Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

J. Mol. Biol. (1987) 196, 963–966

Evidence for a Coiled-coil Structure in the Spike Proteins of Coronaviruses

The amino acid sequences of the spike proteins from three distantly related coronaviruses have been deduced from cDNA sequences. In the C-terminal half, an homology of about 30% was found, while there was no detectable sequence conservation in the N-terminal regions. Hydrophobic "heptad" repeat patterns indicated the presence of two α-helices with predicted lengths of 100 and 50 Å, respectively. It is suggested that, in the spike oligomer, these α-helices form a complex coiled-coil, resembling the supersecondary structures in two other elongated membrane proteins, the haemagglutinin of influenza virus and the variable surface glycoprotein of trypanosomes.

Coronaviruses are enveloped RNA viruses with a single-stranded genome of positive polarity (Siddell *et al.*, 1983; Sturman & Holmes, 1983). They cause considerable economical damage by infecting livestock and other domestic animals. Projecting from their surface are unusually large (~200 Å), petal-shaped spikes. These so-called peplomers mediate the binding of virions to the host cell receptor and are involved in membrane fusion. Further, they are considered the main targets of the protective immune response (Sturman & Holmes, 1983; Cavanagh *et al.*, 1986*a*).

Each peplomer consists of a dimer or possibly a trimer of the peplomer protein (Cavanagh, 1983), a glycoprotein of 180,000 to 210,000 M_r (Sturman & Holmes, 1983; Jacobs *et al.*, 1986; Boyle *et al.*, 1984). The peplomer proteins of mouse hepatitis virus (MHV)† and infectious bronchitis virus (IBV) are post-translationally cleaved into two subunits of similar size (Stern & Sefton, 1982; Sturman & Holmes, 1983). For MHV, cleavage appears essential for fusion activity (Sturman *et al.*, 1985). In contrast, the peplomer protein of feline infectious peritonitis virus (FIPV) is not cleaved and yet capable of inducing membrane fusion (Boyle *et al.*, 1984).

IBV, MHV and FIPV belong to three separate antigenic clusters in the coronavirus family (Siddell *et al.*, 1983). We have cloned and sequenced the peplomer genes of FIPV strain 79-1146 (de Groot *et al.*, unpublished results), IBV strain M41 (Niesters *et al.*, 1986) and MHV strain A59 (Luytjes *et al.*, unpublished results). From the nucleotide sequences, apoproteins were predicted of 1452, 1162 and 1324 residues, respectively. Peplomer sequences from related IBV (Binns *et al.*, 1985, 1986) and MHV (Schmidt *et al.*, 1987) strains have been published by other groups. The proteins are synthesized with an N-terminal signal sequence. A stretch of 20 to 25 hydrophobic residues, found near the C terminus, most probably serves as a transmembrane anchor.

Amino acid sequences have been aligned by the following procedure. Initial amino acid alignments were obtained by FASTP analysis (Lipman & Pearson, 1985). These alignments have been extended by reiterating FASTP with non-aligned parts as query sequence and by DIAGON comparison (Fig. 1(a); Staden, 1982). The results are summarized in Figure 2. Most conservation is observed in the C-terminal half of the proteins, with overall amino acid homologies of 35, 30 and 29% for IBV–FIPV, IBV–MHV and MHV–FIPV, respectively; about 50% of the amino acid substitutions may be considered conservative (Dayhoff *et al.*, 1983). In contrast, we did not find significant homology or matching cysteine residues in the N-terminal segments; amino acid residues that could be aligned by introducing numerous gaps were not conserved in closely related strains of IBV (Niesters, 1987) or MHV (Luytjes *et al.*, unpublished results). Furthermore, insertions or deletions in the N-terminal domains account largely for the differences in size of the peplomer apoproteins.

No experimental data are available on the structure of the peplomers. However, DIAGON plots revealed two repetitious regions in the C-terminal domains with a seven-residue periodicity (Fig. 1(b)). Closer analysis showed the presence of so-called "heptad repeats" (Cohen & Parry, 1986), i.e. a sequence periodicity (a-b-c-d-e-f-g) in which the residues in the a and d positions generally are hydrophobic (Fig. 3). Statistical tests of the predominant occurrence of hydrophobic residues in the a and d positions yielded confidence levels of at least 96%; in the long repetitive regions, the two parts with different heptad phasings have been tested separately. Heptad repeats are indicative of a coiled-coil structure in which the hydrophobic residues form the interface between interlocking α-helices (Cohen & Parry, 1986). In accordance with

† Abbreviations used: MHV, mouse hepatitis virus; IBV, infectious bronchitis virus; FIPV, feline infectious peritonitis virus; HA, haemagglutinin trimer.

0022-2836/87/160963-04 $03.00/0

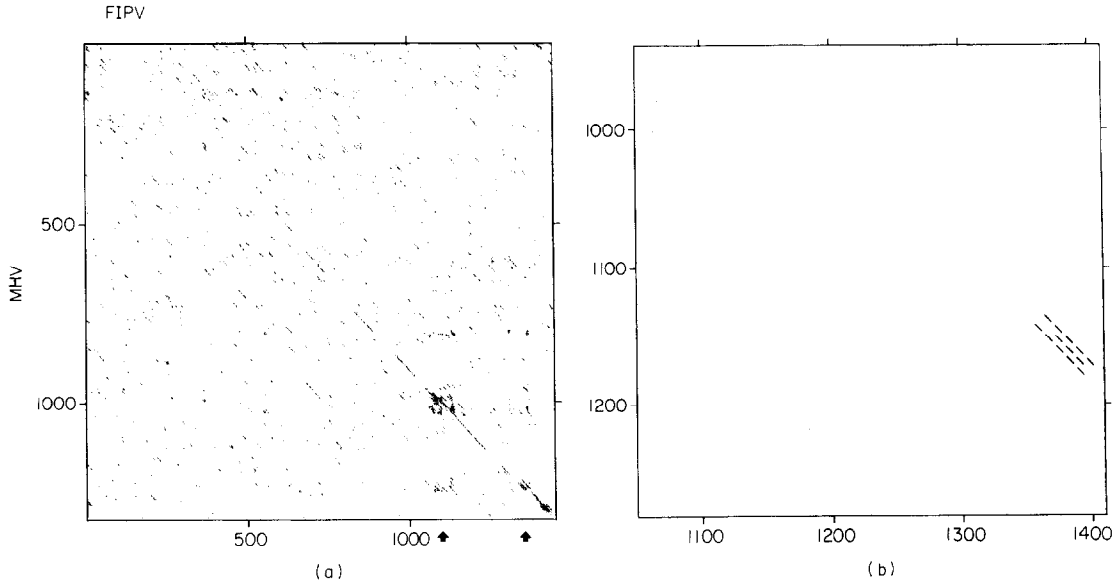

Figure 1. (a) Diagon plot (Staden, 1982) of the amino acid sequences of the peplomer protein from FIPV strain 79-1146 and MHV strain A59. Dots denote a proportional match of segments of 21 residues with a minimal score of 226. Arrows indicate 2 repetitive regions. (b) Enlargement to show the periodicity of the repetitive regions. The broken parallels indicate the spacing corresponding to 7 or 14 residues. Most distances between parallel dotted lines correspond to $7n$ (n = 1, 2, etc.) residues. A similar 7-residue periodicity was observed in proportional DIAGON plots of the HA of influenza virus types A and C (not shown).

the presumptive α-helical conformation, the repeats in the peplomer proteins are located in regions devoid of helix-breaking proline residues.

For the minor repeat near the transmembrane anchor (Figs 2 and 3) an α-helix of 50 (MHV and IBV) or 70 Å (FIPV) may be predicted. The major repeat indicates a helix of at least 100 (IBV and MHV) or 130 Å (FIPV), spanning more than half the peplomer. Note that in FIPV the minor and major repeats contain one insertion of 21 residues and two insertions of seven residues, respectively; thus three and two heptads are added, while the repeat pattern is conserved.

The presence of two heptad repeats suggests an intra-chain coiled-coil. However, this would leave about 50 Å of the predicted major helix unpaired. Therefore, it is assumed that in the oligomer the major helices are involved also in an inter-chain coiled-coil. Such a structure would resemble the complex coiled-coils found in the dimeric variable surface glycoproteins (VSG) of trypanosomes (Metcalf *et al.*, 1987) and the haemagglutinin trimer (HA) of influenza virus (Wilson *et al.*, 1981). In these proteins, bundles of four (VSG) or three (HA) α-helices with lengths of 90 and 76 Å, respectively, are surrounded by shorter helices; the interaction of the long helices stabilizes the oligomer.

The influenza virus HA and the coronavirus peplomer are functionally analogous, both carrying the receptor binding site and mediating membrane

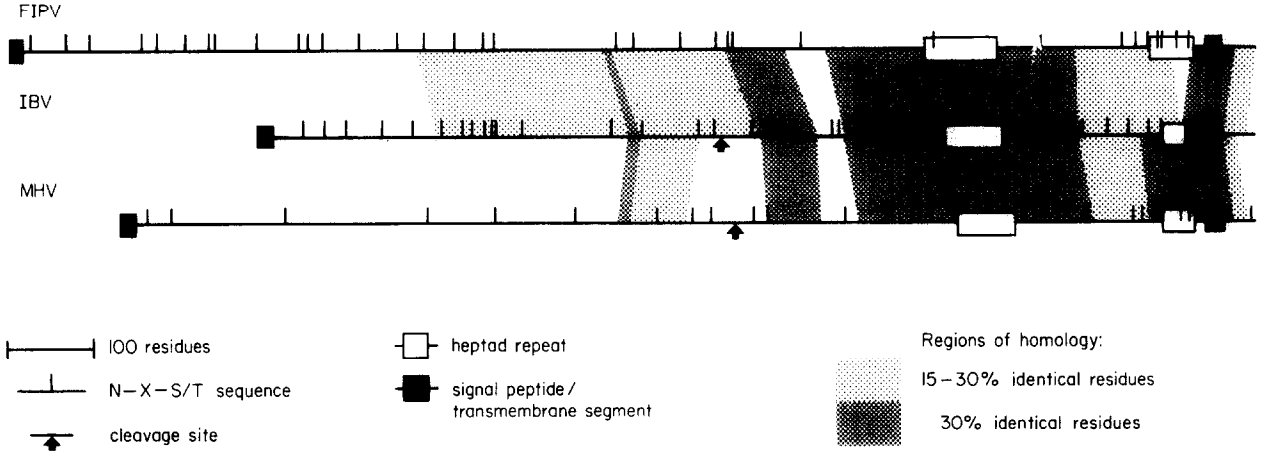

Figure 2. Schematic representation of the homology between the amino acid sequences of the peplomer proteins from FIPV strain 79-1146, IBV strain M41 and MHV strain A59. N-X-S/T (X ≠ Pro) sequences denote potential glycosylation sites. Cleavage sites in the peplomer proteins of IBV (Cavanagh *et al.*, 1986*b*) and MHV (Luytjes *et al.*, unpublished results) are indicated.

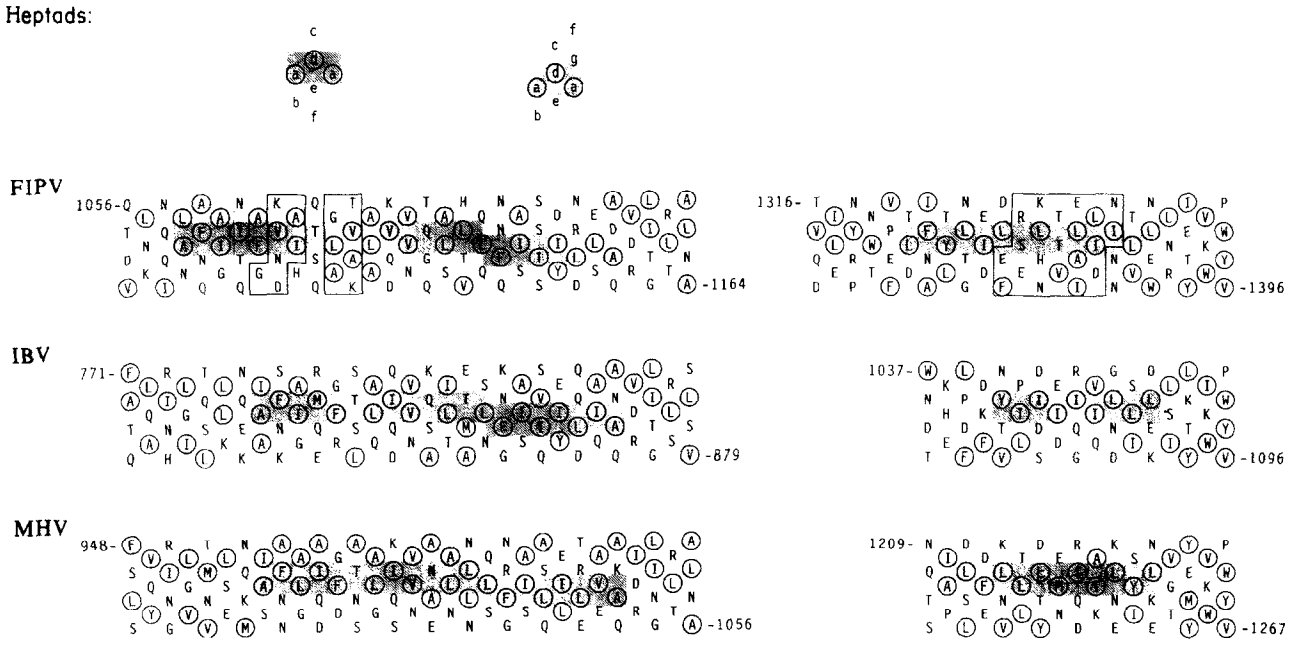

Figure 3. Heptad repeat patterns in the peplomer sequences. The sequences are listed vertically in alternating rows of 3 and 4 residues. Hydrophobic residues are encircled. Boxes indicate insertions in the FIPV sequence. The hatched regions indicate continuous patches of hydrophobic residues, which may interact with the corresponding regions of other α-helices.

fusion. We propose that these surface projections have converged to a similar supersecondary structure in order to position the receptor binding site at some distance from the membrane. Thus, the typical elongated shape of the coronavirus peplomer may be explained by a model (Fig. 4), in which a coiled-coil with a predicted length of 100 to 130 Å forms the connection between the globular part and the viral membrane. As in HA (Wilson *et al.*, 1981), the protein surface near the membrane may carry carbohydrate groups, attached to potential glycosylation sites in the region containing the minor heptad repeat (Fig. 2). The bulbous part of the peplomer protein probably contains the non-conserved N-terminal sequences (Fig. 2; Cavanagh, 1983; Binns *et al.*, 1985). Comparison of peplomer sequences of IBV (Niesters *et al.*, 1986; Binns *et al.*, 1986; Niesters, 1987) and MHV strains (Luytjes *et al.*, unpublished results) indicates that antigenic drift preferentially occurs in these regions. Hence, there is an obvious parallel with the sequence variation in the N-terminal subunit of the HA (Nakada *et al.*, 1984; Wiley *et al.*, 1981).

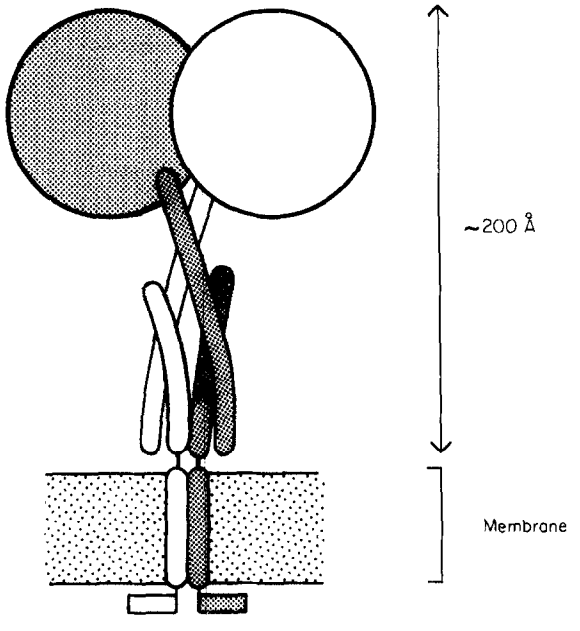

Figure 4. Tentative model of the coronavirus peplomer. The peplomer is represented as a dimer. The transmembrane α-helices and the α-helices in the coiled-coil structure are depicted as rounded cylinders.

This work was supported by a research grant from Duphar B.V., Weesp, The Netherlands.

R. J. de Groot[1]
W. Luytjes[1]
M. C. Horzinek[1]
B. A. M. van der Zeijst[2]
W. J. M. Spaan[1]
J. A. Lenstra[2]

[1] Institute of Virology and [2] Department of Bacteriology, Veterinary Faculty, University of Utrecht, Yalelaan 1, 3584 CL Utrecht, The Netherlands

Received 27 March 1987

References

Binns, M. M., Boursnell, M. E. G., Cavanagh, D., Pappin, D. J. C. & Brown, T. D. K. (1985). *J. Gen. Virol.* **66**, 719–726.

Binns, M. M., Boursnell, M. E. G., Tomley, F. M. & Brown, T. D. K. (1986). *J. Gen. Virol.* **67**, 2825–2831.

Boyle, J. F., Pedersen, N. C., Evermann, J. F., McKeirnan, A. J., Otts, R. L. & Black, J. W. (1984). *Advan. Exp. Med. Biol.* **173**, 133–147.

Cavanagh, D. (1983). *J. Gen. Virol.* **64**, 2577–2583.

Cavanagh, D., Davis, P. J., Darbyshire, J. H. & Peters, R. W. (1986*a*). *J. Gen. Virol.* **67**, 1435–1442.

Cavanagh, D., Davis, P. J., Pappin, D. J. C., Binns, M. M., Boursnell, M. E. G. & Brown, T. D. K. (1986*b*). *Virus Res.* **4**, 133–143.

Cohen, C. & Parry, D. A. D. (1986). *Trends Biochem. Sci.* **11**, 245–248.

Dayhoff, M. O., Barker, W. C. & Hunt, L. T. (1983). *Methods Enzymol.* **91**, 524–545.

Jacobs, L., Van der Zeijst, B. A. M. & Horzinek, M. C. (1986). *J. Virol.* **57**, 1010–1015.

Lipman, D. J. & Pearson, W. R. (1985). *Science*, **227**, 1435–1441.

Metcalf, P., Blum, M., Freymann, D., Turner, M. & Wiley, D. C. (1987). *Nature (London)*, **325**, 84–86.

Nakada, S., Creager, R. S., Krystal, M., Aaronson, R. P. & Palese, P. (1984). *J. Virol.* **50**, 118–124.

Niesters, H. G. M. (1987). Ph.D. thesis, University of Utrecht.

Niesters, H. G. M., Lenstra, J. A., Spaan, W. J. M., Zijderveld, A. J., Bleumink-Pluym, N. M. C., Hong, F., van Scharrenburg, G. J. M., Horzinek, M. C. & Van der Zeijst, B. A. M. (1986). *Virus Res.* **5**, 253–263.

Schmidt, I., Skinner, M. & Siddell, S. (1987). *J. Gen. Virol.* **68**, 47–56.

Siddell, S., Wege, W. & Ter Meulen, V. (1983). *J. Gen. Virol.* **64**, 761–776.

Staden, R. (1982). *Nucl. Acids Res.* **10**, 2951–2961.

Stern, D. F. & Sefton, B. M. (1982). *J. Virol.* **44**, 794–803.

Sturman, L. S. & Holmes, K. V. (1983). *Advan. Virus Res.* **28**, 35–112.

Sturman, L. S. Ricard, C. S. & Holmes, K. V. (1985). *J. Virol.* **56**, 904–911.

Wiley, D. C., Wilson, I. A. & Skehel, J. J. (1981). *Nature (London)*, **289**, 373–378.

Wilson, I. A., Skehel, J. J. & Wiley, D. C. (1981). *Nature (London)*, **289**, 366–373.

Edited by A. Klug