTTGCGTTTTTAAATGGTCAAATTACACTAGTGAAC 1140
 I   T   A   N   I   G   A   G   L   A   F   L   N   G   Q   I   T   L   V   N
                    ↦KNOB
AGCACTCCTTCTGGTTATACAGATTATACTTTGTGGACTACTCCAGACCCTTCTCCTAAT 1200
 S   T   P   S   G   Y   T   D   Y   T   L   W   T   T   P   D   P   S   P   N
GCTAGCATTAAAACTGACTTAGATGCTAAGCTTGTTTTAACTTTATCAAAAGCAGGCAGC 1260
 A   S   I   K   T   D   L   D   A   K   L   V   L   T   L   S   K   A   G   S
ACTGTAATAGGCACAATAGGTATTTTTGCTCTAAAATCTCCTCTTACTCCTATTTCAGAA 1320
 T   V   I   G   T   I   G   I   F   A   L   K   S   P   L   T   P   I   S   E
AATTCTATTAATGTTGAAATTTTTTTTGATGCTAATGGAGAAATTAATTTAACTACTAGT 1380
 N   S   I   N   V   E   I   F   F   D   A   N   G   E   I   N   L   T   T   S
TCGCTAAAAAGCTATTGGGGGATTTAGAGAAGGTGATTCTTATAATCCATCTTCTAACTTA 1440
 S   L   K   S   Y   W   G   F   R   E   G   D   S   Y   N   P   S   S   N   L
AACCCCCTTTACTTGATGCCTAATACTTATGCATACCCTCAAGGTCGGAAAACTATTACA 1500
 N   P   L   Y   L   M   P   N   T   Y   A   Y   P   Q   G   R   K   T   I   T
```

```
CAAGTTTTTCCACTAGAAGTATACTTAAATGGAGACACTGCTAAGCCTGTGCCCTTAGAA 1560
 Q   V   F   P   L   E   V   Y   L   N   G   D   T   A   K   P   V   P   L   E
GTTGCTTTTAATACTTTGTCTTCTACTGGTTTTTCTCTTGAATTTACTTGGAGAAATCTT 1620
 V   A   F   N   T   L   S   S   T   G   F   S   L   E   F   T   W   R   N   L
AATGCTTACACTGGAGAAGCGTTTGCGGTGTCTTTAGGAAATTTTACATACATTAGTCAA 1680
 N   A   Y   T   G   E   A   F   A   V   S   L   G   N   F   T   Y   I   S   Q
TATTAAATTTTAAAACTTTTTATTGCTGATTTTGGTAATACACGACATGTTAACATTCCA 1740
 Y   -
     -   I   K   F   S   K   I   A   S   K   P   L   V   R   C   T   L   M   G   G
END OF E4
CCACCTTCCCATTTTACTTTGTAAACAAATCTATTAAAATGCAATCCAGTGTTGTAATTA 1800
 G   E   W   K   V   K   Y   V   F   R   N   F   H   L   G   T   N   Y   N   D
TCTTGATTGTGTCTTAATTTTACAGCATGTTCTACACATTCAAAATCAGGGCTAGTAATC 1860
 Q   N   H   R   L   K   V   A   H   E   V   C   E   F   D   P   S   T   I   L
AACACAAATCCGGTAAAGCTAGAGCTTAAATCAGGCACACAGTTTAATAATTGGGGTCGT 1920
 V   F   G   T   F   S   S   S   L   D   P   V   C   N   L   L   Q   P   R   L

                                         ↓
AGTAAATGTCGTGCTTCGGACGCTTGTCCGCTTCCTCGTTCTGCGTCAAAAAGCACTGTG 1980
 L   H   R   A   E   S   A   Q   G   S   G   R   E   A   D   F   L   V   T   S
GACATCCGCAGCTGTGTTTCACCATCTCCACCTTGCAAAGCGGGGTGCCCTTCGGGATCA 2040
   M   R   L   Q   T   E   G   D   G   G   Q   L   A   P   H   G   E   P   D   R
CGAGCTCCGCGACGTCTAGCGGTGAGTTCAGCTTCAGCTCTAGCACAGGCTTCGTAAAAG 2100
   A   G   R   R   R   A   T   L   E   A   E   A   R   C   A   E   Y   F   R
CGTCCCATACCTTGCACCCAGAAGATTTTAAATCACCATCTAACCCGCTTAGCTCAGAAA 2160
   G   M
CACTACATTCAGCATGCCAGCCAAAAGGTAAATCAAATTCTACATCAAAGAAAGCGGTAG 2220
CAGAGCCAACTTTTCTTACTACTATATCTTTCTCTGTGCGTAAATGGGCAATATCTTCAT 2280
TATTTTTATCTCTAATAGGAGCAAACTTTGTGCTTTGATACATGCACCTTACACGTTTGT 2340
TCCTATTTTGCCAAGCTAGTCTTTCAGCTTCCGTCATTTTTTCCCATCTACGTTCAAAGC 2400
GTTTCTGATTTCTATCCTCACATTTCTTATGTCCACAATCATCAAACTCACAAGAAAGT 2460
ATCTATACATTAAAATTGGAGTTCCAGCTGGAATTC                          2496
```

Fig. 1. Nucleotide sequence of the BAdV-2 fiber gene and the end of the E4 region. The deduced amino acid sequence of the fiber gene and that of the carboxy-terminus of the putative 17 kD protein encoded on the reverse strand are indicated by one-letter code printed in bold. The different structural parts of the fiber are marked. The probable splice acceptor site in the E4 region is shown by an arrow. The alternative putative amino-terminus of the 17 kD protein is shown in italics. Underlining shows the putative polyadenylation signals for the fiber and the E4 transcripts, respectively

A large portion of the DNA of the subtype A prototype strain (No. 19) of BAdV-2 has been sequenced including E1 and protein IX (Salmon and Haj-Ahmad, 1994), E3 (Esford and Haj-Ahmad, 1994), E4 (Fitzgerald et al., 1997), the polymerase (Yagubi et al., 1998), the core proteins (Rusvai et al., 2000), and the pVI, hexon and protease genes (unpublished data). The purpose of the present study was to sequence the gap between map units 84.4 to 90.5 which region contains the knob of the fiber gene and the end of the E4 region.

The fiber, extruding from each of the 12 vertices of the adenovirus capsid serves as the ligand between the capsid and the host cell. The fibers are homo-trimeric proteins containing an amino-terminal penton base attachment domain,

the 'tail', a long, thin central 'shaft', and a distal head domain, the carboxy-terminal cell attachment 'knob'. The entry of adenovirus into susceptible cells requires distinct sequential steps. The first step is a high affinity binding of adenovirus to the α2 domain of the major histocompatibility complex class I protein via the carboxy-terminal knob domain of the fiber protein (Douglas et al., 1999). The fiber knob carries the type-specific γ-antigen, which determines, together with the ε-antigen of the hexon, the serotype specificity of an adenovirus (Fender et al., 1995; Eiz and Pring-Akerblom, 1997).

In mastadenoviruses, the E4 region is located on the right end of the genome and is transcribed from right to left. In this region, several splicing sites exist, which are temporally regulated during the course of adenovirus infection (Dix and Leppard, 1993). It is shown that the E4 gene products are required for the proper splicing of the tripartite leader (Öhman et al., 1993).

The plasmid pBS112 (Belák et al., 1986) contains the right part of the BAdV-2 subtype A prototype strain (No. 19) genome, a *Sal*I-*Eco*RI fragment covering map units 63–90.5 and contains the fiber gene together with the E3 region and the end of the E4 region. After transferring this fragment into vector pMOB, a transposon insertion method was used for the sequencing (TN1000, Gold Biotechnology, Inc., St. Louis, USA). The pMOB vector carrying the viral fragment was introduced into *E. coli* strain DPWC containing the TN1000 transposon on an F factor. The continuous relocation of the transposon is 'frozen' by conjugation between DPWC and BW26 *E. coli* strains. Two primers (G186 and G187) complementary to the two ends of TN1000 allow the sequencing of the cloned DNA from the direction of the transposon. DNA sequencing was performed on an ABI 373 automated sequencer at the Biological Research Centre of the Hungarian Academy of Sciences (Szeged, Hungary). The overlapping sequences were assembled. The sequence analysis was performed using the PC/Gene (IntelliGenetics Inc.) and LASERGENE program packages (DNASTAR Inc., Madison, Wisconsin). The homology search was performed using the BLASTN and BLASTX programs on the non-redundant data bases of the National Center for Biotechnology Information (NCBI, USA).

The fiber gene of BAdV-2 consists of 1686 nucleotides (Fig. 1). Two-third of this gene was already sequenced (Esford and Haj-Ahmad, 1994). By re-sequencing these parts, several corrections are suggested regarding the conservative part of the fiber similar in all mastadenoviruses. The differences in the nucleic acid sequence resulted in a slightly different amino acid sequence, more consistent with the consensus sequence of the fiber tail region of other adenovirus types. The predicted fiber of BAdV-2 is a 561 amino-acid-long polypeptide. The size of the fiber protein varies extensively among adenovirus types. For example the fiber of BAdV-3 is composed of 976 amino acids (Mittal et al., 1993) which means 40% difference in size compared to BAdV-2. The amino-terminal tail of this predicted polypeptide contains 37 amino acids (Fig. 2).

The high similarity of the tail of the different serotypes reflects the conserved structure essential for the interaction between the tail and the penton base. In all HAdVs, a hydrophobic sequence motif [FNPVYPY(D/E)] is conserved (Caillet-Bouding, 1989). A similar sequence element (FDPVYPYG) was found also in the fiber of BAdV-2 (Fig. 2).

```
A  MKRSVPEDFDPVYPYGKRPLNILPPFYSSNGF
   VEAPT..........................
B  1   38    ETLSLKLA.NPVDFMQN.....
   2   54    GAIGLKLG.GGLSINQDGELES
   3   75    QTITSTVN.PPLYQQN......
   4   90    GGLNLK...YGEEFDIEN....
   5   105   EALKIKTI.APITKTE......
   6   120   NGLTLSIG.DGLELNSN.....
   7   136   NTLQARLS.SGLEIDNQA....
   8   153   IRLRVH...EPLNLNAST....
   9   168   GALQCRIG.NGLTVSD......
   10  183   NSLVVYPH.EPLNLDQTS....
   11  200   GKLQLRVG.NGINVQN......
   12  215   SSLVARIG.QGLAFNN......
   13  230   SDIQINAA.PPFTFSN......
   14  245   NQLSISLG.DGLVTNA......
   15  260   SQLKVNFG.KGLFINSSDS...
   16  278   SKLQVNIR.PPLNYFGNS....
   17  295   NSLTVVTG.NGLGVSGTNLG..
   18  314   SNLYVKTG.NGLEADS......
   19  329   SNVRVKIA.NGLQFTD......
   20  344   GNIEANLG.NGLTFSN......
   21  359   GQITANIG.AGLAFLNGQ....
   22  375   ITLVNSTP.SG...........
C  YTDYTLWTTPDPSPNASIKTDLDAKLVLTLSK
   AGSTVIGTIGIFALKSPLTPISENSINVEIFF
   DANGEINLTTSSLKSYWGFREGDSYNPSSNLN
   PLYLMPNTYAYPQGRKTITQVFPLEVYLNGDT
   AKPVPLEVAFNTLSSTGFSLEFTWRNLNAYTG
   EAFAVSLGNFTYISQY
```

Fig. 2. Amino acid sequence of the BAdV-2 fiber. **A.** Tail. The highly conserved sequence is underlined. **B.** The 22 repeats of the shaft aligned according to the model described by van Raaij et al. (1999). The characteristic and semi-conserved residues are in bold, the apparently essential glycines or prolines are in bold italics. **C.** Knob. The spacer between the shaft and the knob is printed in italics, the conserved motif is in bold

The central part of the fiber protein, the shaft region, consists of repeating motifs that contain well-conserved proline and glycine residues. Each repeat comprises two extended β-strands which are followed by β-turns in between, that run parallel and backward to the shaft axis (van Raaij et al., 1999). The length variation of the fiber protein detectable among the different serotypes is mainly caused by the differences in the length (i.e. number of repeats) of the shaft region. The fiber of BadV-2 consists of 348 amino acids which can be arranged according to the model described by van Raaij et al. (1999) into 22 repeating motifs (Fig. 2).

The carboxy-terminal knob region of the BAdV-2 fiber protein consists of 176 amino acids and begins by the well-conserved YTLWT motif which is the boundary between the shaft and the knob domain. The knob alignment shows several conserved residues (Fig. 3), whereas the regions showing considerable variability may be specially responsible for cell attachment and thus for tissue tropism and host species specificity. There are several conserved residues in the knob of BAdV-2, -3, and -10 and those may be hypothesised to be responsible for the host specificity (bold and italics in Fig. 3).

```
            A                         B                     C
H2    --KLTLWTTPDPSPNCRIH--SDN-DCKFTLVLTK-CGSQ----VLATVA--ALAV-SGDLSS
H3    --NNTLWTGPKPEANCIIEYGKQNPDSKLTLILVKNGGI-----VNGYVT---LMGASDYVNT
H4    -DKLTLWTTPDPSPNCQIL--AEN-DAKLTLCLTK-CDSQ----ILATVS--VLVVRSGNLNP
H5    --KLTLWTTPAPSPNCRLN--AEK-DAKLTLVLTK-CGSQ----ILATVS--VLAVKGSLAP-
H9    --KRTLWTTPDTSPNCKID--QDK-DSKLTLVLTK-CGSQ----ILANVS---LIVVDGKYKI
H12   --PLTLWTTPDPPPNCSLI--QEL-DAKLTLCLTK-NGSI----VNGIVS---LVGVKGNLLN
H31   --PLTLWTTPDPPPNCTLR--QEL-DAKLTLCLTK-NESI----VNGIVS---LIGVKGDLLH
H40   --PTTLWTTADPSPNATFY--ESL-DAKVWLVLVK-CNGM----VNGTISIKAQKGTLLKPTA
C1    --PVTLWTGPDPNVNASIN--GTP-VIRSFISLTR-DSNL----VTVNAS-FTGEGSYQSVSP
C2    --PITLWTGPGPSINGFIN--DTP-VIRCFICLTR-DSNL----VTVNAS-FVGEGGYRIVSP
P3    --PYTLWTGASPTANVILTNTTTP-NGTFFLCLTR-VGGL----VLG--S-FALKSSIDLTSM
B2    YTDYTLWTTPDPSPNASIK--TDL-DAKLVLTLSKA---GST--VIGTIGIFALKSPLTP-IS
B3    ---ETLWTGTGSNANVTWRGYTAP-GSKLFLSLTR----FSTGLVLGNMTIDSNASFGQY-I-
B10   ---IMLWTTPDPSPNLNLE--GER-TAKLFLSLSY-CNRISDWHLLQ----FGGLKEPVETI-


            D                                         E    F
H2    MTGTVASVS--------IFLRF--DQNGV--LMENSSLKKHY---WNFR-NGNSTNANPYTNA
H3    LFKNKN-VSIN------VELYF--DATGHI-LPDSSSLKTD--LELKYK-QTADF------SA
H4    ITGTVS-SAQ-------VFLRF--DANGV--LLTEHSTLKKY---WGYK-QGDSIDGTPYTNA
H5    ISGTV-QSAH-------LIIRF--DENGV--LLNNSFLDPEY---WNFR-NGDLTEGTAYTNA
H9    INNNT-QPALKGFTI---KLLF--DENGV--LMESSNLGKSY---WNFR-NENSIMSTAYEKA
H12   I-----QSTTTTVG---VHLVF--DEQGR--LITSTPTALVPQASWGYR-QGQSVSTNTVTNG
H31   I-----QPTTTTVG---LHLVF--DRQGR--LVTTTPTALVPQASWGYK-QGQSVSSSAVANA
H40   SFISFVMYFYSDGTWRKNYPVF--DNEGI--LANSAT-------WGYR-QGQSANTN-VSNA
C1    T-----QSQFS------LILEF--NQFGQ--LMSTGNLNSTTT--WGEKPWGNNTVQVQPSHT
C2    T-----QSQFS------LIMEF--DQFGQ--LMSTGNINSTTT--WGEKPWGNNTVQPRPSHT
P3    TKKVN-----------FI--F--DGAGR--LQSDSTYKGRFGFRSNDSVIEPTAAGL--S-P
B2    ENSIN-----------VEIFFDAN--GEINLTTSS-LKSY----WGFREGDSYNP-SSNLNP
B3    -NAGHEQIE-------CFILL---N-QGN--LKEGSNLQGT----WEVKNNPSA---SKAAF-
B10   ENSR--QSFK--VI------LF--NHLGQ--LSGG-NLYGY---FGYRFQTNSVLPNSNIQGT


                                    G                     H
H2    V-GFMPNLLAYPKTQSQTA-KNNIVS---QV-------YLHGDKT--KPMIL-TITL--NGTSE
H3    R-GFMPSTTAYPFVLP-----NAGTHN--ENYIFGQCYYKASDGALF-PLEV-TVML--NKRLP
H4    V-GFMPNSTAYPKTQSSTT-KNNIVG---QV-------YMNGDVS--KPMLL-TITL--NGTDD
H5    V-GFMPNLSAYPKS-HGKTAKSNIVS---QV-------YLNGDKT--KPVTL-TITL--NGTQE
H9    I-GFMPNLVAYPKP-TAGS-KKYARD---IVY---GNIYLGGK-----PDQPVTIKTTFNQE--
H12   L-GFMPNVSAYPRPNASEA-KSQMVS---LT-------YLQGDTS--KPI--TMKVAFNGI-
H31   L-GFMPNVSAYPRPNAGEA-KSQMLS---QT-------YLQGDTT--KPI---TMKVVFNGNA-
H40   V-EFMPSSKRYPNE------KGSEV----QNMALT-YTFLQGD-----PNMAISFQSIYNH-A-
C1    WKLCMPNREVYSTPAATLTSCGLNSI---AH---------DGAPN--RSIDC--MLII-NKLRG
C2    WKLCMPNREVYSTPAATISRCGLDSI---AV---------DGAPS--RSIDC--MLII-NKPKG
P3    AWL-MPSTFIYPRN-----TSGSSLTSF--VYINQTYVHVDIKV--------------NTL--
B2    LYL-MPNTYAYP------GRKTITQVFPLEV------YLN-GDTA--KP-VP-LEVA-FNTL--
B3    --L--PSTALYPILNESRGSLPGKNL----VGMQA-IL-GGGG----TCTVIAT----LNGRR-
B10   LML-MPNSVAYPRVKNNVGNYYETTC-----------YLAGNKY----P-VK-LRVS-LNSD--
```

```
                                       I                J
    H2   STETSEVSTYSMS----FT--WSWESGK--YTTETFATN-SY--------TFSYIA-----QE-
    H3   DSRT----SYVMT----FL--WSLNAGL-APETTQATLITSPF-------TFSYIREDD-----
    H4   T--T---SAYSMS----FSYTWT-N-GS--YIGATFGAN-SY--------TFSYIA-----QQ-
    H5   TGDT-TPSAYSMS----FS--WDW-SGH-NYINEIFATS-SY--------TFSYIA-----QE-
    H9   ---TGCE--YSIT----FDFSWAKT-----YVNVEFETT-SF--------TFSYIA-----QE-
    H12  ---TSLNG-YSLT----FM--WS---GLSNYINQPFSTP-SCS--------FSYIT-----QE-
    H31  ---TVDG--YSLT----FM--WT---GVSNYLNQQFSTP-SCS--------FSYIA-----QE-
    H40  ----IEG--YSLK----FT--WRVR----NNERFDIPCC-S----------FSYVTE----Q--
    C1   AA-T-----YTLT----FRFLNF------NKLSSSTVFKTDVL-------TFTYVGEN---Q--
    C2   VA-T-----YTLT----FRFLNF------NRLSGGTLFKTDVL-------TFTYVGEN---Q--
    P3   --ST---NGYS-E----FNFQNMSFSAPFSTSYGTFCYVPRRTTHRPRHGPFSLRERRHLFQLL
    B2   -SST----GFSLE----FT--WR-NL---NAYTGEAFAVSLGN-------FTYIS------QY-
    B3   -SNN-----YPAGQSIIFV--WQEF----NTIARQPLNH-STL-------FSYWT--------
    B10  -SPMVDWG-YSIT----FE--WYEFD---NHI-G-------------------------
```

Fig. 3. Alignment of the amino acid sequence of the knob of BAdV-2 with the corresponding regions of some other adenoviruses. The codes representing different adenovirus serotypes at the beginning of the lines refer to the different adenoviruses isolated from different species, B: bovine, C: canine, H: human, P: porcine. Conserved residues are printed in bold, conserved residues sequence among these BAdV knob regions are printed in bold and italics (the DNA sequence of the BAdV-10 fiber is unpublished yet, K. Ursu). The underlined sequences of HAdV-2 represent β sheets found in its knob structure, and are designated by capital letters from A to J. The corresponding parts of BAdV-2 show similarity and the same structure of its knob is very probable

The presence of a 17 kD protein gene homologue on the complementary strand was shown by the BLASTX homology search program. This gene has been found in HAdV-2, -3 and -40 so far. The 17 kD protein was described to contribute to the activation of the adenovirus early promoters via the activation of the cellular E2F that enhances the transcription activity on the E2 and E1A promoters (Thomas, 1996). The sequence analysis also revealed a possible splicing site. Occurrence of splicing yielding 17 kD protein was described in HAdV-40 (Davison et al., 1993). The possible splice acceptor sites were searched with the PC/Gene program packages. The intron is supposed to end in the presently sequenced genome part but its start is proposed to be somewhere in the earlier sequenced (Fitzgerald et al., 1997) gene of 34 kD protein homologue. Confirmation of the supposed splice acceptor and donor sites needs further experimental proof.

The comparison of the fiber genes of the two subtypes of BAdV-2 would be interesting, since regions responsible for the different haemagglutination patterns and pathogenicity might be identified.

Acknowledgements

References

Belák, S. and Pálfi, V. (1974): An adenovirus isolated from sheep and its relationship to type 2 bovine adenovirus. Arch. ges. Virusforsch. **46,** 366–369.

Belák, S., Berencsi, Gy., Rusvai, M., Lukács, K. and Nász, I. (1983): DNA structure and hemagglutination properties of bovine adenovirus type 2 strains which bypass species specificity. Arch. Virol. **77,** 181–194.

Belák, S., Virtanen, A., Zabielski, J., Rusvai, M., Berencsi, Gy. and Petterson, U. (1986): Subtypes of bovine adenovirus type 2 exhibit major differences in region E3. Virology **153,** 262–271.

Benkő, M., Bartha, A. and Wadell, G. (1988): DNA restriction enzyme analysis of bovine adenoviruses. Intervirol. **29,** 346–350.

Caillet-Bouding, M. L. (1989): Complementary peptide sequence in partner proteins of the adenovirus capsid. J. Mol. Biol. **208,** 195–198.

Davison, A. J., Telford, E. A. R., Watson, M. S., McBride, K. and Mautner, V. (1993): The DNA sequence of adenovirus type 40. J. Mol. Biol. **234,** 1308–1316.

Dix, I. and Leppard, K. N. (1993): Regulated splicing of adenovirus type 5 E4 transcripts and regulated cytoplasmic accumulation of E4 mRNA. J. Virol. **67,** 3226–3231.

Douglas, J. T., Miller, C. R., Kim, M., Dimitriev, I., Mikheeva, G., Krasnykh, V. and Curiel, D. T. (1999): A system for the propagation of adenoviral vectors with genetically modified receptor specificities. Nature Biotech. **17,** 470–475.

Eiz, B. and Pring-Akerblom, P. (1997): Molecular characterisation of the type-specific gamma-determinant located on the adenovirus fiber. J. Virol. **71,** 6576–6581.

Esford, L. and Haj-Ahmad, Y. (1994): Sequence analysis of the putative E3 region of bovine adenovirus type 2. Intervirol. **37,** 277–286.

Fender, P., Kidd, A., Brebant, R., Oberg, M., Drouet, E. and Chroboczek, J. (1995): Antigenic sites on the receptor-binding sites on the receptor-binding domain of human adenovirus type 2 fiber. Virology **214,** 107–110.

Fitzgerald, L., Ansari, M. and Haj-Ahmad, Y. (1997): Cloning and sequencing of the bovine adenovirus type 2 early region 4. Gene **185,** 181–186.

Mittal, S. K., Prevec, L., Babiuk, L. A. and Graham, F. L. (1993): Sequence analysis of bovine adenovirus type 3 early region 3 and fiber protein genes. J. Gen. Virol. **74,** 2825.

Öhman, K., Nordquist, K. and Akusjarvi, G. (1993): Two adenovirus proteins with redundant activities in virus growth factor facilitates tripartite leader mRNA accumulation. Virology **194,** 50–58.

Rusvai, M., Harrach, B., Bánrévi, A., Evans, P. S. and Benkő, M. (2000): Identification and sequence analysis of the core protein genes of bovine adenovirus 2. Virus Res. **70,** 25–30.

Salmon, K. and Haj-Ahmad, Y. (1994): Sequence analysis of bovine adenovirus type 2 early region 1 and PIX gene. Intervirol. **37,** 298–305.

Salmon, K., Esford, L. and Haj-Ahmad, Y. (1993): Subcloning and restriction enzyme mapping of bovine adenovirus type 2. Intervirol. **36,** 72–78.

Thomas, S. (1996): Adenoviridae: the viruses and their replication. In: Fields, B., Knipe, D. and Howley, P. (eds) Fundamental Virology. Third edition. Lippincott-Raven, New York, pp. 979–1016.

van Raaij, M. J., Mitraki, A., Lavigne, G. and Cusack, S. (1999): A triple beta-spiral in the adenovirus fiber shaft reveals a new structural motif for a fibrous protein. Nature **401,** 935–938.

Yagubi, A., Ojkic, D., Bautista, D. S. and Haj-Ahmad, Y. (1998): Sequencing analysis of the region encoding the DNA polymerase of bovine adenovirus serotypes 2 and 3. Intervirol. **41,** 69–79.