The EMBO Journal vol.6 no.4 pp.1099-1105, 1987

Is Sindbis a simple picornavirus with an envelope?

Stephen D.Fuller and Patrick Argos

Biological Structures and Biocomputing Programmes, European Molecular Biology Laboratory, D-6900 Heidelberg, FRG

Communicated by K.Simons

A three-dimensional image reconstruction was performed from cryo-electron micrographs of isolated Sindbis (SNV) nucleocapsids. The isolated capsid is a smooth but fenestrated T=3 structure. Comparison with the nucleocapsid seen within the whole virion indicated that the structure resembles the swollen forms which some non-enveloped viruses adopt after removal of divalent cations. A sensitive comparison method was used to align SNV capsid protein sequences with those of picornavirus vp3 capsid proteins whose high resolution structures display an eight-stranded β-barrel fold found in many icosahedral viruses. The alignment predicted a similar folding for the Sindbis protein which juxtaposes several sets of residues known to be essential for its serine proteolytic activity. These results suggest that the capsid proteins of the enveloped alphaviruses and the non-enveloped picornaviruses may have arisen through divergent evolution from a simple, vp3-like ancestor.

Key words: cryo-electron microscopy/alphavirus/sequence homology/divergent evolution

Introduction

Alphaviruses are among the simplest and best characterized enveloped animal viruses. Sindbis virus (SNV) contains 240 copies of each of two transmembrane glycoproteins (E1 and E2) which span a lipid bilayer surrounding an icosahedral nucleocapsid. The nucleocapsid comprises a single positive strand of RNA and copies of a single protein (C). Negative staining and cryo-electron microscopy have demonstrated that the glycoproteins are arranged as 80 spikes on a T=4 icosahedral axis with each spike made of three copies each of E1 and E2 (Strauss and Strauss, 1977; Simons and Warren, 1984; Kääriäinen and Söderlund, 1978; Fuller, 1987). The closely related Semliki Forest virus (SFV) appears to have an almost identical structure (Vogel *et al.*, 1986) except that each spike also contains three copies of a third non-membrane-spanning glycoprotein E3.

The triangulation number of the alphavirus nucleocapsid has been much more difficult to determine. Some workers have argued for T=4 since the envelope is T=4 and the envelope glycoproteins are known to interact with the nucleocapsid. Electron microscopy of the isolated nucleocapsid did not provide a clear answer (Murphy, 1980). Various workers have interpreted images as evidence of T=3, T=4 and T=9 (von Bonsdorff, 1973; Brown *et al.*, 1972; Brown and Gliedmann, 1973; Enzmann and Wieland, 1979). Recently, Fuller (1987) presented a reconstruction of the whole Sindbis virion from cryo-electron micrographs which showed the T=4 arrangement of the trimeric spikes and an internal structure with T=3 symmetry. The smooth but fenestrated nucleocapsid structure seen in the reconstruction contrasts with that observed for many non-enveloped T=3 viruses. In this paper, we discuss two lines of evidence concerning the relationship of this enveloped virus capsid and the non-enveloped virus capsids. We present a reconstruction of the isolated Sindbis virus nucleocapsid confirming its T=3 symmetry and indicating regions of flexibility for the isolated structure. These regions can be related to the previously characterized structure of 'expanded' TBSV (Robinson and Harrison, 1982). We also present an alignment of the alphavirus capsid protein sequences with those of picornavirus vp3 proteins, suggesting a similar main chain fold.

Results

Cryo-electron microscopy and image reconstruction

Sindbis nucleocapsids, isolated by TX-114 phase partitioning (Bordier, 1981), were examined by cryo-electron microscopy (Figure 1). The isolated nucleocapsid appears to be a relatively smooth structure which is regular and well preserved in these preparations. Preliminary processing of particles (Crowther, 1971; Fuller, 1987) showed that some were ordered to a resolution of 45Å. Despite their regular appearance, they are not as well ordered as the intact virion which shows information to be at least 30Å resolution. Attempts to improve the preservation of the particle by the inclusion of nuclease inhibitors or divalent ions did not result in significantly better preparations.

Pairs of images were taken with 3 μm and 8 μm underfocus to sample the transform to 45Å resolution (Erickson and Klug, 1971). Over 200 focal pairs of images were screened and two sets of ten were selected and combined to give two independent reconstructions using the common lines procedure of Crowther (1971) as modified by Fuller (1987) for low contrast cryo-electron microscopic images. The average icosahedral (self) common lines phase residual for the combined sets was 40° while the interparticle common lines residual was 45° at a resolution of 45Å. The two reconstructions showed similar features.

Figure 2A shows a surface representation of the nucleocapsid reconstruction viewed down the 3-fold axis. The feature of highest contrast is a fenestrated shell between radii 170 and 210Å. Beneath this shell is a low density region which surrounds a non-icosahedral higher density corresponding to the RNA. The volume of the shell is $\sim 4.1 \times 10^6$ Å^3 corresponding to 3.1×10^6 daltons (Matthews, 1968) or 174 copies of the C-terminal domain (see below) of the 30-kd capsid protein. Rotational correlation analysis of the projected shell structures (Figure 2D) was used to test the symmetry of the isolated nucleocapsid. The positions of the peaks corresponding to 6-fold symmetry are indicated by the arrows. A T=4 structure will have local 6-fold symmetry around the icosahedral 2-fold axis while a T=3 structure will display local 6-fold symmetry around the icosahedral 3-fold axis (Casper and Klug, 1962). The map used for the projections was only constrained to have D_5 (522) symmetry so that all peaks are independent. The quality of the 3-fold and 2-fold symmetry revealed by the correlation functions indicates the preservation

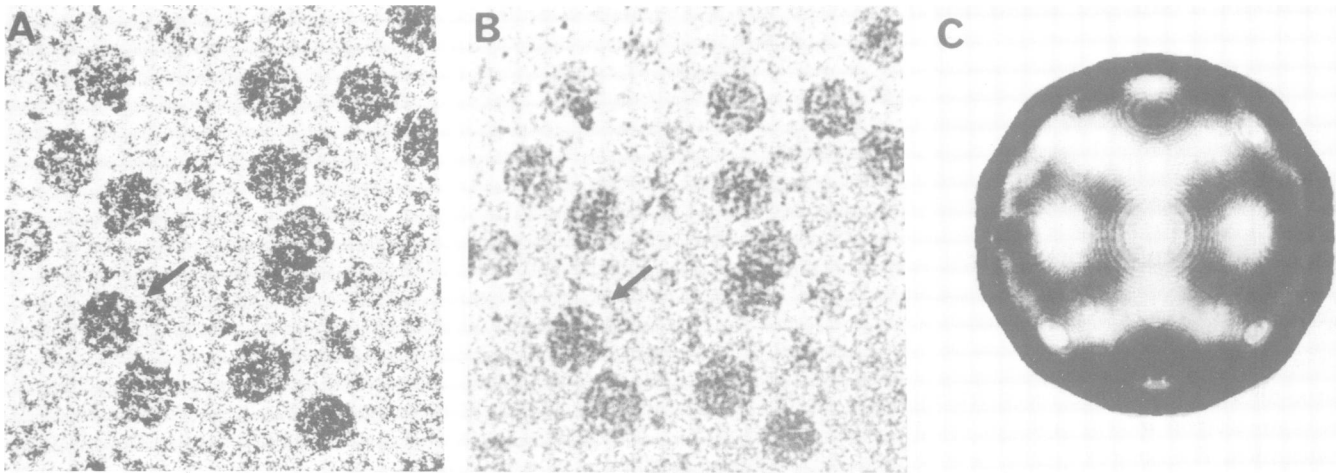

Fig. 1. Cryo-electron microscopy of isolated Sindbis nucleocapsids. **(A)** and **(B)** show the same field of nucleocapsids at two different defocus levels to sample the data to a resolution of 45Å (Erickson and Klug, 1971). In **A**, taken at 6 μm underfocus, the coarser features of the structure are seen while the higher resolution information is provided by images **(B)** which were taken at 3 μm underfocus. A view of a particle near the 2-fold axis is arrowed in each image. It is very similar to the projection of the final image reconstruction along its 2-fold axis **(C)**. The bar represents 1000Å in **(A)** and **(B)** and 100Å in **(C)**.

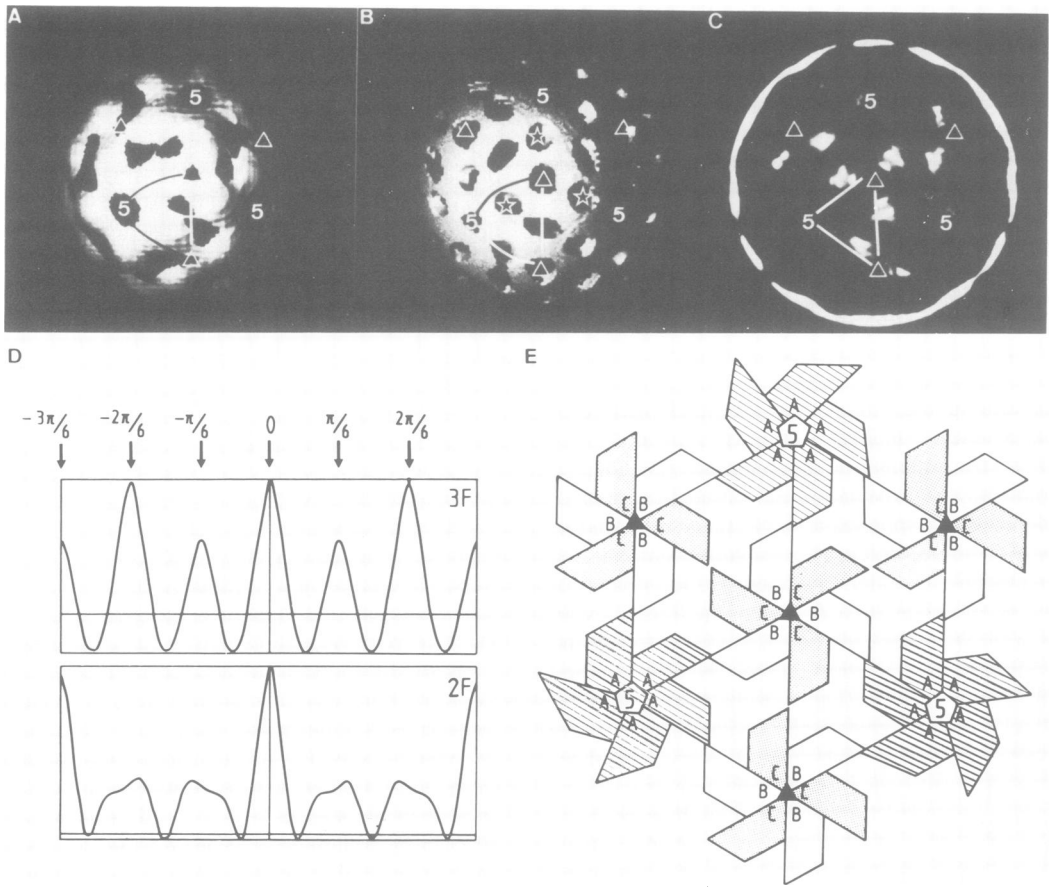

Fig. 2. Image reconstruction of the Sindbis nucleocapsid. **(A)**: surface representation of reconstruction of the isolated Sindbis virus nucleocapsid, with the positions of the 5-fold and 3-fold axes as well as the icosahedral asymmetric unit bounded by a 5-fold and two 3-fold axes. The surface of the shell corresponding to the Sindbis virus nucleocapsid as seen within the three dimensional reconstruction of the intact virion (Fuller, 1987) is shown in **(B)** with the symmetry axes and the asymmetric unit marked. Stars indicate the positions of the interaction of the capsid with the cytoplasmic domains of the spike protein trimers. A difference map between the projected density of the isolated nucleocapsid and the projected density of the nucleocapsid shell seen in the intact virus is shown in **(C)**. The region of greatest difference lies on either side of the line connecting the 3-fold axes. Rotational correlation functions for the projections down the 3-fold (3F) and 2-fold (2F) axes of the top half of the non-averaged isolated nucleocapsid density are shown in **(D)**. The structural changes observed during the expansion of the T=3 TBSV by Robinson and Harrison (1982) are shown in schematic form in **(E)**. The ABC notation refers to the three conformations of the capsid protein seen in the TBSV structure of Harrison *et al.* (1978) and follows their convention. **(A)**, **(B)**, **(C)** and **(E)** are all presented in the same orientation.

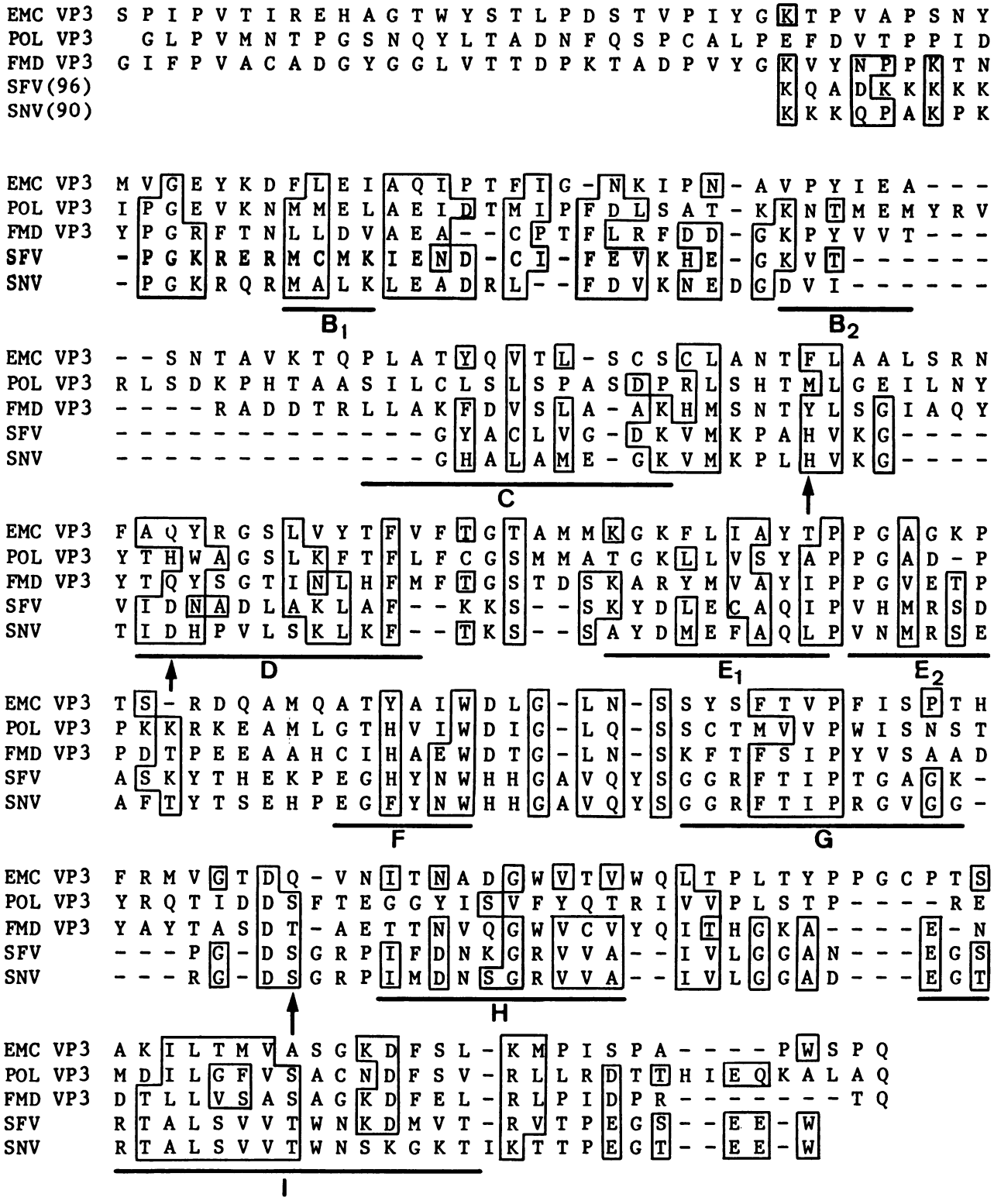

Fig. 3. Picornavirus-alphavirus capsid protein alignment. Alignment of the capsid vp3 protein sequences from polio virus (POL) (Kitamura *et al.*, 1981), foot and mouth disease virus (FMD) (Carrol *et al.*, 1984) and encephalomyocarditis virus (EMC) (Palmenberg *et al.*, 1984) with the nucleocapsid sequences of Sindbis virus (SNV) (Strauss *et al.*, 1984) and Semliki Forest virus (SFV) (Garroff *et al.*, 1980a,b) was performed using the procedure of Argos (1987). The vp3 sequences and the SFV and SNV sequences were aligned visually while the interset alignments followed the search results shown in Figure 4A. Residues are boxed where at least one aligned amino acid contributed by SFV or SNV is conserved relative to at least one amino acid from each of the vp3 sequences according to the following conservation scheme: (ST), (PG), (KR), (EDQN), (FYHW) and (AIVLMC). Arrows mark the conserved His, Asp and Ser in the SFV and SNV sequences, suggested to be essential in the protease activity of the nucleocapsid proteins (Hahn *et al.*, 1985). Approximate locations of the β-strands, expected for the FMD vp3 protein as derived from homology with the known rhinovirus structure (Rossmann *et al.*, 1985), are indicated by underlining; the alphabetic nomenclature for β-strands is borrowed from Rossmann *et al.* (1985) except where βB is broken as βB_1 and βB_2 due to the 'puff' sequences in the vp3 proteins.

of icosahedral symmetry in the isolated capsid particle. The average value for 6-fold peaks which are independent of the icosahedral symmetry is 0.54 in 3F and 0.29 in 2F. Hence, the isolated capsid maintains T=3 icosahedral symmetry.

Figure 2B shows the clear T=3 symmetry of the nucleocapsid structure observed within the intact virion (Fuller, 1987). The positions occupied by interaction with the spike protein trimers are indicated by stars. The virion and isolated nucleocapsids are very similar, particularly near the 3-fold and 5-fold axes and differ most near the 2-fold axes (Figure 2C). This difference may reflect increased protein flexibility in this region of isolated nucleocapsids and is discussed below.

Several features of the T=3 Sindbis nucleocapsid suggest that it resembles the swollen structure assumed by other T=3 viruses upon increase of pH and removal of divalent cations. A correspondence between the Sindbis nucleocapsid reconstruction and the positions of the capsid proteins found by Robinson and Harrison (1982) in their 8Å structure of the expanded tomato bushy stunt virus (TBSV) can be shown by superimposing like symmetry axes in the two structures. Figure 2 displays the structures of the isolated Sindbis nucleocapsid (Figure 2A), the nucleocapsid seen within the virion (Figure 2B), the difference map showing the position of greatest change between the intact and isolated structure (Figure 2C) and a schematic diagram of the same region of the expanded TBSV structure (Figure 2E) in the same orientation. The region of greatest difference corresponds to that near the C subunits, which are found to be the most disordered in the TBSV structure.

Sequence alignment of alphavirus and picornavirus capsid proteins

The tertiary structures of the capsid proteins for several non-enveloped, positive-strand RNA icosahedral viruses from plants and mammals show a common eight-stranded antiparallel β-strand architecture (cf. Rossmann *et al.*, 1985; Hogle *et al.*, 1985). These include the T=3 plant viruses, southern bean mosaic virus (SBMV) (Abad-Zapatero *et al.*, 1980) and TBSV (Harrison *et al.*, 1978) the T=1 satellite tobacco necrosis virus (STNV) (Liljas *et al.*, 1982), and the human picornaviruses, rhinovirus (Rossmann *et al.*, 1985) and the homologous poliovirus (POL) (Hogle *et al.*, 1985) which build a pseudo T=3 shell from three different coat proteins (vp1, vp2, vp3) with similar folds. A sensitive comparison program (Argos, 1987, and Materials and methods) was used to determine whether the structural homology detected by image reconstruction was reflected at the level of sequence homology.

The SNV and related SFV capsid sequences were compared with those mentioned above as well as with the related picornavirus, foot-and-mouth disease (FMD) and encephalomyocarditis (EMC). The strongest relationship was found between FMD vp3 and SFV coat proteins. Figure 3 shows this alignment while Figure 4A displays the search matrix showing a clear diagonal

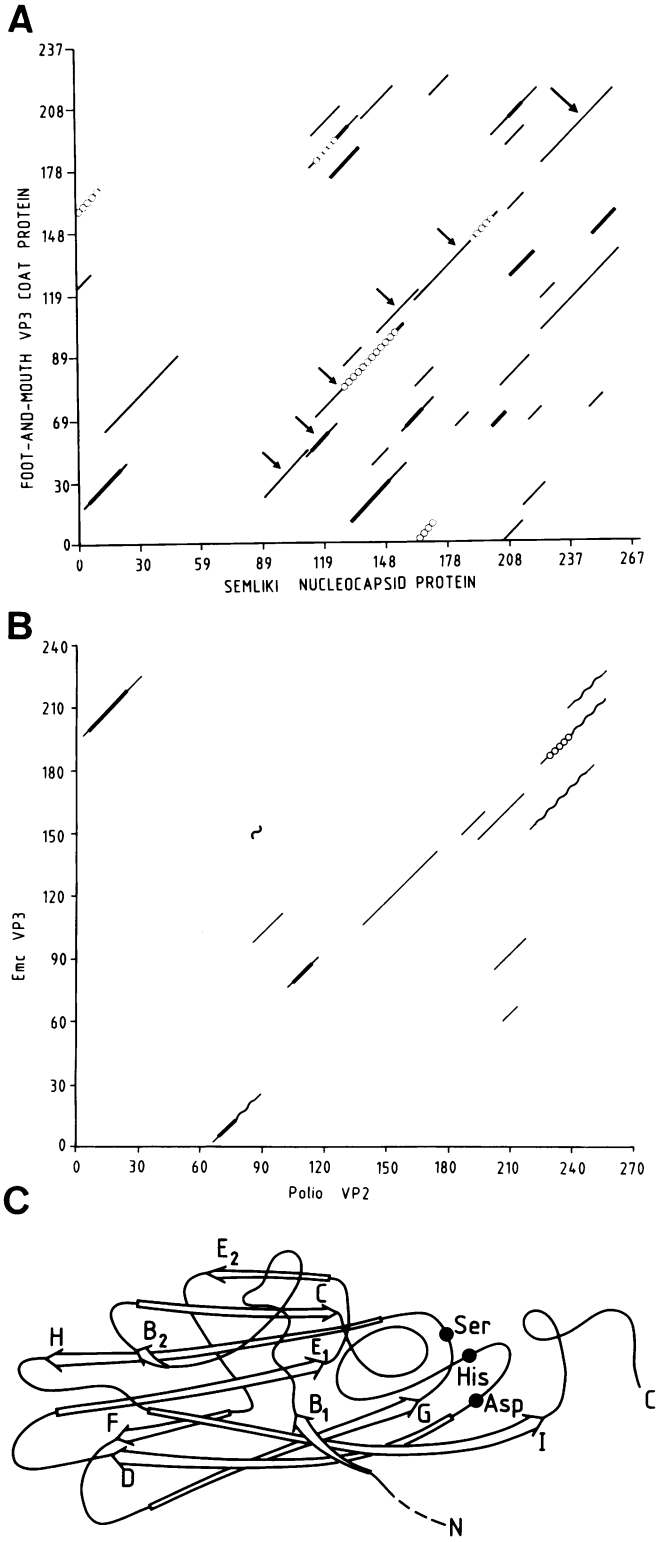

Fig. 4. **(A)** Homology search matrix between the SFV nucleocapsid protein sequence and the FMD vp3 coat protein sequence. The search window lengths ranged from 7−35 in steps of 2. Symbols are used to indicate fractional standard deviation ranges (σ) of the peak heights (S): thin lines ($3.6\sigma <= S < 4.0\sigma$), thick lines ($4.0\sigma <= S < 4.2\sigma$) and circles ($4.2\sigma <= S \leq 5.6\sigma$). The matrix shows the maximum standard deviation values accummulated from searches using all the probe lengths with the symbols plotted over the entire window length. The path, collinear with the main diagonal and used to effect the alignment of Figure 3, is indicated by arrows. The remaining aligned regions not indicated by the search matrix were determined visually. **(B)** The search matrix for the comparison of EMC vp3 and POL vp2 sequences. The symbols used to indicate the peak heights (S) in fractional standard deviation ranges are: thin lines ($4.0\sigma <= S < 4.2\sigma$), thick lines ($4.2\sigma <= S < 4.3\sigma$), wavy lines ($4.2\sigma <= S < 4.8\sigma$) and circles ($4.8\sigma <= S \leq 5.5\sigma$). **(C)** is based on the β-barrel mainchain trace of the rhinovirus vp3 sequence by Rossmann *et al.* (1985). the trace has been somewhat altered in the loop region to reflect the deletions predicted for the SFV nucleocapsid protein. The fold indicated for SFV and SNV is intended only as a rough approximation. The predicted spatial proximity of the His, Asp and Ser involved in the autolytic activity of Sindbis virus are indicated; their positions were derived from the alignments of Figure 3 and the predicted structural relationships between rhinovirus vp3 and the alphavirus capsid proteins.

path of high peaks. Twenty-four percent of the alignment positions, excluding insertions and deletions, show identical amino acids in the SFV nucleocapsid protein and FMD vp3, a value close to the mean identity level for the vp3 sequences of FMD, POL and EMC (29%). There are 175 positions where at least one aligned amino acid is contributed to by SFV or SNV and at least one from the vp3 sequences. Fifty-seven percent of these positions show conservation between the two sets of sequences (Figure 3). The mean correlation for SFV and FMD vp3 over five residue characteristics is 0.32, while that for the three vp3 relationships yields a mean of 0.48.

The significance of the alphavirus-picornavirus homology was tested in two ways. First, the same alignment technique was used to search for the known structural homology between picornaviral vp2 and vp3 proteins (Rossmann *et al.*, 1985; Hogle *et al.*, 1985). The POL-EMC search matrix (Figure 4B) suggests a relationship between vp2 and vp3. A second alignment, taken from a control group of unrelated proteins described in Materials and methods, yielded a mean trial correlation of 0.000 with standard deviation of 0.054 and maximum correlation of 0.230. The alphavirus-picornavirus homology is statistically significant (6.0 above the control mean and 1.7 above the control maximum) and more apparent than the known vp2-vp3 homology (Figure 4). Other correlations near the 4.0 level included FMD vp3 with SNV (0.28) and EMC vp3 with SFV (0.23) and SNV (0.22).

The alignment with known three-dimensional structures allows a prediction of the folding of the homologous parts of the Sindbis virus capsid protein. One striking feature of the predicted folding is that the majority of the carboxy-terminal two-thirds of the protein is involved in the β-barrel fold. Such a folding leaves little of the protein to contribute to the puffs and projecting domains responsible for the surface relief in the previously characterized T=3 virus structures. This is consistent with the smooth appearance of the isolated Sindbis nucleocapsid surface (Figure 2).

Discussion

The structure of the Sindbis nucleocapsid seen within the whole virion (Fuller, 1987) or in isolation is a smooth but fenestrated shell. This differs from the closed and relatively rough-surfaced shells formed by the capsid proteins of simple non-enveloped viruses (cf. Harrison *et al.*, 1978). Part of this difference can be explained by the hypothesis that the Sindbis nucleocapsid corresponds to the expanded forms assumed by some simple spherical viruses on removal of divalent ions. Several pieces of evidence support this correspondence. The Sindbis nucleocapsid diameter is 30% larger than that of the compact form of SBMV (Abad-Zapatero *et al.*, 1980) whose capsid protein has a similar mol. wt. The Sindbis genome is degraded by RNase treatment of the isolated nucleocapsid (Söderlund *et al.*, 1972; Söderlund, 1979). The compact capsids of non-enveloped viruses which protect their RNAs from nuclease usually lose this ability upon expansion. Furthermore, exposure of the isolated Sindbis nucleocapsid to low pH results in a 60Å contraction (Söderlund *et al.*, 1975; Söderlund, 1979).

Often the expanded form of the capsid is less well ordered than the compact form so that our understanding of the structural interaction which maintains an expanded shell remains limited. Robinson and Harrison (1982) have published an 8Å resolution map of the expanded form of TBSV. Their map shows that expansion is accomplished by separation of the protein subunits in the ABC trimer made from the different conformations of the capsid protein (Figure 2E). Contacts around the 3-fold (BC) and those around the 5-fold (AA) are maintained while the AB, AC and CC interactions are mostly lost upon expansion. Some density seen between the A and B subunits is suggested to be a result of an extension of an N terminal arm from each. The AC and CC contacts appear to be completely lost (Figure 2E). The most disordered regions in the expanded structure are the P domains of the C subunits which display a temperature factor of 600Å^2 in contrast to the 80Å^2 seen for A and B. Though the individual subunits are not resolved in the Sindbis reconstruction, superposition of the symmetry axes reveals that the position of greatest change (Figure 2D) correlates with that of the C subunit in expanded TBSV. The positions of the holes in the Sindbis nucleocapsid, while consistent with an expanded structure, do not align precisely with those seen in TBSV. A 10−15° rotation of the subunit clusters would align these positions. A rotation of this magnitude has been observed in the T=1 satellite tobacco necrosis virus (Liljas *et al.*, 1982) and in the T=3 alfalfa mosaic virus (Rossmann *et al.*, 1985) relative to the TBSV positions. There are six holes around the 3-fold axes. The three closest to the 5-fold axes coincide with the positions of nucleocapsid interaction with the cytoplasmic domains of the viral spike protein trimers (Figure 2B). This interaction between trimeric spikes and trimers of capsid units appears to be a close one in the structure of the intact Sindbis virion. The interaction with the spike on the 3-fold axis appears looser (Fuller, 1987). The increased flexibility of the isolated Sindbis nucleocapsid can be explained by an expanded structure stabilized in the intact virion by interaction with the spike proteins. The fact that these interactions occur at or near the interfaces containing the divalent ions which stabilize the TBSV and SBMV structures supports this interpretation. An alternative interpretation is that the flexibility in the isolated nucleocapsid results from some digestion of the viral RNA during isolation of the capsids. The shape of the isolated nucleocapsid is very sensitive to RNase treatment (Söderlund *et al.*, 1975, 1979). Nonetheless, our hypothesis concerning the expanded nature of the nucleocapsid and the role of spike interactions in maintaining its integrity would remain valid, even if digestion contributes to the disorder. The significance of the complementary interactions between the T=3 nucleocapsid and the T=4 envelope is discussed elsewhere (Fuller, 1987).

The comparison of the alphavirus capsid protein sequences and those from the picornavirus provides a second line of evidence relating these T=3 and T=3-like structures. We found that the carboxy-terminal two-thirds of the alphavirus proteins aligned with the β-barrel portion of the picornavirus vp3 proteins. This alignment predicts a fold for the protein which is consistent with the relatively smooth structure observed for the alphavirus nucleocapsid. Further similarities between alpha- and picornavirus include sequence homology of their RNA-dependent RNA polymerases and their position at the 5′ end of the viral genome (Kamer and Argos, 1984, and references therein).

Biochemical and genetic data can be used to test this predicted folding. The alphavirus capsid protein displays serine protease activity during its autolytic release from the initial polyprotein (reviewed in Hahn *et al.*, 1985). Three eight-residue stretches are strongly conserved in alphaviruses and are also homologous to similar spans in several serine proteases which contain histidine, asparagine and serine residues essential for activity. Mutations in these regions destroy the proteolytic activity of the capsid protein (Hahn *et al.*, 1985; Melançon and Garoff, 1987). These three residues coincide in a natural cleft on one side of the β-barrel exposed in the known picornaviral capsid

structure (Rossmann *et al.*, 1985; Hogle *et al.*, 1985) and exposed in the expanded nucleocapsid structure. Modifications at the other end of the β-barrel to accommodate the cleavage active geometry could disrupt the tight interactions around the local 6-fold axis. This localization of residues essential for activity supports the folding pattern of Figure 4C since the odds for the random occurrence of such an event are $< 1{:}800$. This result should not be construed as evidence that the capsid protein must adapt the same folding as a serine protease; our results suggest that precisely the opposite is true and that the protease active site is accommodated on the picornaviral fold. Examples abound of proteins which adapt a different main chain fold but a similar active site geometry for similar function (Garavito *et al.*, 1977; Argos *et al.*, 1978; Kraut *et al.*, 1971).

Given the structural similarities of the alphavirus capsid and that of simple non-enveloped RNA viruses, we propose that both arose by divergent evolution from a simple, non-enveloped T=3 prototype. The basic structure was elaborated, probably to escape immune surveillance and allow receptor recognition. For the picornaviruses this involved three separate but related genes for the three capsid proteins allowing more effective modulation of surface features (Rossmann *et al.*, 1985; Hogle *et al.*, 1985). In the alphavirus case, interaction sites were developed in the capsid for the cytoplasmic domains of membrane glycoproteins allowing for its envelopment with a lipid bilayer. This was accompanied by the acquisition of genes for the glycoproteins with their capacities for surface recognition, modulation of antigenicity and membrane fusion.

Materials and methods

Virus and preparation of capsids

Virus was grown and purified as described previously (Fuller, 1987). Nucleocapsids were isolated by treatment of a 1 mg/ml solution of virus with 5% (w/v) Triton X-114 in phosphate-buffered saline (pH 7.6) containing 1 mM Ca^{2+} and 1 mM Mg^{2+}, followed by phase partitioning at 37°C and low speed centrifugation. The nucleocapsids remained in the detergent-poor supernatant phase while the envelope proteins were concentrated in the detergent-rich lower phase.

Cryo-electron microscopy

Cryo-electron microscopy was performed as described previously (Dubochet *et al.*, 1982; Fuller, 1987) except that a carbon support film was used to overcome the influence of residual detergent on the thinning of the water film during preparation. The behavior of the particles under the electron beam as well as the change upon warming above the −160°C required to maintain the vitrified state showed that the particles were indeed embedded in a water layer. Micrographs were taken on a Phillips EM 400 microscope equipped with a Phillips cold stage (PW 6591/7000) and an improved blade anticontaminator (Homo *et al.*, 1984) at a nominal magnification of 33 000 at 80 kV using a 25-μm objective aperature and a fluence of 5−10 electrons/Å^2. When a focal series was taken, the closer to focus image was always taken first.

Image reconstruction and image presentation

Images were screened by eye and optical diffractometry for intactness of particles, proper underfocus, lack of astigmatism and drift. Data from each image was only used within the first zero of the contrast transfer function so that no phase changes would be introduced by the level of defocus. Images were scanned with an Optronics P1000 drum scanner using a step size of 25 μm which is equivalent to 8.04 Å. Initial processing of images, including the selection and centering of particles and the subtraction of the background level, was performed with SEMPER (Image Techniques of Cambridge Limited, Cambridge, UK). Subsequent processing steps used the suite of programs for icosahedral image reconstruction originally written by Dr R.A.Crowther (Crowther, 1971) implemented in FORTRAN 77 by Dr Timothy Baker (Department of Biology, Purdue University, West Lafayette, IN, USA) and modified as described in Fuller (1987). A surface representation program (originally obtained from Dr G.Vigers, Medical Research Council Laboratory of Molecular Biology, Cambridge, UK) was used to generate the representations in Figures 2A and 2B.

Alignment of sequences

Alignment of the capsid vp3 protein sequences from POL (Kitamura *et al.*, 1981) FMD (Carrol *et al.*, 1984), and EMC (Palmenberg *et al.*, 1984) with the nucleocapsid sequences of SNV (Strauss *et al.*, 1984) and SFV (Garroff *et al.*, 1980a,b) was performed using the sensitive alignment procedure of Argos (1987). The protein sequences were correlated by comparing every possible span of length L residues in one protein with all such spans in the second protein. Two scoring procedures were used (Argos, 1985): the Dayhoff relatedness odds matrix (Barker *et al.*, 1978) and the mean correlation coefficient over five residue physical characteristics including β-strand and reverse turn conformational preferences (Palau *et al.*, 1982), residue bulk (Jones, 1975), amino acid refractivity index (Jones, 1975), and the 'surrounding' hydrophobicity (Manavalan and Ponnuswamy, 1978). The five parameters used gave the highest correlations amongst 11 000 residue pairs aligned through mainchain atom spatial superposition of known tertiary protein structures (cf. Rossmann and Argos, 1975) and were themselves intercorrelated at less than 0.7. The search matrix for a given probe length L was constructed by scaling and averaging the scores from the two techniques. The matrix values were stored as a number of standard deviations (σ) above the matrix mean. The final search matrix was produced by calculating standard deviation matrices for each of the different window lengths and storing the scores at each matrix position corresponding to the entire probe length; finally, only the maximum value was saved at each search position from the several probe lengths.

A control experiment was performed to test the significance of these relationships. Sixty-three sequences of proteins whose tertiary structures are known to be unrelated were successively duplicated to achieve a length of 1000 residues. Segments with length 170 residues (which is close to the number of aligned residues between SFV and FMD vp3) were randomly selected from each of the 63 sequences, aligned as such for all possible pairs, and mean correlations for the five characteristics calculated. This procedure was repeated 30 times for a total of nearly 55 000 individual trials. The mean trial correlation and standard deviation were 0.000 and 0.054 respectively with the maximum achievable correlation at 0.230. Sequences from actual proteins were used rather than random shuffling of the sequences as biases built over evolutionary time are accounted for.

Acknowledgements

The authors wish to thank Dr Kai Simons (EMBL), Dr Jacques Dubochet (EMBL) and Dr Stephen Harrison (Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA, USA) for encouragement and useful discussions during the course of this work. We also thank Christine Barber for her help with the word processing of the manuscript.

References

Abad-Zapatero,C., Abdel-Meguid,S.S., Johnson,J.E., Leslie,A.G.W., Rayment,I., Rossmann,M.G., Suck,D. and Tsukihara,T. (1980) *Nature*, **286**, 33−39.

Argos,P. (1985) *EMBO J.*, **4**, 1351−1355.

Argos,P. (1987) *J. Mol. Biol.*, **193**, 385−396.

Argos,P., Garavito,R.M., Eventhoff,W. and Rossmann,M.G. (1978) *J. Mol. Biol.*, **126**, 141−158.

Barker,W.C., Ketchlam,L.K. and Dayhoff,M.O. (1978) In Dayhoff,M.O. (ed.), *Atlas of Protein Sequence and Structure.* National Biomedical Research Foundation, Washington, Vol. 5, Suppl. 3, pp. 359−362.

Bordier,C. (1981) *J. Biol. Chem.*, **256**, 1604−1607.

Brown,D.T. and Gliedmann,J.B. (1973) *J. Virol.*, **12**, 1534−1539.

Brown,D.T., Waite,M.F. and Pfefferkorn,E.R. (1972) *J. Virol.*, **10**, 524−536.

Carroll,A.R., Rowlands,D.J. and Clarke,B.S. (1984) *Nucleic Acids Res.*, **12**, 2969−2985.

Casper,D.L.D. and Klug,A. (1962) *Cold Spring Harbor Symp. Quant. Biol.*, **24**, 1−24.

Crowther,R.A. (1971) *Phil. Trans. Roy. Soc. Lond.*, **B261**, 221−230.

Dubochet,J., Lepault,J., Freeman,R., Berriman,J.A. and Homo,J.-Cl. (1982) *J. Microsc.*, **128**, 219−237.

Enzmann,P.J. and Wieland,F. (1979) *Virology*, **95**, 501−510.

Erickson,H.P. and Klug,A. (1971) *Philos. Trans. Roy. Soc. Lond.*, **A317**, 319−340.

Fuller,S.D. (1987) *Cell*, **48**, in press.

Garavito,R.M., Rossmann,M.G., Argos,P. and Eventhoff,W. (1977) *Biochemistry*, **16**, 5065−5071.

Garoff,H., Frischauf,A.-M., Simons,K., Lehrach,H. and Delius,H. (1980a) *Proc. Natl. Acad. Sci. USA*, **77**, 6373−6380.

Garoff,M., Frischauf,A.-M., Simons,K., Lehrach,H. and Delius,H. (1980b) *Nature*, **288**, 236−241.

Hahn,C.S., Strauss,E.G. and Strauss,J.H. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 4648−4652.

Harrison,S.C., Olson,A.J., Schutt,C.E., Winkler,F.K. and Bricogne,G. (1978) *Nature*, **276**, 368−373.

Hogle,J.M., Chow,M. and Filman,D.J. (1985) *Science*, **220**, 1358−1365.

Jones,K.K. (1975) *J. Theor. Biol.*, **50**, 167−183.

Jones,T.A. and Strandberg,B. (1983) *J. Mol. Biol.*, **165**, 711–736.
Kamer,G. and Argos,P. (1984) *Nucleic Acids Res.*, **12**, 7269–7282.
Kitamura,N., Semler,B.L., Rothberg,P.G., Larson,G.R., Adler,C.J., Dorner,A.J., Emini,E.A., Hanecak,R., Lee,J.J., van der Werf,S., Anderson,C.W. and Wimmer,E. (1981) *Nature*, **291**, 541–553.
Kraut,J., Robertus,J.D., Birktoft,J.J., Alden,R.A., Wilcox,P.E. and Posers,J.C. (1971) *Cold Spring Harbor Symp. Quant. Biol.*, **36**, 117–123.
Kääriäinen,L. and Söderlund,H. (1978) *Curr. Top. Microbiol. Immunol.*, **82**, 16–57.
Liljas,L., Unge,T., Jones,T.A., Fidborg,G., Lavgren,S., Skoglund,U. and Strandberg,B. (1982) *Nature*, **159**, 93–108.
Matthews,B.W. (1968) *J. Mol. Biol.*, **33**, 491–497.
Melancon,P. and Garoff,H. (1987) *J. Virol.*, in press.
Murphy,F.A. (1980) In Schlesinger,R.W. (ed.), *The Togaviruses.* Academic Press, New York, pp. 241–310.
Palau,J., Argos,P. and Puigdomenech (1982) *Int. J. Peptide Protein Res.*, **19**, 394–401.
Palmenberg,A.C., Kirby,E.M., Janda,M.R., Drake,N.L., Duke,G.M., Potratz,K.F. and Collet,M.S. (1984) *Nucleic Acids Res.*, **12**, 2969–2985.
Robinson,I.K. and Harrison,S.C. (1982) *Nature*, **297**, 563–568.
Rossmann,M.G. and Argos,P. (1975) *J. Biol. Chem.*, **250**, 7523–7532.
Rossmann,M.G., Abad-Zapatero,C., Murphy,M.R.N., Liljas,L., Manavalan,P. and Ponnuswamy,P.K. (1978) *Nature*, **275**, 673–674.
Rossmann,M.G., Arnold,E., Erickson,J.W., Frankenberger,E.A., Griffith,J.P., Hecht,H.-J., Johnson,J.E., Kamer,G., Luo,M., Mosser,A.G., Rueckert,R.R., Sherry,B. and Vriend,G. (1985) *Nature*, **317**, 145–153.
Simons,K. and Warren,G. (1984) *Adv. Protein Chem.*, **36**, 79–132.
Söderlund,H., Kääriäinen,L., von Bonsdorff,C.-H. and Weckström,P. (1972) *J. Virol.*, **47**, 753–760.
Söderlund,H., Kääriäinen,L. and von Bonsdorff,C.-H. (1975) *Medical Biology*, **53**, 412–417.
Söderlund,H., von Bonsdorff,C.-H. and Ulmanen,I. (1979) *J. Gen. Virol.*, **45**, 15–26.
Strauss,E.G., Rice,C.M. and Strauss,J.H. (1984) *Virology*, **133**, 92–110.
Strauss,J.H. and Strauss,E.G. (1977) In Nayak,D. (ed.), *Molecular Biology of Animal Viruses.* Dekker, New York, Vol. I, pp. 111–166.
Vogel,R.H., Provencher,S.W., von Bonsdorff,C.-M., Adrian,M. and Dubochet,J. (1986) *Nature*, **320**, 533–535.
von Bonsdorff,C.-H. (1973) *Commentationes Biologicae*, **74**, 1–53.

Received on December 10, 1986; revised on January 20, 1987