VIRUS GENES 2:3, 283–291, 1988
© Kluwer Academic Publishers, Manufactured in The Netherlands

Sequences of Capsid Protein VP1 of Two Type A Foot-and-Mouth Disease Viruses

OTFRIED MARQUARDT AND KARL-HEINZ ADAM
Federal Research Centre for Virus Diseases of Animals, P.O. Box 1149, D-7400 Tübingen, Federal Republic of Germany

Received October 19, 1988
Accepted November 25, 1988

Requests for reprints should be addressed to Otfried Marquardt, Federal Research Centre for Virus Diseases of Animals, P.O. Box 1149, D-7400 Tübingen, Federal Republic of Germany.

Key words: FMDV A_5Bernbeuren/1984, FMDV A Iran/1987, VP1, sequences, 2-D structure

Abstract

We have sequenced the nucleotides of the regions that encode the capsid protein VP1 of the foot-and-mouth disease viruses (FMDV) A_5Bernbeuren/1984 and A Iran/1987. Amino acid sequences and secondary protein structures are provided. Both proteins consist of 212 amino acids. The sequences and secondary structures are compared to those of FMDV A_{22}/CCCP/64, a strain previously endemic in the Near East. Nucleotide divergency among the three sequences is highest for FMDV A_5Bernbeuren/1984 (18% compared to 13% for each other case). Thirty amino acid divergencies are observed between A_{22}/CCCP/64 and A_5 Bernbeuren/1984 or A Iran/1987, whereas the latter two differ by 27 residues. The secondary structures of all three proteins are different. A Iran/1987 is considered to belong to a thus far unknown subtype.

Introduction

Foot-and-mouth disease viruses (FMDV) infect cloven-hooved animals. The disease is distributed worldwide. Seven serotypes of FMDV are distinguished: A, O, C, Asia 1, South African Territory (SAT) 1, SAT2, and SAT3. Each serotype comprises a varying number of immunologically distinguishable subtypes, which leads to a number of more than 70 (for review see 1).

Since the major immunogenic sites of FMDV have been found to reside in distinct parts of the capsid protein VP1 (2), extensive efforts have been made to describe its nucleotide and amino acid sequences (3-11). The aim of the descriptions is the understanding of the antigenic variation among related subtypes or among

individual isolates of one subtype. Knowledge of the VP1 sequences of FMDV strains that prevail in certain areas of the world may provide additional information on FMDV evolution.

In 1987 a FMDV outbreak occurred among cattle exported from the Federal Republic of Germany to Iran. The Iranian authorities suspected that the virus was introduced by the cattle from Germany. Therefore, the nucleotide and amino acid sequences of the Iranian virus isolate were compared with those of the most recent outbreak of FMD (A_5Bernbeuren) in the Federal Republic of Germany in 1987 (12) and with the equivalent sequence of FMDV A_{22}/CCCP (9), which was isolated in 1964 in the Near East.

Materials and methods

Virus propagation

The viruses A_5Bernbeuren and Iran/1987 were grown in baby hamster kidney cells of the clone Tübingen (13). Dense monolayers of cells grown in rolling tubes (6 × 35 cm) were infected with 10 ml of virus at a multiplicity of 2. After 1 hr of adsorption maintenance, medium (Glasgow modification of Eagle's basal medium supplemented with tryptose phosphate broth) was added until there was 50 ml per tube and incubation was continued at 37°C.

Extraction and analysis of RNA

Infected cells were harvested about 3.5 hr after infection when a beginning cytopathic effect became visible, the exact time depending upon the virus strain. The medium was discarded and the flasks were chilled. Then the cells were detached from the glass and lysed in 20 ml of 25 mM sodium citrate buffer, pH 7.0, containing 4 M acid guanidinium thiocyanate, 0.5% sarcosyl, and 1 mM dithiotreitol per rolling tube.

Total cellular RNA was extracted according to a recently published protocol (14). It consists briefly of phenol extraction of cell lysate (1 volume of phenol, chloroform, and isoamylalcohol, 25:24:1, pH 4.8) and precipitation of RNA from the aqueous phase by isoamylalcohol (1 volume). The precipitate was resuspended in guanidinium thiocyanate solution and reprecipitated by isoamylalcohol. Salts were removed by resuspension of the precipitate in H_2O and subsequent ethanol precipitation. The final solution of RNA in H_2O was mixed with 5 units of RNase inhibitor (Boehringer, Mannheim, FRG) per milligram of RNA. RNA concentration was deduced from the UV absorption at 260 nm. Analytical horizontal agarose gel electrophoresis of glyoxal-denatured RNA samples (20 μg/slot) was carried out as described by Thomas (15). FMDV RNA was routinely detected as a discrete band that migrates slightly more slowly than RNA purified from polio

virus kindly provided by Dr. Dernick, Hamburg (results not shown).

Synthesis of oligonucleotide primers

Oligodeoxynucleotide primers were synthesized with an automated DNA synthesizer (Biosearch 8700; New Brunswick, Heusenstamm, FRG), following the protocol recommended by the supplier. The sequences of the primers are complementary to parts of the VP1-encoding RNA of the FMDV strains A_5Westerwald (A_5Ww) and O_1Kaufbeuren (O_1K), which are highly conserved according to the available data (11). The primers are described in Table 1.

cDNA synthesis in the Sanger reaction (16)

The cDNA reactions were carried out with 100 µg total RNA of FMDV-infected cells. Air-dried ethanol precipitates were resuspended in 10 µl transcript system buffer (50 mM NaCl, 34 mM Tris-HCl, pH 8.3, 6 mM $MgCl_2$, 5mM dithiotreitol, 0.1% Triton X-100) and 0.2 µl primer (5 ng), being thus present in an approximately twofold molar excess over FMDV template RNA. Following incubations for 10 min at 60°C and for 30 min at 42°C, usually 60 µCi of [α ^{35}S]dATP (100 pmol) were added. The mixture was divided into four 2.75 µl aliquots, and one out of the four transcript system nucleotide mixtures containing different dideoxy nucleotides (Promega Biotec, Madison, Wisconsin, USA) was added in 3 µl volumes to each aliquot. Five units of reverse transcriptase (Promega Biotec) in 0.5 µl were added, and cDNA synthesis was allowed to proceed for 20 min at 42°C. Following a 15 min chase reaction with highly concentrated dNTPs, the reaction was stopped by 10 min incubation at 42°C with 2 µg RNase A and by adjustment to 37% formamide, 8 mM EDTA, and 0.12‰ each of bromophenol blue and xylencyanol. The samples were now ready for application on sequencing gels.

Table 1. FMDV VP1-specific primer oligodeoxynucleotide sequences

Strain	Codons	Complementary to sequence (5′ to 3′)
O_1K	"225-231"*	CAA GGG CCC AGG GTT GGA CTC
A_5Ww	171-180	TT CAT GCG CAC GAG AAG CTC GTG GAT GG
A_5Ww	126-133	T CCC GTT GTA CAC GGT TGC CAA
O_1K	74- 80	C TGC TAT CTC CAA GTC AG
A_5Ww	36- 41	GTT CAC AAA TCT GTC CAT

*C-terminally outside VP1

Sequence analysis

1.5 µl aliquots of each cDNA reaction were loaded onto polyacrylamide-urea gels, which ran for 2 to 4 hr at 75 W in a LKB (Freiburg, FRG) unit. We found gels of 0.4 mm thickness and 6% acrylamide most suitable when using combs with 36 teeth. Resolution of cDNA products by electrophoresis was controlled and increased by sample application at 2 hr intervals. Gels were soaked for 15 min in 10% acetic acid to remove urea, dried at 70°C for 45 min, and exposed to autoradiography overnight. X-ray films (XAR5, Kodak) were automatically developed by a Sakura QX60 unit. Autoradiograms allowed the reading of 80 to 180 nucleotides beginning 1 to 13 nucleotides upstream of the primer. Each nucleotide was determined at least twice by independent experiments. Computer analysis of nucleic acid sequences was carried out by using the genetics computer group sequence analysis software package (17), version 5.3, on a Microvax station (Digital Equipment Corporation, Munich, FRG).

Results and discussion

The nucleotide sequences that encode the capsid protein VP1 of FMDV A_5Bernbeuren/1984 and A Iran/1987 have been described almost completely and are presented in Fig. 1 in a comparison with the analogous sequence of FMDV A_{22}/CCCP/64 (9). Preliminary descriptions of partial sequences (11,18) are thus confirmed and substantially extended. Five silent nucleotide variances within the A_5/Bernbeuren/1984 sequences were, however, observed (codons 139, 144, 145 U instead of C, 190, and 191 C instead of A and U). Such differences are probably due to different passage history and have also been found in other A_5 isolates (9). The homology between the nucleotide sequences of A_{22}/64 and A Iran/1987 is higher (87%) than that between the sequences of A_5Bernbeuren/1984 and A_{22}/64 or A Iran/1987 (81% and 83%, respectively). The A Iran/1987 sequence was further compared with other available type A FMDV sequences (4,5,9) and found to be dissimilar to more than at least 12% (not shown).

Despite two unresolved nucleotides, translation of the VP1-specific sequence of A Iran/1987 is feasible. Efforts to determine codons 2 to 6 of A_5Bernbeuren/1984, however, were unsuccessful. Precise determination of certain parts of the VP1-encoding sequence of a variety of FMDV strains was also found to be difficult by others (9-11). Fortunately, the sequence of A_5Westerwald/1951 is available (9,11). The sequence of its codons 2 to 6 is ACU GCU GUU GGG GAG and is identical to the one we resolved (Fig. 1). It is assumed that both A_5 isolates are identical in their 5'-terminal sequence.

A comparison of the amino acid sequences of VP1 of A_{22}/64, A Iran/1987, and A_5Bernbeuren/1984 is shown in Fig. 2. The VP1 sequences of A Iran/1987 and A_5Bernbeuren/1984 comprise 212 amino acids and are therefore shorter than that of A_{22}/64 due to the lack of one residue at position 142. VP1 sequences of 212

Comparison of sequences encoding FMDV seotype A VP1

```
            1                                                  31
A22/CCCP/64 AAC ACU ACC ACC GGG GAG UCU GCA GAC CCU GUC ACC ACC ACC GUU GAA AAC UAC GGC GGU
A/Iran/87   .C. ..C G.U .. ... ... ..C ... ... ..C ... ... ... ... ... ..G ... ... ..U ...
A5/Bb/84    .CU .. .U UU .. .. .. ..C ... ... ... ... ... ... ... ..G ..G ... ... ... ...

           61                                                  91
A22/CCCP/64 GAG ACA CAA GUC CAA CGA CGU CAG CAC ACC GAC GUU ACU UUC AUA AUG GAC AGA UUU GUA
A/Iran/87   ... ... ... ... .GG ..G ... ... ... ..U ... ..C .GC ..U ... ... ... ... ... ...
A5/Bb/84    ... ... ... AC. ..G A.. ..G ..C ... ..G ..U ..C GG. ... ..C ... ... ... ... ..G

          121                                                 151
A22/CCCP/64 AAG AUA CAA AAU UUG AAC CCC AUA CAU GUC AUU GAC CUC AUG CAA ACC CAC CAA CAC GGG
A/Iran/87   ..A ..U A.C CC. G.. GC. ... .C. ... ... ... ... ... ... ..G ... ... ... ... .C.
A5/Bb/84    ... ... A.C .G. ... .GU ..U .CG ... ... ... ... ... ... ..G ... ... ..G ... ...

          181                                                 211
A22/CCCP/64 UUG GUA GGU GCC CUG UUA CGU GCU GCU ACG UAC UAC UUC UCU GAC CUG GAG AUU CUG GUA
A/Iran/87   ... ... ..C ..U ... ..G ... ..G ... ... ... ... ... ... ... ... ... ... G.. ...
A5/Bb/84    C.  ..G ... ..G ... ..G ... ..A ..C ... ... ... ... ... ... U.. ... ... G.U ..G

          241                                                 271
A22/CCCP/64 CGC CAU GAC GGU AAC CUA ACC UGG GUA CCC AAC GGA GCA CCC GAG GCA GCU CUG UCU AAC
A/Iran/87   ... ... ..A ..C ... ..G ..G ... ..G ..U ... ... ... ... ... ... ... ..C ..U
A5/Bb/84    ..G ... ... ..C ..U U.G ..U ... ..G ... ... ..U ..C ..U ..A ... ... U.. ..A ..U

          301                                                 331
A22/CCCP/64 AUG GGC AAC CCC ACC GCC UAC CCC AAG GCA CCA UUU ACG AGG CUC GCG CUC CCC UAC ACC
A/Iran/87   .CA A.. ... ... ..U ... ... AA. ... .A. ... ... ... ... ... ... ... ... ... ...
A5/Bb/84    .CC A.. ... ... ..U ... ... AA. ... ... ..G ..C ... ... ... ..U ... ..U ... ..U

          361                                                 391
A22/CCCP/64 GCG CCA CAC CGC GUA UUG GCG ACA GUG UAC AAC GGG ACG AGC AAG UAC UCC GCA GGU GGU
A/Iran/87   ... ... ... ..U ..G ... ..C ... ... ... ... ... ... .A. ... ... G.U ... ACC ...
A5/Bb/84    ... ... ... ... ..G ... ..A ..C ... ... ... ... ... .A. ... ... ... A.G .A. ...

          421                                                 451
A22/CCCP/64 AUG GGC AGA CGG GGC GAC CUA GAG CCU CUC GCG GCG AGG GUC GCC GCU CAG CUU CCU ACU
A/Iran/87   GCU *** ... ... ... ... ... .G. U.. ... ... ..A ... ... ... ..G ..A ..C ... U..
A5/Bb/84    CC. *** ... ..U ..U ... G.G .G. U.G ... ... ... C.. .C. ..G AAA ..A ... ... G.C

          481                                                 511
A22/CCCP/64 UCU UUC AAC UUU GGU GCA AUU CAA GCC ACG ACC AUC CAC GAG CUC CUC GUG CGC AUG AAG
A/Iran/87   ..C ..U ... ... ..G ... ... .GG ... ... ... ... ... ... ..U ... ... ... ... .G.
A5/Bb/84    ... ..U ... .AC ... ... ..C AGG ... GAC ... ... ... ... ..U ... ... ... ... ..A

          541                                                 571
A22/CCCP/64 CGU GCC GAA CUC UAC UGC CCC AGA CCA CUG UUG GCA GUG GUG GUG UCG UCU CAA GAC AGA
A/Iran/87   ..C ... ..G ... ... ... ... ..G ..G ... C.. ... A.. .A. ... ... G.G G.. .G. ...
A5/Bb/84    ..G ..A ..G ... ... ... ... ..G ... ..A C.. ... A.A .A. ... ..U ..A ... ... ..G

          601                                                 631
A22/CCCP/64 CAC AAA CAG AAG AUC AUU GCA CCU GCA AAA CAA CUU UUG
A/Iran/87   ... ... ... ... ... ... ... ... ... ..G ..G ... C.
A5/Bb/84    ... ..G ..A ... ... ... ... ... ... .G. ..G U.G C..
```

Fig. 1. The nucleotide sequence encoding capsid protein VP1 of FMDV A_{22}/CCCP/64 (9) is shown and numbered; those of A Iran/1987 and A_5Bernbeuren/1984 (A_5/Bb/1984) are aligned with it. Sequence redundancy is indicated by dots, and sequence deviations are indicated by letters, which abbreviate nucleotides. Asterisks mark gaps in sequences; no symbol is given where a nucleotide could not be determined.

residues have been described for European isolates (5,11) and South American isolates (A_{27}/Colombia/1976; 9), and those of 213 residues have been described for other South American isolates (9). The length of VP1 does not, therefore, indicate a close relationship of certain FMDV isolates.

The sequence of A_{22}/64 diverges by 30 amino acids each from those of A_5Bernbeuren/1984 and A Iran/1987; the sequences of A_5Bernbeuren/1984 and A Iran/1987 diverge by 27 amino acids. Evolutionary divergence of all three viruses is thus indicated. The differences are predominantly located where hypervariability has already been observed (9,11): at the amino terminus, between residues 23 and 33, 41 and 49, and 133 and 150. In addition, each virus exhibits sequence individuality elsewhere: A_5Bernbeuren/1984 between residues 153 and 171,

```
Comparison of FMDV serotype A VP1 sequences

                1
A22/CCCP/64     NTTTGESADP VTTTVENYGG ETQVQRRQHT DVTFIMDRFV KIQNLNPIHV
A/Iran/87       T.A...... .......... ....R..... ..S....... ..NPVA.T..
A5/Bb/84        T.av...... .......... ...T...H.. ..G....... ..NS.S.T..

               51
A22/CCCP/64     IDLMQTHQHG LVGALLRAAT YYFSDLEILV RHDGNLTWVP NGAPEAALSN
A/Iran/87       .........A .......... ........V. ..E....... ..........
A5/Bb/84        .......... .......... ........V. .......... ..........

              101
A22/CCCP/64     MGNPTAYPKA PFTRLALPYT APHRVLATVY NGTSKYSAGG MGRRGDLEPL
A/Iran/87       TS.....N.E .......... .......... ...N..A.T. A*.....GS.
A5/Bb/84        TS.....N.. .......... .......... ...N...TD. P*....VGS.

              151
A22/CCCP/64     AARVAAQLPT SFNFGAIQAT TIHELLVRMK RAELYCPRPL LAVVVSSQDR
A/Iran/87       .........S .......R.. .........R .......... ..ME..AEG.
A5/Bb/84        ...A.K...A ...Y...R.D .......... .......... ..IE......

              201
A22/CCCP/64     HKQKIIAPAK QLL
A/Iran/87       .......... ...
A5/Bb/84        .........R ...
```

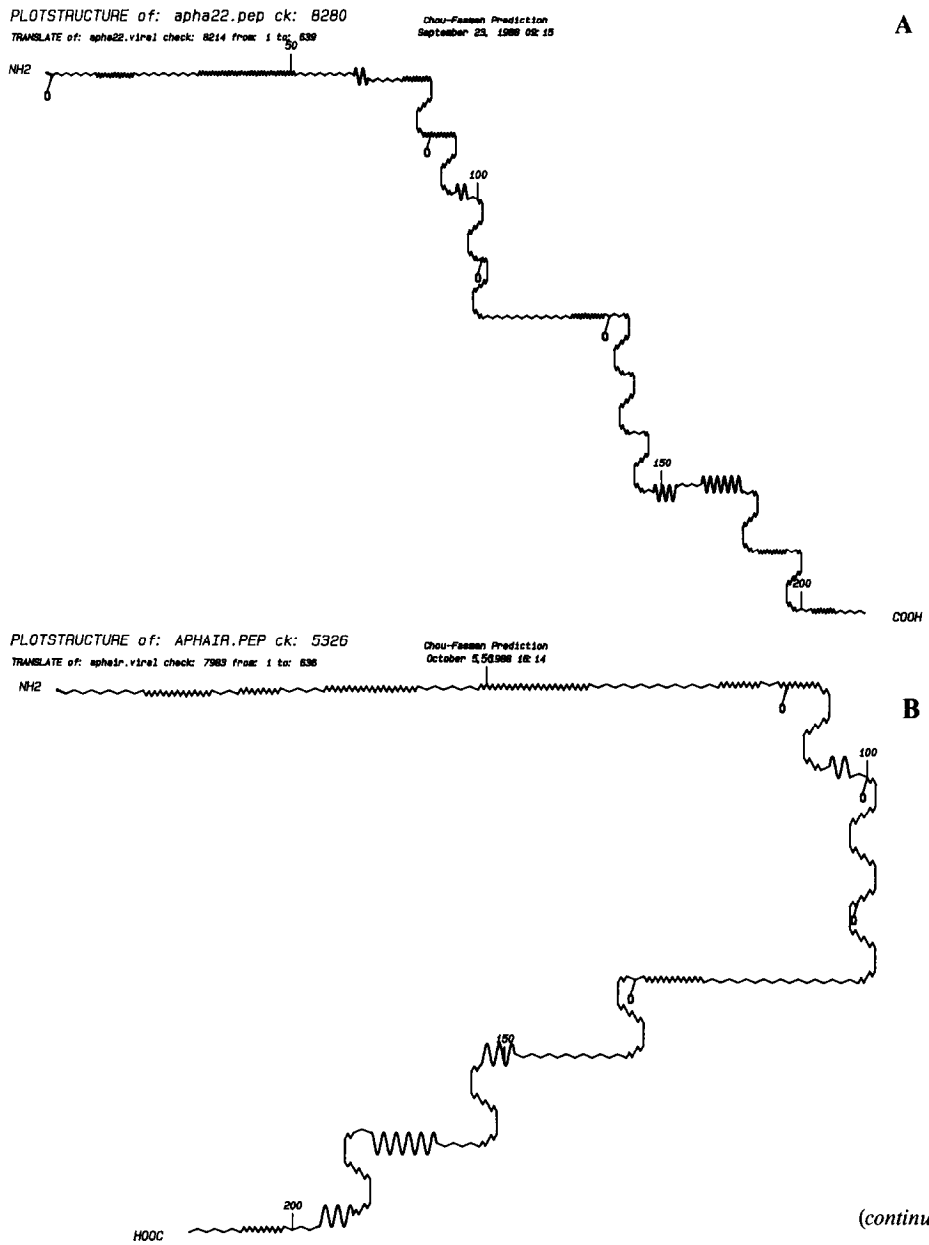
Fig. 2. Amino acids are given in the one-letter code for VP1 of FMDV A_{22}/CCCP/64 (9) and numbered. The sequences of A Iran/1987 and A_5Bernbeuren/1984 (A_5/Bb/1984) are aligned with it. Sequence deviations are indicated by letters, and sequence redundancy is indicated by dots. Asterisks mark gaps in sequences. Amino acids 2 to 6 of A_5/Bb/1984 have actually been determined for A_5Westerwald/1951 (9,11). Sequence deviations are therefore indicated by small letters. The major neutralization epitope is formed by amino acids 134 to 145.

and A Iran/1987 between residues 192 and 200. Amino acids 133 to 150 of VP1 form a neutralization epitope (2) The differences found here make cross-neutralization unlikely. In fact, anti-A Iran/1987 serum is poor in neutralizing the other viruses mentioned (R. Ahl, Tübingen, personal communication).

The significance of amino acid exchanges within related sequences is partially revealed by the calculation of their secondary structure, for example, according to Chou and Fasman (19). The secondary structures of VP1 in all three viruses were therefore calculated and are shown in Fig. 3 (A = A_{22}/64; B = A Iran/1987; C = A_5Bernbeuren/1984). The aminoterminal halves of A_5Bernbeuren/1984 and A_{22}/64 VP1 exhibit substantial similarity, whereas the analogous structure of A Iran/1987 differs. The difference is mainly due to alanine instead of glycine at position 60 (Fig. 2), which favors elongation of the beta-sheet structure beginning at position 49, to the disadvantage of the alpha-helical structures that might be formed by residues 64 to 70.

The secondary structures of the carboxy-terminal halves of all three proteins differ to a greater extent than their amino-terminal halves, although amino acid exchanges are equally scattered in both halves. This is consistent with the finding herein of major immunogenic epitopes (2).

Description of the VP1 secondary structures of six different FMDV subtypes and different isolates of two of them are available (20). The comparison led to the conclusion that the secondary structure predictions of different subtypes of the



(*continued*)

Fig. 3. The VP1 primary sequences of A_{22}/CCCP (A), A Iran/1987 (B), and A_5/Bb/1984 were subjected to the peptide-structure and plot-structure programs included in the genetics computer group sequence analysis software package (17) and were plotted on an HP 7550. Amino (NH_2) and carboxy (COOH) termini, as well as each 50th amino acid, are indicated. Beta sheets are plotted as ~~~~, alpha helices as /\/\/\, and turns as ⌒. The programs further indicate putative glycosylation sites, however irrelevant in this context, as side chains.

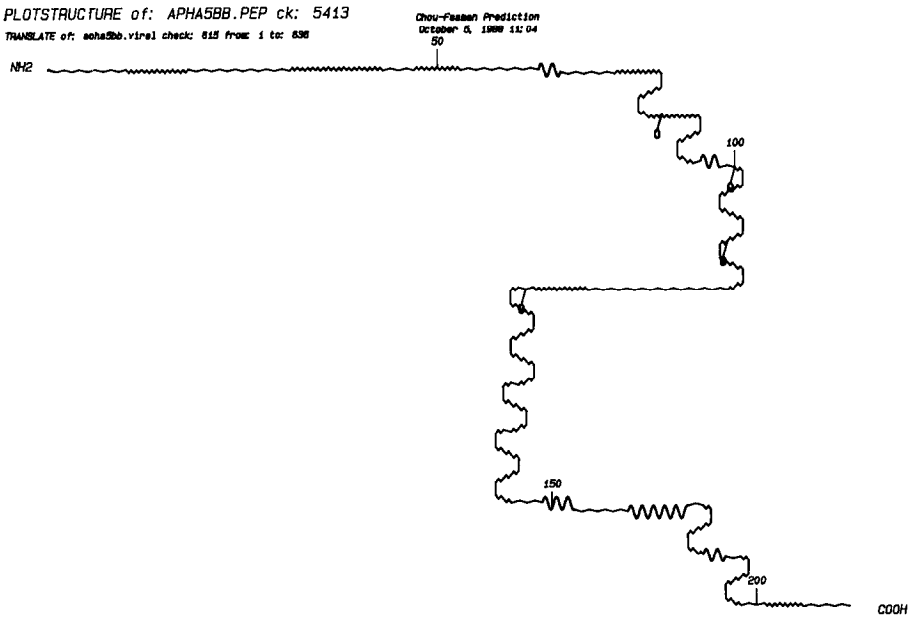


Fig. 3. (Continued)

same serotype differ significantly, whereas those of different isolates of one subtype are very similar. Our study supports such a conclusion. The plots shown in Fig. 3 are dissimilar, whereas those of A_5Westerwald/51 (20) and A_5Bernbeuren/1984 are almost identical (not shown).

Such secondary structure predictions are of limited reliability, as revealed by comparison with three-dimensional picornavirus structures (21,22). All capsid proteins have thus been found to be wedge-shaped due to eight-stranded antiparallel beta barrels, despite considerable heterogeneity in amino acid sequences. The secondary structures shown in Fig. 3 do not indicate the beta sheets between amino acids 151 and 200 observed in Mengo virus crystals. However, FMDV VP1 amino acids 130 to 155 have been found to loop out of the protein. The structure of the loop may therefore be truly described by calculations, according to Chou and Fasman (19).

Based upon the comparison of primary and secondary protein structures, we suggest considering FMDV A Iran/1987 as belonging to a thus far undescribed subtype of FMDV serotype A.

Acknowledgments

We thank Drs. R. Dernick, Hamburg, for the supply with purified RNA from polioviruses and R. Ahl, Tübingen, for the delivery of viruses A_5Bernbeuren/1984

and A Iran/1987 as well as of baby hamster kidney cells. We further thank B. Hohnsbeen, Tübingen, for support in computer analysis.

References

1. Brooksby J.B., Intervirology *18:*1-23, 1982.
2. Strohmaier K., Franze R. and Adam K.-H., J Gen Virol *59:*295-306, 1982.
3. Kurz C., Forss S., Küpper H., Strohmaier K. and Schaller H., Nucl Acids Res *9:*1919-1931, 1981.
4. Makoff A.J., Paynter C.A., Rowlands D.J. and Boothroyd J.C., Nucl Acids Res *10:*8285-8295, 1982.
5. Boothroyd J.C., Harris T.J.R., Rowlands D.J. and Loeve P.A., Gene *17:*153-161, 1982.
6. Cheung A., Delamarter J., Weiss S. and Küpper H., J Virol *48:*451-459, 1983.
7. Beck E., Feil G. and Strohmaier K., EMBO J *2:*555-559, 1983.
8. Villanueva N., Davila M., Ortin J. and Domingo E., Gene *23:*185-194, 1983.
9. Weddell G.N., Yansura D.G., Dowbenko D.J., Hoaltin M.E., Grubman M.J., Moore D.M. and Kleid D.G., Proc Natl Acad Sci USA *82:*2618-2622, 1985.
10. Sobrino F., Palma E.L., Beck E., Davila M., de la Torre J., Negro P., Villanueva N., Ortin J. and Domingo E., Gene *50:*149-159, 1986.
11. Beck E. and Strohmaier K., J Virol *61:*1621-1629, 1987.
12. Ahl R., Tierärztl Umschau *42:*204-210, 1987.
13. Ahl R., Arch Gesamte Virusforsch *46:*302-314, 1974.
14. Chomczynski R. and Sacci N., Anal Biochem *162:*156-159, 1987.
15. Thomas P.S., Proc Natl Acad Sci USA *77:*5201-5205, 1980.
16. Sanger F., Nicklen S. and Coulson A.R., Proc Natl Acad Sci USA *74:*5463-5467, 1977.
17. Devereux J., Haeberli P. and Smithies O., Nucl Acids Res *12:*387-395, 1983.
18. Samuel A.R., Knowles N.J. and Kitching R.P., Epidem Inf. in press.
19. Chou P.Y. and Fasman, G.D., Adv Enzymol *47:*45-147, 1978.
20. Becker Y., in: Foot and Mouth Disease 17th Conference, Paris 1986, Office Internat. des Epizooties, France, 1987, pp. 152-168.
21. Hogle J.M., Chow M. and Filman, D.J., Science *229:*1358-1365, 1985.
22. Luo M., Vriend G., Kamer G., Minor I., Arnold E., Rossmann M.G., Boege U., Scraba D.G., Duke G.M. and Palmenberg A.C., Science *235:*182-193, 1987.