NUCLEOTIDE SEQUENCE OF THE PROTEIN E GENE OF THE TICK-BORNE ENCEPHALITIS VIRUS STRAIN 595 ISOLATED IN SLOVAKIA

M. KALUZOVÁ[1], Š. KALUZ[1], O. KOŽUCH[1], J. PASTOREK[1], M. LABUDA[2]

[1]Institute of Virology, and [2]Institute of Zoology, Slovak Academy of Sciences, Dúbravská cesta 9, 842 46 Bratislava, Slovak Republic

Received April 4, 1995; revised May 26, 1995

Summary. – Tick-borne encephalitis (TBE) virus, strain 595 was isolated from *Ixodes ricinus* ticks in southern Slovakia. A part of the protein E gene was sequenced and compared with the prototype strain Neudorfl. Seventeen silent mutations and two amino acid changes (Ile →Val, residue 167; Asn → Thr, residue 366) were found. The nucleotide homology in the sequenced part of protein E gene of the strain 595 and the prototype strain Neudorfl is 98.6%. These findings indicate that the strain 595 is closely related to the strain Neudorfl.

Key words: tick-borne encephalitis virus; protein E

Tick-borne encephalitis (TBE) virus and other closely related tick-borne viruses form TBE virus subgroup within family *Flaviviridae* (Westaway *et al.*, 1985). In Europe the most prominent endemic areas have been described in Hungary, Austria, Slovenia, Switzerland, Germany, Poland, Sweden, Bohemia and Slovakia. The association of the antigenic types of TBE viruses with tick vectors *Ixodes ricinus* (Western subtype) and *Ixodes persulcatus* (Eastern subtype) was well established (Rubin & Chumakov, 1980). TBE virus strains isolated throughout Europe exhibit a high degree of homogeneity, as revealed by peptide mapping and monoclonal antibodies (Heinz and Kunz, 1981, 1982; Guirakhoo *et al.*, 1987).

The major surface protein of the virus, the E protein (containing about 500 amino acids) mediates several important viral functions during early stages of the viral life cycle, such as the receptor binding and fusion activity. E protein induces the formation of neutralizing and protective antibodies, and active immunization with isolated forms of E protein can provide solid protection against the disease (Heinz *et al.*, 1981). Phenotype changes of flaviviruses are frequently connected with observed mutations in E protein (Holzmann *et al.*, 1990, Cecilia and Gould, 1991). No sequence analysis of a TBE strain isolated from ticks in Slovakia has been performed so far. In order to determine the relationship of the strain 595 to the prototype strain Neudorfl of TBE virus on the nucleotide level we sequenced the gene for protein E of this virus strain.

TBE virus strain 595 was isolated from *Ixodes ricinus* females collected in southern Slovakia, locality Gbelce, in July 1980 (Kožuch *et al.*, 1987, 1990) and passaged in suckling and adult mice. Virus stock was prepared as 10% (w/v) suspension of infected suckling mouse brains.

cDNA synthesis. Total cellular RNA was prepared from infected suckling mouse brain suspensions using the guanidinium thiocyanate method (Chomczynski and Sacchi, 1987). Total cellular RNA (1.7 µg) with antisense primer 5'-TCGACTCCAAGGGTCATGGCCA-3' (600 ng) corresponding to the 3'-end of E protein (nt 2415-2433, Mandl *et al.*, 1988) were heat-denatured at 70 °C for 10 mins and maintained on ice. Reverse transcription was performed with 200 U of Moloney murine virus reverse transcriptase (Pharmacia). The mixture was incubated at 37 °C for 90 mins, then at 94 °C for 5 mins, and cooled on ice.

PCR was performed with 5 µl of the cDNA reaction mixture as a template, 200 ng of sense primer (5'-GATCGCGTTGCACACTTGGA-3', nt 953-972), 200 ng of antisense primer (the same as in cDNA synthesis) in final volume of 50 µl as follows: starting at 95 °C for 5 mins, then 40 cycles at 94 °C for 1 min, 58 °C for 1.10 min, and 72 °C for 1.40 min. The final extension was at 72 °C for 10 mins. The PCR product was analyzed in electrophoresis and the fragment of expected length was purified from agarose gel by Gene Clean (BIO 101).

Cloning and sequencing. The PCR product was phosphorylated with 5 U of T4 polynucleotide kinase (BRL Gibco), made blunt-ended with 2.5 U of Klenow fragment of DNA polymera-

Abbreviations: IPTG = isopropyl-ß-D-thiogalactopyranoside; PCR = polymerase chain reaction; TBC = tick-borne encephalitis

```
          S  R  C   T  H  L  E   N  R  D   F  V  T   G  T  Q  G
Neu     1 UCGCGUUGCA CACACUUGGA AAACAGGGAC UUUGUGACUG GUACUCAGGG
595       .......... .......... .......... .......... ..........

          T  T  R   V  T  L   V  L  E  L   G  G  C   V  T  I
Neu    51 GACUACGAGG GUCACCUUGG UGCUGGAACU GGGUGGAUGU GUUACUAUAA
595       .......... .......... ....A..... .......... .....C....

          I  A  E  G   K  T  S   M  D  V   N  L  D  A   I  Y  Q
Neu   101 CAGCUGAGGG GAAGCCUUCA AUGGAUGUGU GGCUUGACGC CAUUUAGCAG
595       .......... ...A...... .......... .......... ..........

          E  N  P   A  K  T  R   E  Y  C   L  H  A   K  L  S  D
Neu   151 GAGAACCCUG CUAAGACACG UGAGUACUGU UUACACGCCA AGUUGUCGGA
595       .......... .......... .......... .......... ..........

          T  K  V   A  A  R   C  P  T  M   G  P  A   T  L  A
Neu   201 CACUAAGGUU GCAGCCAGAU GCCCAACAAU GGGACCAGCC ACUUUGGCUG
595       .......... .......... .......G.. .......... ..........

          E  E  H  Q   G  G  T   V  C  K   R  D  Q  S   D  R  G
Neu   251 AAGAACACCA GGGUGGCACA GUGUGUAAGA GAGAUCAGAG UGAUCGAGGC
595       .......... ...C...... .......... .......... ...C......

          W  G  N   H  C  G  L   F  G  K   G  S  I   V  A  C  V
Neu   301 UGGGGCAACC ACUGUGGACU GUUUGGAAAG GGUAGCAUUG UGGCCUGUGU
595       .......... .......... .......... .......... ..........

          K  A  A   C  E  A   K  K  K  A   T  G  H   V  Y  D
Neu   351 CAAGGCGGCU UGUGAGGCAA AAAAGAAAGC CACAGGACAU GUGUACGACG
595       .......... .......... .......... .......... ..........

          A  N  K  I   V  Y  T   V  K  V   E  P  H  T   G  D  Y
Neu   401 CCAACAAAAU AGUGUACACG GUCAAAGUCG AACCACACAC GGGAGACUAU
595       .......... .......... ..U....... .......... ..........

          V  A  A   N  E  T  H   S  G  R   K  T  A   S  F  T  I
Neu   451 GUUGCCGCAA ACGAGACACA UAGUGGGAGG AAGACGGCAU CCUUCACAAU
595       .......... .......... .......... .......... .........G.
                                                                 V
          S  S  E   K  T  I   L  T  M  G   E  Y  G   D  V  S
Neu   501 UUCUUCAGAG AAAACCAUUU UGACUAUGGG UGAGUAUGGA GAUGUGUCUU
595       .......... .........C .......... .......... ..........

          L  L  C  R   V  A  S   G  V  D   L  A  Q  T   V  I  L
Neu   551 UGUUGUGCAG GGUCGCUAGU GGCGUUGACU UGGCCCAGAC CGUCAUCCUU
595       .......... .......... ..U....... .......... ..........

          E  L  D   K  T  V  E   H  L  P   T  A  W   Q  V  H  R
Neu   601 GAGCUUGACA AGACAGUGGA ACACCUUCCA ACGGCUUGGC AGGUCCAUAG
595       .......... .......... .......... .......... ........C..

          D  W  F   N  D  L   A  L  P  W   K  H  E   G  A  Q
Neu   651 GGACUGGUUC AAUGAUCUGG CUCUGCCAUG GAAACAUGAG GGAGCGCAAA
595       .......... .......... .......... .......... ..........

          N  W  N  N   A  E  R   L  V  E   F  G  A  P   H  A  V
Neu   701 ACUGGAACAA CGCAGAAAGA CUGGUUGAAU UUGGGGCUCC UCACGCUGUC
595       .......U.. .......... .......... .......... ..........
```

Fig. 1

Comparison of the nucleotide and deduced amino acid sequences of protein E of TBE virus strain 595 to that of the prototype strain Neudorfl

Identical nucleotides are dotted, amino acid changes are shown below the sequence of the strain 595. Neu = Neudorfl.

```
          K  M  D   V  Y  N  L   G  D  Q   T  G  V   L  L  K  A
Neu  751  AAGAUGGACG UGUACAACCU CGGAGACCAG ACUGGAGUGU UACUGAAGGC
595       .......... .......... .......... .......... ..........

          L  A  G   V  P  V   A  H  I  E   G  T  K   Y  H  L
Neu  801  UCUCGCUGGG GUUCCUGUGG CACACAUUGA GGGAACCAAG UACCACCUGA
595       ..........  .....C.... .......... .......... ..........

          K  S  G  H   V  T  C   K  V  G   L  E  K  L   K  M  K
Neu  851  AGAGUGGCCA CGUGACCUGC GAAGUGGGAC UGGAAAAACU GAAGAUGAAA
595       ... ......  U.....A... .......... .......... ..........

          G  L  T   Y  T  M  C   D  K  T   K  F  T   W  K  R  A
Neu  901  GGUCUUACGU ACACAAUGUG UGACAAAACA AAGUUCACAU GGAAGAGAGC
595       .......... .......... .......... .......... ..........

          P  T  D   S  G  H   D  T  V  V   M  E  V   T  F  S
Neu  951  UCCAACAGAC AGUGGGCAUG AUACAGUGGU CAUGGAAGUC ACAUUCUCUG
595       .......... .......... .C........ .......... ..........

          G  T  K  P   C  R  I   P  V  R   A  V  A  H   G  S  P
Neu 1001  GAACAAAGCC CUGUAGGAUC CCAGUCAGGG CAGUGGCACA UGGAUCUCCA
595       .......... .......... .......... .......... ..........

          D  V  N   V  A  M  L   I  T  P   N  P  T   I  E  N  N
Neu 1051  GAUGUGAACG UGGCCAUGCU GAUAACGCCA AACCCAACAA UUGAAAACAA
595       .......... .......... .......... .......... .......C...
                                                              T
          G  G  G   F  I  E   M  Q  L  P   P  G  D   N  I  I
Neu 1101  UGGAGGUGGC UUCAUAGAGA UGCAGCUGCC CCCAGGGGAU AACAUCAUCU
595       .......... .......... .......... .......... ..........

          Y  V  G  E   L  S  H   Q  W  F   Q  K  G  S   S  I  G
Neu 1151  AUGUUGGGGA ACUGAGUCAU CAAUGGUUCC AAAAAGGGAG CAGCAUCGGA
595       .......... .,........ .......... .......... ..........

          R  V  F   Q  K  T  K   K  G  I  E  R  L  T   V  I  G
Neu 1201  AGGGUUUUCC AAAAGACCAA GAAAGGCAUA GAAAGAGUGA CAGUGAUAGG
595       .......... .......... .......... .......... ..........

          E  H  A   W  D  F   G  S  A  G   G  F  L   S  S  I
Neu 1251  AGAGCACGCC UGGGACUUCG GUUCUGCUGG AGGCUUUCUG AGUUCAAUUG
595       .......... .......... .......... .......... ..........

          G  K  A  V   H  T  V   L  G  G   A  F  N  S   I  F  G
Neu 1301  GGAAGGCGGU ACAUACGGUC CUUGGUGGCG CUUUCAACAG CAUCUUCGGG
595       .......... G......... .........U. .......... ..........

          G  V  G   F  L  P  K   L  L  L   G  V  A   L  A  W  L
Neu 1351  GGAGUGGGGU UUCUACCAAA ACUUUUAUUA GGAGUGGCAU UGGCUUGGUU
595       ...

          G  L  N   M  R  N   P  T  M  S   M  S  F   L  L  A
Neu 1401  GGGCCUGAAC AUGAGAAACC CUACAAUGUC CAUGAGCUUU CUCUUGGCUG

          G  G  L  V   L  A  M   T  L  G   V  G  A
Neu 1451  GAGGUCUGGU CUUGGCCAUG ACCCUUGGAG UGGGGGCG
```

Fig. 1

se I (Boehringer), cloned into *Eco*RV-cut pBluescript KS$^-$ (Stratagene) and transfected into competent *E. coli* DH5αMCR cells (Gibco BRL) utilizing white blue X-gal and IPTG (Boehringer) selection. Putative positive clones were selected by the quick-screen method comparing mobilities of covalently closed circular forms of plasmids. After digestion with *Sal*I and *Xba*I (Boehringer) positive clones containing the fragment of expected size were chosen and sequenced (T7 sequencing kit, Pharmacia) by the dideoxy termination method (Sanger *et al.*, 1977), using double-stranded template and gene-specific primers. At first, the sequence of a single clone was determined and mutations against the prototype sequence were confirmed by sequencing another clone.

The sequencing of 1353 nucleotides of the protein E gene of TBE virus strain 595 revealed 19 nucleotide changes as compared to the prototype TBE strain Neudorfl (Mandl *et al.*, 1988; Fig. 1) resulting in two amino acid changes: A→G transition at nucleotide position 499 leading to amino acid substitution of isoleucine by valine, and A→C transversion at nucleotide position 1097 leading to amino acid substitution of asparagine by threonine. The former mutation retained the neutral non-polar character of this locus and occured within the region of variable amino acid residues of domain C. The latter mutation preserved neutral polar amino acid at this position and was located within domain B occuring also within the region of variable amino acid residues (Holzmann *et al.*, 1990). The nucleotide homology in the sequenced region of protein E gene between the strain 595 and the prototype strain Neudorfl was 98.6%, while the deduced amino acid homology was 99.56%.

The homology of strains 595 and 4387 (isolated from the bank vole in the same locality Gbelce, southern Slovakia) was 99.5% at the nucleotide level with the single deduced amino acid substitution of isoleucine by threonine at the amino acid position 319 (U→C transition at nucleotide position 956) and one silent mutation.

The prototype strain Neudorfl of TBE virus was isolated in Burgenland, Austria from *I. ricinus* ticks (Heinz *et al.*, 1982). Strain 595 was also isolated from *I. ricinus* ticks. The first of two amino acid mutations disclosed in the protein E gene (position 499, Ile→Val) was found in these tick-borne flaviviruses: TBE virus, strain Sofjin (Pletnev *et al.*, 1990), Langat virus, strain TP 21 (Mandl *et al.*, 1991), louping ill virus (Shiu *et al.*, 1991), and TBE virus strain 4387 (Labuda *et al.*, 1994).

In contrast to these the findings, second mutation (Asn→Thr) seems to be specifically linked with the geographic area where the virus strain was detected. Apart from strain 4387 isolated from organs of a bank vole in the same locality (Kožuch *et al.*, 1990), no other TBE strain sequenced so far harbors this mutation. Overall 98.6% nucleotide homology of the prototype strain Neudorfl and the strain 595 clasifies these two strains as closely related. According to the computer analysis it is unlikely that these mutations would have any impact on biochemical, structural and antigenic features of protein E. It is interesting that geographically more remote strain Kumlinge A52, isolated in Finland (Whitby *et al.*, 1993) has only one amino acid change of valine for isoleucine at the amino acids position 167 (A→G transition, nucleotide position 499) as compared to the prototype strain Neudorfl protein.

References

Cecilia D, Gould EA (1991): Nucleotide changes responsible for loss of neuroinvasiveness in Japanese encephalitis virus neutralization-resistant mutants. *Virology* **181**, 70–77.

Chomczynski P, Sacchi N (1987): Single-step method of RNA isolation by acid quanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156–159.

Guirakhoo F, Radda AC, Heinz FX, Kunz C (1987): Evidence for antigenic stability of tick-borne encephalitis virus by the analysis of natural isolates. *J. Gen. Virol.* **68**, 859–864.

Heinz FX, Tuma W, Kunz C (1981): Antigenic and immunogenic properties of defined physical forms of tick-borne encephalitis virus structural proteins. *Infect. Immun.* **33**, 250–257.

Heinz FX, Kunz C (1981): Homogeneity of the structural glycoprotein from European isolates of tick-borne encephalitis virus: Comparison with other flaviviruses. *J. Gen. Virol.* **57**, 263–274.

Heinz FX, Kunz C (1982): Molecular epidemiology of tick-borne encephalitis virus: peptide mapping of large non-structural proteins of European isolates and comparison with other flaviviruses. *J. Gen. Virol.* **62**, 271–285.

Heinz FX, Berger R, Majdic O, Knapp W, Kunz C (1982): Monoclonal antibodies to the structural glycoprotein of tick-borne encephalitis virus. *Infect. Immun.* **37**, 869–874.

Holzmann H, Heinz FX, Mandl CW, Guirakhoo F, Kunz C (1990): A single amino acid substitution in envelope protein E of tick-borne encephalitis virus leds to attenuation in the mouse model. *J. Virol.* **64**, 515-6-5159.

Kožuch O, Lysý J, Guryčová D, Labuda M, Vyrostekova V (1987): Mixed natural focus of arboviruses and tularemia in South Slovakia. *Biológia (Bratislava)* **42**, 1083–1089.

Kožuch O, Labuda M, Lysý J, Weismann P, Krippel E (1990): Longitudinal study of natural foci of Central European encephalitis virus in West Slovakia. *Acta Virol.* **34**, 537–544.

Labuda M, Jiang WR, Kaluzová M, Kožuch O, Nuttal PA, Weismann P, Elečková E, Žuffová E, Gould EA (1994): Change in phenotype of tick-borne encephalitis virus following passage in *Ixodes ricinus* ticks is accompanied by amino acid substitution in viral envelope protein. *Virus Res.* **31**, 305–315.

Mandl CW, Heinz FX, Kunz C (1988): Sequence of the structural proteins of tick-borne encephalitis virus (Western subtype) and comparative analysis with other flaviviruses. *Virology* **166**, 197–205.

Mandl CW, Iacono-Connors L, Wallner G, Holzmann H, Kunz C, Heinz FX (1991): Sequence of the genes encoding the structural proteins of the low-virulence tick-borne flaviviruses Langat TP-21 and Yelantsev. *Virology* **185**, 891–895.

Pletnev AG, Yamaschikov VF, Blinov VM (1990): Nucleotide sequence of the genome and complete amino-acid sequence of the polyprotein of tick-borne encephalitis virus. *Virology* **174**, 250–263.

Rubin SG, Chumakov MP (1980): New data on the antigenic types of tick-borne encephalitis (TBE) virus. In *Arboviruses in the Mediterranean Countries*. Fisher, Stuttgart and New York, pp. 231–236.

Sanger F, Nicklen S, Coulson AR (1977): DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.

Shiu SYW, Ayres MD, Gould EA (1991): Genomic sequence of the structural proteins of louping ill III virus: Comparative analysis with tick-borne encephalitis virus. *Virology* **180**, 411–415.

Westaway EG, Brinton MA, Gaidamovich SY, Horzinek MC, Igarashi A, Kaariainen L, Lvov DK, Porterfield JS, Russell PK, Trent DW (1985): Flaviviridae. *Intervirology* **24**, 183–192.

Whitby JE, Jennings AD, Barrett ADT (1993): Nucleotide sequence of the envelope protein gene of the tick-borne flaviviruses, Kumlinge A52. *Virus Genes* **7**, 145–149.