Studies on the Structural Proteins of Semliki Forest Virus

By S. I. T. KENNEDY AND D. C. BURKE

Division of Biological Sciences, School of Molecular and Biological Sciences, University of Warwick, Coventry CV4 7AL

(Accepted 14 September 1971)

SUMMARY

A method is described for the growth of Semliki Forest virus in suspensions of primary chick embryo fibroblasts and for its subsequent purification. The overall purification of the virus was approximately 250-fold with a 40 to 50% recovery of haemagglutination activity. Analysis of purified Semliki Forest virus by SDS-polyacrylamide gel electrophoresis showed the presence of two distinct protein bands. The protein of the virus envelope has a molecular weight of approximately 51,000 and is a glycoprotein. The lighter protein, with a molecular weight of approximately 32,000, is associated with the nucleocapsid core of the virion and does not appear to contain carbohydrate. Preparative SDS-polyacrylamide gel electrophoresis through a discontinuous gel system yielded mg. quantities of each of the two structural proteins. Amino acid analysis revealed that the envelope protein is relatively rich in hydrophobic amino acids, whereas the core protein is rich in hydrophilic amino acids, particularly lysine and glutamate. The N-terminal amino acid of the envelope protein is valine and that of the core protein is lysine. By employing the dansyl technique, tryptic peptide maps of the envelope and core proteins of Semliki Forest virus were obtained. Approximately 25 soluble tryptic peptides were obtained from the envelope protein and approximately 38 from the core protein.

INTRODUCTION

Semliki Forest virus is a group A arbovirus which contains a molecule of single-stranded RNA of molecular weight $4.1 \times 10^6$ (A. Meager, unpublished observation). Structurally, Semliki Forest virus is spherical with a diameter of 700 to 800 Å and consists of a nucleocapsid core containing RNA of diameter approximately 280 Å (Acheson & Tamm, 1967) surrounded by an envelope containing lipid (Osterrieth & Calberg-Bacq, 1966) which carries the haemagglutinin of the virus (Kääriäinen, Simons & von Bonsdorff, 1969; Kennedy & Burke, unpublished observation). Until recently some controversy has existed as to the number of structural proteins present in Semliki Forest virus. Hay, Skehel & Burke (1968) reported finding two structural proteins: an envelope protein of molecular weight 50,000 and a nucleocapsid protein of molecular weight 35,000, while Friedman (1968) reported the presence of three proteins, two in the nucleocapsid and one in the envelope. However, Kääriäinen et al. (1969), while confirming the presence of a single species of protein in the virus envelope, showed that, in the presence of a sulphydryl reductant, the nucleocapsid contained only a single protein species. This was confirmed by Acheson & Tamm (1970a, b), who isolated and purified Semliki Forest virus nucleocapsids from infected chick embryo cells. Analysis of these nucleocapsids by polyacrylamide electrophoresis revealed the presence of a single pro-
tein species of molecular weight approximately 32,000. These authors point out that failure to include a sulphydryl reductant in the polyacrylamide gel may result in partial dimerization of the nucleocapsid protein. Presumably therefore the minor protein band of molecular weight approximately 65,000, observed by Friedman (1968) in the absence of reductant, represents the dimeric form of the nucleocapsid protein. On the basis of its structural proteins, Semliki Forest virus is thus very similar to Sindbis virus, which has an envelope protein of molecular weight ~53,000 and a nucleocapsid protein of molecular weight ~30,000 (Strauss et al. 1968; Strauss Burge & Darnell, 1969).

This paper confirms that Semliki Forest virus contains only two protein species and reports that, like Sindbis (Strauss, Burge & Darnell, 1970), a carbohydrate moiety is covalently bound to the envelope protein. We also report the amino acid composition, identification of the N-terminal amino acid and characterization by tryptic peptide mapping of the two structural proteins of Semliki Forest virus.

METHODS

Materials. Sodium dodecyl sulphate (specially pure grade), dansyl chloride (1-dimethylaminonaphthalene-5-sulphonyl chloride) and dansyl amino acid standards were obtained from British Drug Houses, Poole, Dorset. Earle's medium in powdered form was obtained from Oxoid Ltd., London, and hydrated according to the maker's instructions. Colomycin was supplied by Pharmax Ltd., Crayford, Kent. Protamine sulphate (grade I, histone free), Dowex 50 × 8–100 (dry mesh 50 to 100, hydrogen form), reduced glutathione, and β-2-thienylalanine were all purchased from Sigma Chemical Co. Ltd., London. For tryptic digestion, 1-(tosylamido-2-phenyl) ethyl chloromethyl ketone treated trypsin was obtained from the Worthington Biochemical Corporation, Freehold, New Jersey. Cleland's reagent (A grade) and L-α-amino-β-guanidinopropionic acid (A grade) were supplied by Calbiochem., London. [4,5-3H]-L-lysine HCl (4.62 c/m-mole), [14C]-L-lysine HCl (290 mc/m-mole), [14C]-L-valine (225 mc/m-mole) and [1-3H]-D-glucosamine HCl (2.4 c/m-mole) were obtained from the Radiochemical Centre, Amersham, Bucks. Acrylamide (purum) was supplied by Fluka AG, Buchs, Switzerland and NN'-methylenebisacrylamide and NNN’N’-tetramethylethylenediamine by Eastman Organic Chemicals, Rochester, New York. Acrylamide and NN’-methylenebisacrylamide were recrystallized from chloroform and acetone respectively.

All other chemicals were the best grade which could be obtained commercially.

Virus. Semliki Forest virus, the ts+ wild-type strain, was kindly supplied by Professor F. Fenner, John Curtin School of Medical Research, Australian National University, Canberra, and passaged in the brains of suckling mice as previously described (Walters, Burke & Skehel, 1967).

Infectivity and haemagglutination assays. The infectious titre of Semliki Forest virus preparations was determined by plaque assay on primary chick embryo fibroblast monolayers as described by Walters et al. (1967). The haemagglutination (HA) activity of virus suspensions was determined by employing a modification of the method of Clarke & Casals (1958) in which, because of its greater buffering capacity at pH 5.8, citrate-saline buffer replaced phosphate-saline buffer.

Growth of Semliki Forest virus in suspensions of primary chick embryo fibroblasts. The method was a modification of that described by Zwartouw & Algar (1968). A suspension of primary chick embryo fibroblasts (7 × 10⁶ cells/ml), prepared as previously described (Walters et al. 1967), was infected with Semliki Forest virus at a m.o.i. of 1 for 1 hr at 4°C.
The infected cells were diluted to 2 to 3 x 10^7 cells/ml. with Earle's medium containing 0.35 % glucose, 0.4 % sodium bicarbonate, 0.001 % phenol red and 180 units/ml. of colo-
mycin, adjusted to pH 7.5 by bubbling with 90 % O_2/10 % CO_2. The infected suspension was
dispensed in 180 ml. portions in 1 l. Erlenmeyer flasks from which the air was then dis-
paced with 95 % O_2/5 % CO_2 and the flasks tightly sealed. The cultures were incubated at
35° on an orbital shaker moving at 80 rev./min. After 18 hr the suspension was clarified by
centrifugation at 15,000 g for 30 min at 4° and the virus-containing supernatant fluid,
containing 2 x 10^9 to 10^10 p.f.u./ml., was either used immediately or stored at −20°. In
order to grow radioactively labelled virus, the isotope was added to the supplemented
Earle's medium to 10 μc/ml. or as specified in the text.

**Virus purification.** Throughout the purification procedure the temperature was maintained
between 3° and 5°. The clarified virus suspension was continuously stirred during the drop-
wise addition of saturated ammonium sulphate (prepared in Earle's medium and adjusted to
pH 7.5) to 60 % (v/v) saturation. The preparation was stirred for a further 15 min. and the
precipitate, recovered by centrifugation at 15,000 g for 30 min., was resuspended in 0.05 M-
borate, 0.12 M-NaCl, pH 9.0 (borate-saline), to approximately one-twentieth of the starting
volume and dialysed against 10 l. of borate-saline for 12 to 16 hr. The dialysed preparation
was centrifuged at 12,000 g for 30 min. and protamine sulphate was added to the supernatant
fluid to a final concentration of 1 mg./ml. The treated suspension was immediately centrifuged
at 12,000 g for 30 min. and the supernatant fluid retained. The pellet was resuspended in
10 ml. borate-saline, centrifuged at 12,000 g for 15 min. and the supernatant fluid pooled with
that of the first protamine sulphate centrifugation. Virus was pelleted from the combined
protamine sulphate supernatant fluids by centrifugation at 120,000 g for 2 hr and gently
resuspended in 5 ml. of borate-saline. The virus suspension was layered on to a
48 ml. preformed linear 20 to 50 % (w/v) sucrose gradient, prepared in borate-saline, and cen-
trifuged for 14 to 16 hr at 65,000 g. The light-scattering band of virus, some two-thirds of the
way down the tube, was collected, diluted with borate-saline and the virus pelleted by centri-
fugation at 120,000 g for 2 hr and resuspended in 3 ml. of borate-saline. This virus suspension
was layered on to a 19 ml. preformed linear 20 to 50 % (w/v) potassium tartrate gradient
prepared in borate-saline and centrifuged at 62,000 g for 12 to 14 hr. The virus band (Fig. 1)
was collected, diluted with borate-saline, and the virus sedimented by centrifugation at
120,000 g for 2½ hr. The final pellet was resuspended in 2 ml. borate-saline and stored at
−20°.

Using this purification procedure virus was purified from up to 3 l. of suspension culture
fluid. For volumes greater than this (up to 10 l.) the number of tubes employed for the sucrose
and tartrate gradients were appropriately increased.

**Extraction of virus structural proteins.** The structural proteins of Semliki Forest virus were
extracted by treatment of purified suspension with 1 % (w/v) SDS, 0.5 M-urea, 10 % (v/v)
glacial acetic acid and 1 % (v/v) 2-mercaptoethanol, incubating at 37° for 90 min., and dialys-
ing overnight against 1000 vol. of either 0.01 M-phosphate buffer, pH 7.2, containing
0.1 % SDS and 0.1 % 2-mercaptoethanol for analytical electrophoresis or against 0.005 M-
tris + 0.038 M-glycine, pH 8.4, containing 0.1 % SDS and 0.01 % Clelands reagent for pre-
parative electrophoresis (Summers, Maizel & Darnell, 1965; Maizel, Summers & Scharff,
1970).

**Analytical polyacrylamide gel electrophoresis.** Electrophoresis was done in 10 % (w/v)
acrylamide gels cast to a height of 90 mm. in 6 mm. internal diameter perspex tubes. The gel
and buffer systems were those described by Hay et al. (1968) except that the gels were over-
laid with 0.5 ml. of 0.1 M-reduced glutathione and pre-electrophoresed for 3 hr before use.
After electrophoresis the gels were stained with Coomassie blue (Maizel et al. 1970), destained by washing with 7 % (v/v) glacial acetic acid in 10 % (v/v) methanol and scanned at 570 nm. in a Chromoscan densitometer (Joyce, Loebl and Co. Ltd., Gateshead, Durham). The gels were then frozen briefly at −70° and sectioned with the aid of a mechanical chopper (the Mickle Laboratory Engineering Co., Gomshall, Surrey) into segments 1-0 mm. thick. The segments were dissolved in 0·1 ml. 100 vol. H2O2 at 60° for 2 hr in scintillation vials and, after the addition of 10 ml. of Triton-toluene scintillation fluid (4 g. PPO, 0·05 g. POPOP in 1 l. of Triton X-100-toluene (1:2, v/v)), counted in a Packard Tricarb Model 3320 scintillation spectrometer. The radioactivity due to [3H] and [14C] in gels containing both isotopes was determined using the channels ratio method of Hendler (1964). All values are corrected for background.

**Preparative polyacrylamide gel electrophoresis.** Cylindrical polyacrylamide gel columns (height 95 mm., wall width 12·5 mm.) were used at 8° in the apparatus marketed by Quickfit and Quartz (Stone, Staffordshire). A single gel system was employed containing 10 % (w/v) acrylamide, 0·27 % N,N'-methylenebisacrylamide, 0·07 % ammonium persulphate, 0·0575 % N,N,N',N'-tetramethylethlenediamine, 0·1 % SDS, 0·5 M-urea in 0·375 M-tris-chloride pH 8·9. The electrophoresis buffer was 0·05 M-tris + 0·384 M-glycine containing 0·1 % SDS and 0·01 M-reduced glutathione, pH 8·4, and the elution buffer was 0·43 M-tris + 0·23 M-acetic acid containing 0·1 % SDS and 0·01 % Cleland’s reagent, pH 8·6. Gels were pre-run for 5 hr at 100 v potential to remove unreacted persulphate and to introduce reduced glutathione into the gel matrix.

Extracted virus proteins (up to 25 mg. in 2 ml.) made 5 % (w/v) with respect to sucrose and containing bromophenol blue as tracking dye was layered on to the gel with the aid of a peristaltic pump. Electrophoresis was started at 60 v (to minimize current induced convective disturbance of the sample) and after 2 hr was increased to 100 v and maintained at this fixed potential throughout the remainder of the separation. The flow rate of the elution buffer was 32 ml/hr. The E280 of the eluate was continuously measured and fractions of 1·0 ml. were collected. A 0·1 ml. sample of each fraction was added to 10 ml. of Triton-toluene scintillation fluid, counted and appropriate fractions pooled (see Fig. 2). The pooled fractions were concentrated by ultrafiltration (Amicon W.V., Oosterhout, Holland), dialysed against 2 x 1000 vol. of distilled water for 36 hr at room temperature and lyophilised over P2O5.

**Amino acid analysis.** Samples of lyophilized protein, each approximately 2 mg. in 2 ml. of 6 M-HCl, were hydrolysed in vacuo for 18, 48 and 60 hr at 108°. Amino acid analyses were performed in a Biocal BC 100 amino acid analyser according to the method of Spackman (1963), using β-2-thienylalanine and L-α-amino-β-guanidinopropionic acid as internal standards. Half cystine was determined as cysteic acid after dimethylsulphoxide oxidation (Spencer & Wold, 1969). Tryptophan was determined after barium hydroxide hydrolysis (Noltman, Mahowald & Kuby, 1962). Destructive losses of serine and threonine were corrected by extrapolation to zero time and all values are expressed as moles of amino acid per 100 moles of all amino acid residues recovered.

**N-terminal amino acid determination.** The dansyl chloride technique of Gray & Hartley (1963) was used. Separation of the dansyl amino acid derivatives was achieved by two-dimensional thin-layer chromatography according to the method of Woods & Wang (1967).

**Peptide mapping.** Approximately 500 µg. of lyophilized protein in 250 µl. of 10 mM-NH4OH was incubated for 90 min. at 37° with 5 µl. of a freshly prepared 1 mg./ml. solution of L-(tosylamido-2-phenyl) ethyl chloromethyl ketone treated trypsin dissolved in 1 mM-CaCl2, 6H2O. A further 5 µl. of trypsin and 100 µl. and 10 mM-NH4OH were added, and
Structural proteins of Semliki Forest virus

incubation continued for 12 hr. Any trace of insoluble material was removed by centrifugation at 5000g for 15 min. and the digested sample lyophilized over P₂O₅. Separation of the tryptic peptides was performed using a modification of the method of Schmer & Kreil (1967). Approximately 250 μg. of the lyophilized sample in 50 μl. of 0.1 M-NaHCO₃ in a 30 × 5 mm. borosilicate tube was dansylated by the addition of 50 μl. of 50 mg./ml. dansyl chloride in acetone (a dansyl chloride:sample weight ratio of 10). The tube was sealed and incubated at 37° for 6 hr. After drying in vacuo over NaOH the sample was taken up in 1 ml. of 0.01 M-acetic acid, the residue allowed to settle and the clear supernatant solution transferred to a 75 × 10 mm. borosilicate tube. To this tube was added 600 μl. of a concentrated suspension of Dowex 50 × 8-100 which had previously been equilibrated with 0.01 M-acetic acid. The mixture was incubated at room temperature for 30 min. with occasional mixing, the resin allowed to settle and the supernatant solution removed and discarded. The resin was resuspended in 400 μl. of 0.01 M-acetic acid, allowed to settle and the supernatant solution again discarded. The resin was washed a further 4 times with 400 μl. portions of 0.01 M-acetic acid. The dansylated peptides were eluted from the resin by adding 500 μl. of 3.0 M-NH₄OH in 25 %(v/v) acetone, mixing and incubating at room temperature for 30 min. The resin was removed by centrifugation at 10000g for 30 min. and the supernatant solution lyophilized over P₂O₅. The sample was redissolved in 50 μl. of acetone-glacial acetic acid (3:2, v/v) and 10 μl. samples of the dansylated peptides separated by two-dimensional chromatography on polyamide thin layers (Polyamide-6 from Macherey-Nagel and Co., Duren, Germany) using water-formic acid (200:3, v/v) in the first vector and, after drying for 5 min. at 110°, benzene-acetic acid (9:1, v/v) in the 90° vector. The separated peptides were directly visualized by viewing the developed plate under 254 nm. radiation. A record of the peptide map was obtained by tracing the position of each fluorescent spot on to a sheet of clear perspex ½ in. thick, interposed between the u.v. lamp and the thin-layer plate.

Protein determinations. Protein concentrations were determined by the method of Lowry et al. (1952), using the continuous-flow system of Gibbs & Bright (1968). Bovine serum albumin (fraction V from Armour Pharmaceuticals Co. Ltd., Eastbourne) was used as standard.

RESULTS

In order to study the chemistry of the structural proteins of Semliki Forest virus the large-scale growth and purification procedures detailed in the Methods section were developed. The result of the final purification step, equilibrium centrifugation through potassium tartrate, of [3H]-lysine-labelled virus is shown in Fig. 1. The virus assayed by its HA activity formed a sharp band at a mean buoyant density of 1.18 g./ml. Purity of the virus was indicated by the coincidence of the HA activity distribution with both the E₂₈₀ and radioactivity profiles, no other u.v. absorbing or radioactive material being present. Typically, the overall purification of the virus was 250-fold with a 40 to 50 % recovery of HA activity and a 15 to 20 % recovery of infectivity. Routinely up to 4 mg. of purified virus (on a protein basis) was obtained per l. of suspension culture fluid.

Analytical electrophoresis of the structural proteins of Semliki Forest virus

Purified virus, labelled with [¹⁴C]-lysine and [³H]-glucosamine, was disrupted and the structural proteins fractioned by electrophoresis through 10 %(w/v) polyacrylamide gels. Following electrophoresis the gels were stained, scanned and then sectioned and the distribution of radioactivity along the gel determined as described in the Methods section. The composite result of one such electrophoretic separation is shown in Fig. 2. Two protein bands were
Fig. 1. Equilibrium centrifugation of Semliki Forest virus on a potassium tartrate gradient. Virus, labelled with $[^{3}H]$-lysine, was purified as described in the Methods section before gradient analysis. Fractions of 10 drops were collected and assayed for HA activity ($\bigcirc - \bigcirc$), $E_{max}$ ($\triangle - \triangle$) and radioactivity ($\square - \square$). Densities were calculated from refractive index measurements.

Fig. 2. Analytical SDS-polyacrylamide gel electrophoresis of the structural proteins of Semliki Forest virus, labelled with $[^{3}H]$-glucosamine and $[^{14}C]$-lysine. Following electrophoresis the gel was stained with Coomassie blue, scanned at 570 nm. ($\bigcirc - \bigcirc$), frozen, sliced and radioactivity of both $[^{14}C]$-lysine ($\triangle - \triangle$) and $[^{3}H]$-glucosamine ($\square - \square$) determined.
observed by optical scanning (designated 1 and 2). The position of these two bands corresponded exactly with those of the [14C]-lysine peaks. This pattern is in close agreement with that obtained by other workers (Hay et al. 1968; Kääriäinen et al. 1969; Acheson & Tamm, 1970b), who have identified protein 1 as the envelope protein of the virus and protein 2 as the core or nucleocapsid protein with mol. wts of 51,000 and 32,000 respectively. It was also observed that the [3H]-glucosamine profile displayed only a single peak completely coincident with the envelope protein, indicating that the envelope protein of Semliki Forest virus carries a covalently bound carbohydrate moiety. Since the [3H]-glucosamine:[14C]-lysine ratio was constant across the width of the envelope protein (fractions 38 to 54) the possibility that this band, because of its width and occasionally anomalous splitting (Acheson & Tamm, 1970b) is composed of more than a single protein species, can probably be discounted. However, since nothing is known either about the molecular weight of the carbohydrate moiety or the number of such moieties per protein molecule, estimations of parameters based on a molecular weight of 51,000 for the protein moiety should be made with some caution. Indeed a molecular weight of 51,000 for the glycoprotein may be inaccurate since little is known regarding the electrophoretic behaviour of glycoproteins through SDS-polyacrylamide gels.

![Electrophoretic separation of structural proteins](image)

**Preparative separation of the structural proteins of Semliki Forest virus**

In view of the resolving power of polyacrylamide gel electrophoresis in separating the two structural proteins of Semliki Forest virus at an analytical level, the application of this technique at a preparative level was investigated. After preliminary experiments, a discontinuous tris-glycine system based on that of Ornstein (1964) and Davis (1964) was adopted, in preference to the phosphate system (Summers et al. 1965) employed hitherto, largely because of the lower conductivity and higher applied potential—factors which resulted in a more
rapid and reproducible separation. The results of a typical preparative separation of the two viral structural proteins, labelled with [\(^{3}H\)]-lysine, are illustrated in Fig. 3. The separation of the envelope from the core protein was essentially complete. After pooling appropriate fractions the two proteins were concentrated and analysed by analytical polyacrylamide gel electrophoresis. Both the envelope and core protein preparations were judged to be homogeneous (Fig. 3, inserts 1 and 2).

**Amino acid composition of the envelope and core proteins**

The results of amino acid analyses of the envelope and core proteins of Semliki Forest virus are shown in Table 1. In general, they are similar to those of Simons & Kääriäinen (1970), except for threonine and serine where the present values are higher. Both these amino acids are slowly destroyed under the conditions used for acid hydrolysis (Duggan, 1951), and higher, corrected values are therefore obtained by extrapolation to zero time: no such correction was applied by Simons & Kääriäinen (1970). The core protein is relatively rich in hydrophilic amino acids, notably lysine and glutamate, probably reflecting the structural role of the core protein in encapsulating the virus RNA. The envelope protein is more hydrophobic in character, as would be expected of a protein which interacts with lipid.

**Table 1. Amino acid composition of the structural proteins of Semliki Forest virus**

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Envelope protein</th>
<th>Core protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tryptophan</td>
<td>1.08 (±0.03)*</td>
<td>0.89 (±0.04)</td>
</tr>
<tr>
<td>Lysine</td>
<td>5.74 (±0.02)</td>
<td>12.84 (±0.05)</td>
</tr>
<tr>
<td>Histidine</td>
<td>3.91 (±0.03)</td>
<td>2.60 (±0.02)</td>
</tr>
<tr>
<td>Arginine</td>
<td>3.93 (±0.03)</td>
<td>4.97 (±0.04)</td>
</tr>
<tr>
<td>Aspartate</td>
<td>8.45 (±0.02)</td>
<td>8.01 (±0.02)</td>
</tr>
<tr>
<td>Threonine</td>
<td>9.03 (±0.07)</td>
<td>6.27 (±0.07)</td>
</tr>
<tr>
<td>Serine</td>
<td>7.14 (±0.10)</td>
<td>5.03 (±0.06)</td>
</tr>
<tr>
<td>Glutamate</td>
<td>8.17 (±0.04)</td>
<td>10.16 (±0.04)</td>
</tr>
<tr>
<td>Proline</td>
<td>7.21 (±0.06)</td>
<td>8.23 (±0.05)</td>
</tr>
<tr>
<td>Glycine</td>
<td>6.98 (±0.01)</td>
<td>9.14 (±0.02)</td>
</tr>
<tr>
<td>Alanine</td>
<td>8.37 (±0.02)</td>
<td>8.78 (±0.04)</td>
</tr>
<tr>
<td>l-Cystine</td>
<td>2.09 (±0.06)</td>
<td>1.54 (±0.06)</td>
</tr>
<tr>
<td>Valine</td>
<td>7.72 (±0.02)</td>
<td>6.62 (±0.01)</td>
</tr>
<tr>
<td>Methionine</td>
<td>1.85 (±0.03)</td>
<td>2.55 (±0.04)</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>4.30 (±0.02)</td>
<td>3.92 (±0.03)</td>
</tr>
<tr>
<td>Leucine</td>
<td>5.68 (±0.03)</td>
<td>3.78 (±0.04)</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>4.18 (±0.01)</td>
<td>2.30 (±0.02)</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>3.17 (±0.03)</td>
<td>2.37 (±0.06)</td>
</tr>
</tbody>
</table>

* Values are expressed as moles/100 moles of all amino acid residues recovered, and have been corrected for destructive loss to threonine and serine. The mean and standard deviation of four analyses are shown.

**Identification of the N-terminal amino acid of the envelope and core proteins**

By employing the dansyl technique the N-terminal amino acid of the envelope protein was identified as valine and that of the core protein as lysine. This result was confirmed by N-terminal analysis of envelope and core proteins labelled with \([^{14}C]\)-valine and \([^{14}C]\)-lysine respectively. When the fluorescent spots due to the N-terminal amino acids were recovered from the thin-layer plates, both proved to be radioactive (3714 counts/min. and 2084 counts/min. respectively). Suitable control experiments showed that over 90% of the total radioactivity incorporated into the two structural proteins was present in the same amino acid as had been added to the growth medium. There was therefore no conversion of these amino acids into others.
Further characterization of the two structural proteins of Semliki Forest virus was done by tryptic peptide mapping. Each tryptic peptide was dansylated, and after removal of interfering quantities of the reaction by-products, dansylamine and dansyl-hydroxide, the fluorescent dansylated peptides were separated by two-dimensional thin-layer chromatography. The envelope protein (Fig. 4) gave 25 spots which were reproducibly positioned on 6 separate maps from 4 different samples of envelope protein. Three spots, however, the unshaded areas, were less reproducible, both in intensity and position. This may be due to incomplete solubilization of the envelope protein as shown by the presence of an insoluble 'core' remaining after tryptic digestion. The presence of the 'core' may explain why only 28 peptides were obtained rather than the 35 expected from the amino acid composition. In contrast, the 'core' protein (Fig. 5), which was completely digested by trypsin, gave 38 spots – almost the anticipated number of 41. These were reproducibly positioned on 4 separate maps from 3 different samples of core protein. The spot marked with an asterisk is α-di dansyl-lysine (Woods & Wang, 1967) derived from the N-terminal lysine of the core protein.

Fig. 4. Tryptic peptide map of the envelope protein of Semliki Forest virus. The shaded areas show the position of dansylated peptides, the cross-hatched areas the position of by-products and unshaded areas the position of less reproducible peptides.
DISCUSSION

Recent investigations into the structural proteins of group A arboviruses have indicated that these small RNA viruses contain two electrophoretically distinct protein species, one the core protein associated with the virus RNA, the other the envelope protein forming part of the outer lipid-containing membrane (Hay et al. 1968; Strauss et al. 1968; Igarashi et al. 1970; Acheson & Tamm, 1970b). The results presented in this paper demonstrate that in Semliki Forest virus one of these two structural proteins, the envelope protein, carries a covalently bound carbohydrate moiety. Strauss et al. (1970) have demonstrated that the envelope protein of Sindbis virus is also a glycoprotein. At present, however, nothing is known regarding the functional rule of the carbohydrate moiety but it is tempting to speculate by analogy with influenza virus (Laver & Webster, 1966) that the carbohydrate may be a component of the virus haemagglutinin. Furthermore, incorporation of monosaccharide precursor added at the onset of viral growth indicates the de novo synthesis of the carbohydrate moiety rather than incorporation of pre-existing polysaccharide. Grimes & Burge (1971) have recently shown that the addition of carbohydrate to Sindbis virus envelope protein is probably a host-controlled function.

Preparative polyacrylamide gel electrophoresis permitted the isolation of mg. quantities of the structural proteins of Semliki Forest virus and some of the chemical properties of these

---

**Fig. 5.** Tryptic peptide map of the core protein of Semliki Forest virus. The symbols are the same as for Fig. 4. The significance of the spot marked with an asterisk is described in the text.
proteins have now been studied. Detailed amino acid analysis of the structural proteins, whilst being of interest themselves, also permit the relative intracellular pool sizes of a given pair of amino acids, one labelled with [3H] and other with [14C], to be determined under conditions of infection when protein synthesis is virtually entirely virus directed. This in turn permits the ratio of these two amino acids to be determined in the non-structural viruses-induced proteins without necessitating their purification, and by studying sufficient numbers of such amino acid pairs an amino acid analysis of each of the non-structural proteins can be deduced.

By employing the dansyl technique tryptic peptide maps of the envelope and core proteins of Semliki Forest virus have been obtained. This technique offers two advantages over more classical methods of peptide mapping. Firstly, the intense fluorescence of the dansyl moiety permits complete maps to be obtained with as little as 75 μg. of protein. Secondly, comparison of maps is greatly facilitated by the two reaction by-products dansyl-amine and dansyl-hydroxide, the positions of which act as extremely reproducible internal markers. The dansyl mapping technique is being employed to examine the reported observation (Lomniczi & Burke, 1970) that a number of temperature-sensitive mutants of Semliki Forest virus may contain an aberrant envelope protein.

We thank the Medical Research Council for financial support and Miss C. Bruton for excellent technical assistance.

REFERENCES


(Received 2 August 1971)