Aromatic interactions define the binding of the alphavirus spike to its nucleocapsid

Ulrica Skoging, Mauno Vihinen†, Lennart Nilsson and Peter Liljeström*

Background: Most enveloped viruses bud from infected cells by a process in which viral intracellular core components interact with cytoplasmic domains of transmembrane spike glycoproteins. We have demonstrated previously that a tyrosine motif in the cytoplasmic domain of the Semliki Forest virus (SFV) spike glycoprotein E2 is absolutely essential for budding. In contrast, hardly anything is known regarding which region of the capsid protein is involved in spike binding. Therefore, the mechanism by which spikes are selectively sorted into the viral bud or by which energy is provided for envelopment, remains unclear.

Results: Molecular models of the SFV capsid protein (SFCP) and the cytoplasmic domain of the spike protein were fitted as a basis for a reverse genetics approach to characterizing the interaction between these two proteins. Biochemical analysis of mutants defined a hydrophobic pocket of the capsid protein that is involved both in spike binding and nucleocapsid assembly.

Conclusions: We suggest that aromatic residues in the capsid protein serve to bind the side chain of the essential E2 tyrosine providing both specificity for spike incorporation and energy for budding. The same hydrophobic pocket also appears to play a role in capsid assembly. Furthermore, the results suggest that budding may occur in the absence of preformed nucleocapsids. This is the first demonstration of the molecular mechanisms of spike–nucleocapsid interactions during virus budding.

Address: Department of Biosciences, Karolinska Institute, S-141 57 Huddinge, Sweden.

†Present address: Department of Biosciences, Division of Biochemistry, PO Box 56, Viikinkaari 5, FIN-00014 University of Helsinki, Finland.

*Corresponding author.

Key words: alphavirus, Semliki Forest virus, tyrosine signal, virus assembly, virus budding

Received: 26 Jan 1996
Revisions requested: 20 Feb 1996
Revisions received: 29 Feb 1996
Accepted: 29 Feb 1996

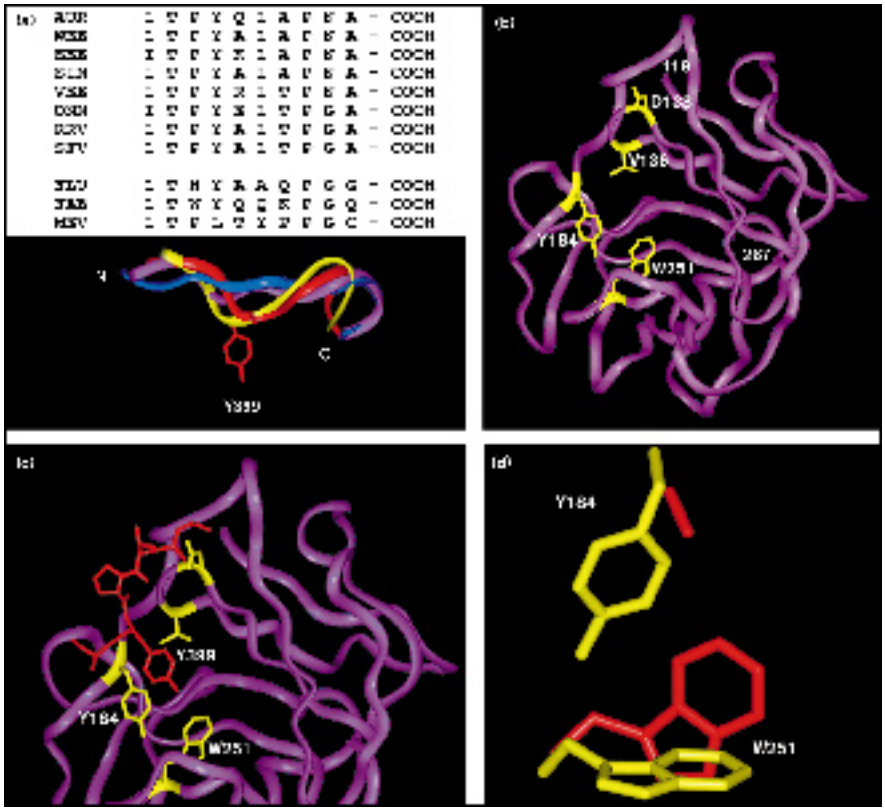
Structure 15 May 1996, 4:519–529

Introduction

Most enveloped viruses are released from infected cells by budding at the plasma membrane by a process in which viral intracellular core components are thought to interact with cytoplasmic domains of transmembrane spike glycoproteins [1]. Although retroviruses may form virus-like particles in the absence of viral spikes [2,3], both core and spike are believed to be required for budding of all other viruses. However, only in the cases of alphaviruses [4] and hepadnaviruses [5] has this been directly shown. For influenza virus, the neuraminidase protein [6] or the cytoplasmic tail of hemagglutinin [7] (both are glycoproteins associated with the viral envelope) are dispensable for budding, although the process is elevated in the presence of hemagglutinin [7]. For vesicular stomatitis virus, it has been suggested that the cytoplasmic domain of the G glycoprotein is required for budding [8,9], although further observations have indicated that budding may occur by the sole action of the membrane-associated M protein [10–12]. Inclusion of spike proteins in the viral envelope is essential for making any virus infectious, but virtually nothing is known about how spikes are sorted into the viral bud or how they may bind to core proteins.

The assembly of alphaviruses has been extensively studied, and their biology has recently been reviewed [13]. The alphavirus particle is composed of a nucleocapsid (NC) surrounded by an envelope containing two transmembrane spike glycoproteins, E1 and E2. The viral structural proteins are made from a polyprotein precursor in the order C, E2, 6K and E1. The capsid protein (C) contains a serine protease activity which is responsible for the autoproteolytical cleavage of the protein from the nascent chain. The remaining polypeptide is translocated into the endoplasmic reticulum where signal peptidase cleavages generate the individual transmembrane proteins. These are subsequently transported to the cell surface where budding occurs. The E1 protein, which carries the fusion function of the spike, has a very short cytoplasmic domain of two arginine residues, which can be removed without affecting budding [14]. The 6K protein is dispensable for budding and virus maturation, but its removal slows down virus release [15,16]. It has been suggested that the E2 protein interacts with the capsid protein during budding via its 31-residue-long cytoplasmic tail [17,18]. While the distal half of the tail functions as a signal peptide for the membrane translocation of the following 6K polypeptide [19], mutations in the proximal part disturb budding [20–23]. It has been proposed that a pentapeptide repeat in this region is important for budding, with a single conserved tyrosine residue (Tyr399) within this repeat being essential [23]. It has

Figure 1



Molecular modelling. (a) The E2 tail tyrosine pentapeptide motif of SFV compared with that of other alphaviruses ([13]; AUR, Aura virus; WEE, Western Equine Encephalitis virus; EEE, Eastern Equine Encephalitis virus; SIN, Sindbis virus; VEE, Venezuelan Equine Encephalitis virus; ONN, O'Nyong-Nyong virus; RRV, Ross River virus) and the three sequences found in the search (influenza B neuraminidase [FLU], human immunoglobulin Fab fragment [FAB] and Mengo encephalomyocarditis virus coat protein [MEV]). In the superimposition of the peptides, FLU is yellow, FAB is blue, MEV is purple and SFV is red. (b) The modelled C-terminal (protease) domain of SFCP. Residues mutated in this study are marked in yellow. The N-terminal residue 119 and the C-terminal residue 267 are indicated. (c) The E2 tail motif (red) docked into the cavity of the capsid protein. (d) The change in Trp251 orientation in the Tyr184→Ala mutant. Yellow side chains denote the wild type situation and red residues depict that of the mutant.

been suggested that palmitoylation of cysteine residues on both sides of the tyrosine motif anchor the E2 tail to the membrane [20,21], which might help to orient the tail for optimal binding to the NC.

Very little is known about which region of the capsid protein is involved in spike recognition. The Sindbis virus (SINV) capsid protein (SCP) has been crystallized and its C-terminal half was found to have a chymotrypsin-like structure [24,25]. The serine protease activity of this domain is independent of the N-terminal half of the protein [26]. This N-terminal half is believed to bind the genomic RNA [26–28], but its 3D structure is unknown. In the mature virion, 240 copies of the spike protein bind to 240 copies of the capsid protein. In these T=4 structures, E1E2 heterodimers are grouped into 80 trimeric complexes such that each complex binds capsid monomers from three separate capsomeres (morphological units) [29–33].

In this study, the structures of the Semliki Forest virus (SFV) capsid protein (SFCP) and the E2 cytoplasmic tail were studied with the aid of molecular modelling. The models formed a basis by which the spike–capsid interactions could be characterized by reverse genetics. We show that the E2 tail binds to a hydrophobic cavity of the capsid protein and that the side chain of the essential E2 tail tyrosine residue penetrates into this cavity where aromatic interactions constitute the main binding force. The results also indicate that budding may occur in the absence of preformed NCs.

Results

Modelling the SFV E2 tail pentapeptide repeat

We previously suggested that the pentapeptide repeat Leu–Thr–Pro–Tyr–Ala–Leu–Thr–Pro–Gly–Ala in the 31-residue cytoplasmic tail of the SFV E2 spike protein is involved in NC binding, and that Tyr399 is essential for budding [23]. The repeat sequence, with its two invariant proline residues, is highly conserved among alphaviruses (Fig. 1a), suggesting that this stretch may adopt a preferred fold. To find examples of such conformations, the sequence was used as a query to search the sequences in the Protein Data Bank. This search resulted in five peptides that showed high sequence similarity, four of which (80%) had the same fold. This is noteworthy as statistical analysis of sequence-similar pentapeptides in unrelated proteins showed that main-chain folds were conserved in only 20% of cases [34]. The three showing the highest sequence similarity with the SFV peptide were influenza B neuraminidase [35], immunoglobulin Fab fragment [36] and Mengo encephalomyocarditis virus coat protein [37]. Modelling

of the SFV peptide was based on superimposition of these known folds (Fig. 1a). The folds of the C-terminal part of the superimposed peptides were quite variable, particularly in the region of the proline and glycine residues near the C termini. These residues cause a strong bend of each peptide chain. Therefore, in subsequent analysis only the first half of the repeat, Leu–Thr–Pro–Tyr–Ala, was tested for capsid interaction.

Modelling the protease domain of the SFCP

To find the capsid protein region that could bind the modelled tail peptide, the known SINV capsid protein crystal structure, spanning the C-terminal residues 114–264 (the protease domain) [24,25], was first used as a template to model the corresponding region of SFCP. The SCP and SFCP polypeptides have almost identical polypeptide lengths, sharing 80% sequence similarity and 68% identity, suggesting that they would adopt similar overall structures. The residues forming the hydrophobic core domains of the SINV protease are fully conserved except for His148 and Leu169 which in SFV are substituted by asparagine and isoleucine, respectively. The catalytic triad and the residues forming the specificity pocket that is occupied by the last tryptophan residue (Trp267) in the mature protein are invariant. Unpublished data by Cheng and coworkers (referred to in [30]) recently reported that the atomic structures of the SCP and SFCP are similar with a root mean square (rms) deviation of 0.7 Å in Cα atoms, further suggesting that a molecular modelling approach would be feasible. In practice, the modelling turned out to be quite straightforward. The amino acids substitutions were performed based on information provided by a side chain rotamer library, and deletions were modeled by searching a database for structural fragments of required length and end-to-end distance. Fragments were selected that gave a low rms deviation from the target molecule and interfered least with the core region. The energy minimized structure (Fig. 1b) was tested in terms of stereochemistry, polarity features and 3D profiles. The test results, together with the structural and functional conservation of the molecule, suggest that the model is valid.

Docking the tail to the capsid

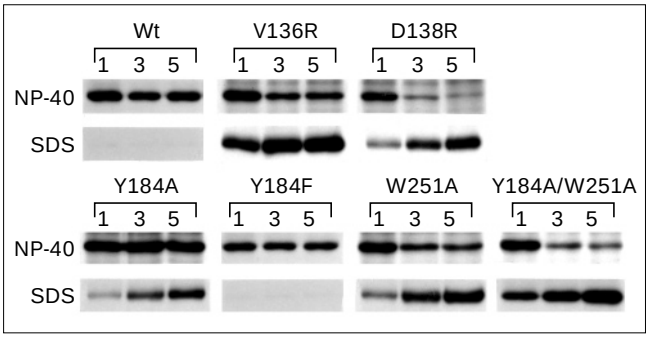
Finding the region of the capsid protein structure which could bind the E2 tail motif would have been laborious without two previous observations. Firstly, of the four tyrosine surface residues present in intact SINV NCs, only Tyr180 (Tyr184 in SFV) could be iodinated [38]. Secondly, the recent cryo-electronmicroscopy (cryo-EM) structure of the Ross River virus (RRV) provided an orientation of the capsid protein in the capsomeres which was consistent with the iodination data [30]. Together, these findings indicated which side of the molecule faces outward towards the surrounding membrane and the penetrating spike complex, and thus restricted the area of search. Although the Tyr399 residue in E2 is necessary for budding, other hydrophobic residues with bulky side chains (tryptophan, phenylalanine, leucine and methionine) are transiently tolerated [23]. Therefore, first we looked for cavities of the SFCP which could easily accommodate the Tyr399 side chain. Two such cavities were identified, one which contained Tyr132 and Cys134, and the other included Val136, Tyr184 and Trp251 (Fig. 1b). However, only the latter appeared to have reasonable interactions and structural complementarity with the modelled flanking pentapeptide residues. The tail peptide was docked into the latter cavity by taking into account spacial and electrostatic aspects and complementarity. The complete complex was energy minimized and subjected to molecular dynamics (MD) simulation to explore the conformational space. The MD showed a stable overall conformation of the bound peptide and changes during simulation were mainly fluctuations, suggesting a stable positioning of the tail peptide onto the capsid protein. According to the model, the side chain of the E2 tyrosine residue penetrates into the cavity of the core protein where the main interactions are defined by the aromatic residues Tyr184 and Trp251 of the capsid protein (Fig. 1c). Although the Tyr399 side chain appeared to be quite stably anchored, the cavity seemed to allow for some flexibility. On the surface, the N-terminal leucine residue of the peptide was in close proximity or contact with residue Asp138 of the core.

Reverse genetics

The docked model formed the basis for designing mutations that should perturb virus budding. Although the first cavity did not appear to be a serious candidate for the E2–capsid interaction, we nevertheless constructed three mutations, Tyr132→Trp, Cys134→Ile and Cys134→Met, which according to our model would disturb the binding of an inserted Tyr399 side chain. Subsequent assays showed that none of these mutants affected budding (data not shown), supporting our assumption that this cavity was irrelevant for this process. In the other candidate cavity, two mutations were designed to sterically block binding: Asp138→Arg on the surface to impair recognition and positioning of the pentapeptide, and Val136→Arg at the entrance to the cavity to prevent the Tyr399 side chain from entering (Fig. 1c). Mutations inside the cavity, Tyr184→Ala, Tyr184→Phe, Trp251→Ala, and Tyr184→Ala/Trp251→Ala (double mutant), were designed to directly affect the binding of the Tyr399 side chain (Fig. 1).

As mutation of the C gene might also affect its protease activity thus epistatically prohibiting the eventual budding by blocking spike production, plasmid derivatives of the full-length infectious cDNA clone of SFV [15] were used so that the capsid protein was expressed from one construct and the spike proteins from another. Previous studies using these constructs had shown that NC

Figure 2



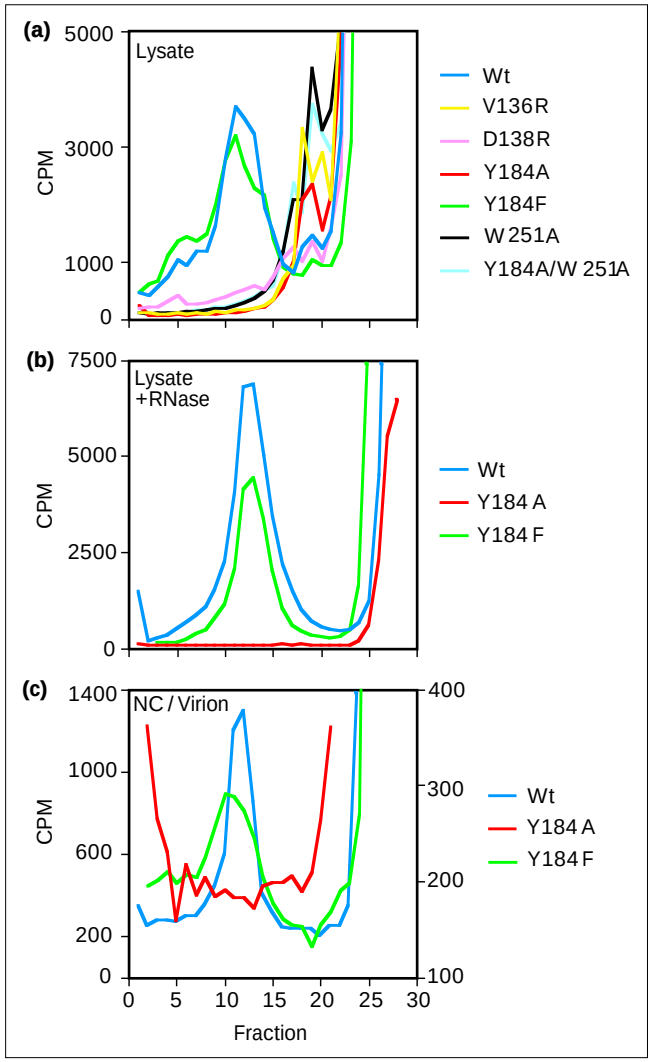
Expression of wild type and mutated capsid (C) protein analyzed by SDS-PAGE. BHK cells co-transfected with RNAs encoding the C protein or the spike proteins were pulse-labelled for 15 min and chased for 1, 3 or 5 hours. Cells were first lysed in non-ionic detergent (NP-40). Remaining detergent-insoluble material was dissolved in SDS and analyzed separately (SDS). The same results were obtained when the capsid protein was expressed alone or together with the spike proteins.

assembly, spike protein synthesis and cell surface expression were as for the wild type [4]. Expression of the capsid protein from each individual construct was first analyzed by transfecting the corresponding RNAs into BHK cells, followed by pulse-labelling. Subsequent gel analysis of the NP-40 cell lysates revealed that although the wild type and Tyr184→Phe mutant capsid proteins were stable for several hours, all other mutants resulted in apparently unstable capsid proteins (Fig. 2). However, further analysis found these capsid protein species in a sodium dodecyl sulphate (SDS)-soluble fraction. Quantitation showed that most of the initially synthesized capsid protein could be recovered (Table 1), suggesting that the disappearance of the mutant capsid proteins from the NP-40 lysates was mostly due to aggregation rather than to degradation.

NC assembly and virus release

Next, we tested whether the various mutants were able to form NCs. Cell lysates of pulse-labelled BHK cells transfected with wild type or mutant RNAs were analyzed by sucrose density gradient centrifugation under conditions in which the wild type NC bands in fractions 10–15 (Fig. 3). However, except for Tyr184→Phe, none of the mutant constructs formed any NCs (Fig. 3a). Instead, these species migrated slightly into the gradient, forming a shoulder in fractions 18–20. We suspected that this could be due to co-migration with the 60S ribosomal subunit, which is known to bind newly synthesized capsid protein [39,40]. To verify that this was the case, and that the migration was not due to multimeric forms of the capsid protein, we treated the lysates with RNase before loading on the gradient (Fig. 3b). This treatment removed this (60S) shoulder, suggesting that the capsid protein in this fraction was indeed monomeric.

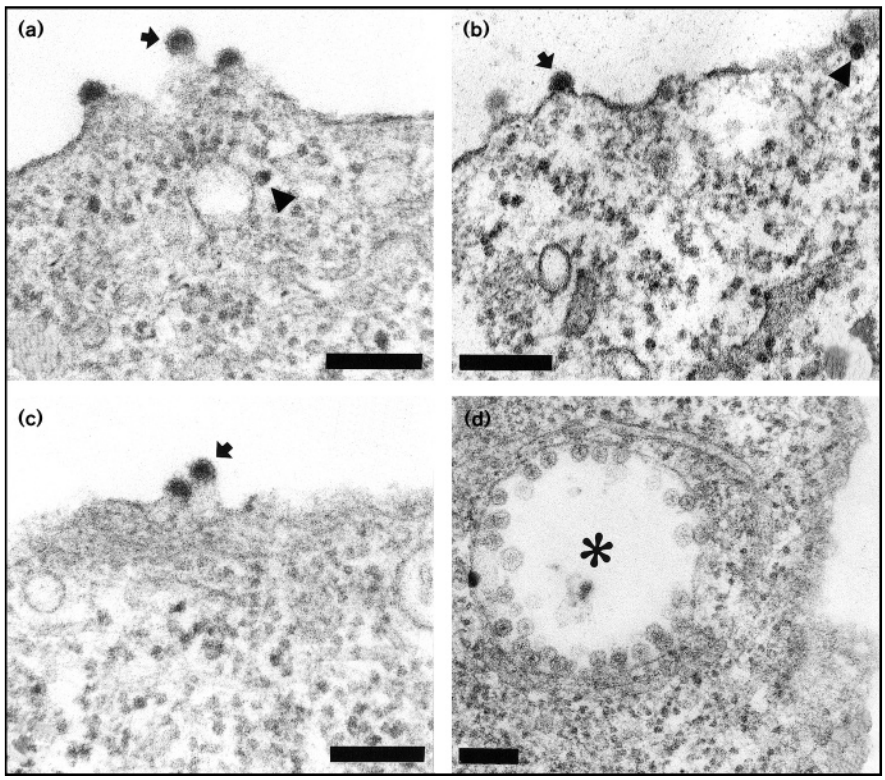
Figure 3



NC analysis by sucrose density gradient centrifugation. **(a)** Lysate loaded directly on the gradient. **(b)** Lysates pretreated with RNase before loading on the gradient. **(c)** Budded virus (nucleocapsid/virion) treated with NP-40 before loading on the gradient. In (c), TCA (trichloroacetic acid) precipitation of the fractions showed that all capsid protein had entered the gradient, and that virtually none of it was in the top fractions 25–30 (data not shown). Also in (c), the scale on the left Y axis denotes CPM (counts per minute) for wild type, and that on the right Y axis denotes Tyr184→Ala and Tyr184→Phe samples.

The lack of detectable NC in the cytoplasm could either be due to lack of assembly, or due to disassembly during preparation of the lysate. We therefore performed electron microscopy of transfected cells which revealed that, except for Tyr184→Phe, none of the mutants showed NC formation in the cytoplasm (Fig. 4). Wild type (Fig. 4a), Tyr184→Phe (Fig. 4b), and Tyr184→Ala (Fig. 4c) particles were seen budding at the plasma membrane, whereas such intermediates were absent from all the other mutants (Fig. 4d). However, CPVI (cytoplasmic vacuole type I)

Figure 4



Electron microscopy of transfected BHK cells. Thin sections of cells transfected with: (a) wild type spike and wild type capsid, (b) wild type spike and Tyr184→Phe, (c) Tyr184→Ala or (d) Trp251→Ala capsid. Budding virions at the plasma membrane of wild type transfected cells are indicated by arrows, and cytoplasmic NCs by arrowheads. A CPV I structure is indicated by a star in (d). The scale bar represents 0.2 μm.

structures stemming from RNA replication complexes [41] were present in all cases, indicating that transfection had been successful (see Fig. 4d).

To assay whether the different mutants were able to produce virus, BHK cells were co-transfected with capsid and spike RNAs, pulse-labelled and the virus particles produced were collected from the medium by centrifugation. Virus production could be detected only in the wild type, the Tyr184→Phe, and the Tyr184→Ala constructs (data not shown), thus confirming our EM results. Quantitation showed that both Tyr184→Phe and Tyr184→Ala budded with about 20% efficiency compared with wild type. None of the other mutants appeared to produce virus, or the production was too low to be detected by this method. The budding of Tyr184→Ala was particularly interesting as this mutant did not appear to produce any NC (Figs 3,4), and as the intracellular capsid protein had a tendency to aggregate (Fig. 2). We therefore treated budded Tyr184→Phe and Tyr184→Ala virus with detergent (NP-40) to strip the particles from their spikes and membrane, and assayed the preparation by sucrose density centrifugation (Fig. 3c). The Tyr184→Phe NC banded at the wild-type position in the gradient, although some aggregated forms may have been present. In contrast, for Tyr184→Ala most capsid protein appeared at the bottom of the gradient, indicating that the NC structure had disassembled and aggregated.

The results with the Tyr184→Ala mutant suggested that virus budding can occur in the absence of preformed NCs, and thus it was of interest to analyze the budding of the different mutants in greater detail. For this, we applied a more sensitive procedure which enabled very small amounts of virus to be monitored. BHK cells were co-transfected with capsid and spike RNAs, and after extensive washes at seven hours post transfection, release of new virions into fresh growth medium was monitored for the following two hours. This one-step transfection protocol combined with short incubation times was used to assure that the assay was measuring mutant rather than revertant phenotypes [23]. The 2 h medium was titrated on BHK cell monolayers and infection was monitored by indirect immunofluorescence (Table 1). This analysis showed that Tyr184→Phe and Tyr184→Ala budded with 26% and 20% efficiencies, respectively, compared with wild type. Val136→Arg did not produce any virus at all, and all other constructs produced small amounts of virus, with Asp138→Arg being the least proficient of these (Table 1).

The most striking budding result was that for Tyr184→Phe, in which removal of a single hydroxyl group reduced budding fourfold, without affecting NC formation or protein stability to any great extent. Of specific interest was the budding ability of the spike mutants in which Tyr399 was substituted for either phenylalanine

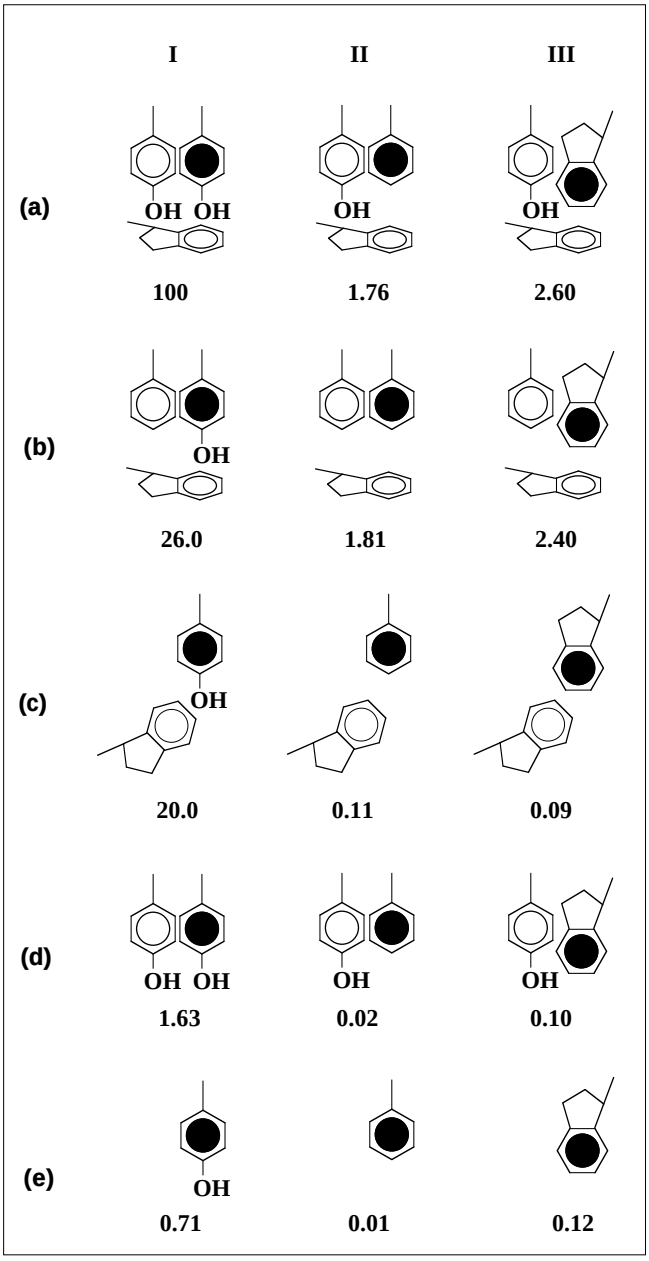
Table 1

Virus particle formation.

Construct		Virus release*	Intracellular C protein†		
Capsid	Spike (E2)	% of wt	NP-40‡	SDS‡	Total§ (%)
C_{am} (wt)#	Tyr399 (wt)	100	96/97	4/3	91
Val136→Arg	Tyr399	0	86/37	14/63	54
Asp138→Arg	Tyr399	0.05	92/36	8/64	41
Tyr184→Ala	Tyr399	20.0	92/85	8/15	88
Tyr184→Ala	Tyr399→Phe	0.11	nd	nd	nd
Tyr184→Ala	Tyr399→Trp	0.09	nd	nd	nd
Tyr184→Phe	Tyr399	26.0	96/96	4/4	103
Tyr184→Phe	Tyr399→Phe	1.81	nd	nd	nd
Tyr184→Phe	Tyr399→Trp	2.40	nd	nd	nd
Trp251→Ala	Tyr399	1.63	91/58	7/42	74
Trp251→Ala	Tyr399→Phe	0.02	nd	nd	nd
Trp251→Ala	Tyr399→Trp	0.10	nd	nd	nd
Tyr184→Ala/Trp251→Ala	Tyr399	0.71	90/30	10/70	64
Tyr184→Ala/Trp251→Ala	Tyr399→Phe	0.01	nd	nd	nd
Tyr184→Ala/Trp251→Ala	Tyr399→Trp	0.12	nd	nd	nd
C_{am} (wt)#	Tyr399→Phe	1.76	nd	nd	nd
C_{am} (wt)#	Tyr399→Trp	2.60	nd	nd	nd

*2 h medium titrated by immunofluorescence of BHK infected cells using either anti-capsid or anti-E2 antibodies. (As both RNA species can be packaged, infected cells will express either capsid or spike proteins. There were no significant differences between individual experiments using either antibody type.) †RNA encoding the capsid protein was transfected into BHK cells which were pulse-labelled at 7 h for 10 min, then chased for 10 mins or 130 min. NP-40 or SDS lysates were analyzed by SDS-PAGE and the amount of capsid protein quantitated. ‡Amount (%) of capsid protein in the 10 min/130 min chase samples. §Amount of capsid protein present in the 130 min chase sample as a % of that found in the 10 min chase sample. #Wild type capsid gene followed by a double amber stop codon. Any results not determined are marked nd.

Figure 5



Summary of aromatic interactions in the spike-binding site of the capsid protein. Only side chains are depicted. The Tyr399 residue of the E2 spike cytoplasmic domain is highlighted in black. The other residues depict Tyr184 and Trp251 of the capsid protein. Numbers under the various groups indicate budding efficiencies compared with wild type, which is given as 100% (compare with Table 1). (a) wild-type capsid protein; (b) Tyr251→Phe capsid mutant; (c) Tyr184→Ala capsid mutant; (d) Trp251→Ala capsid mutant; (e) Tyr184→Ala and Trp251→Ala capsid double mutant. I, wild type (Tyr399) E2 tail residue; II, Try399→Phe and III, Tyr399→Trp.

or tryptophan. Specifically, budding occurred in these mutants although at efficiencies significantly lower than the wild type [23]. This prompted us to further characterize the spike–NC interaction by testing the Tyr184→Phe, Tyr184→Ala and Trp251→Ala capsid mutants in combination with the Phe399 and Trp399 spike variants. In these combinations, budding was reduced to 2–3% when the capsid was of wild type or Tyr184→Phe, but much larger reductions were obtained for the Tyr184→Ala (reduced to ≈0.1%) or Trp251→Ala (≤0.1%) capsid variants (Table 1 and Fig. 5).

Discussion

Budding can occur in the absence of preformed NCs

Our reverse genetics analysis clearly defined the region of the SFCP that interacts with the Tyr399 during virus assembly. At the onset of our study it appeared important to show that the mutants were able to efficiently form NCs, so that mutational effects on budding could be confined to envelopment rather than to NC formation or stability. The Tyr184→Phe mutant was, therefore, central in showing that it was possible to significantly affect budding by the removal of a single hydroxyl group from the predicted binding cavity, while maintaining the stability of the protein monomer and corresponding NC. All other mutations resulted in an unstable capsid protein that aggregated in the cell. Although the aggregated forms are

unlikely to function as precursors for NC formation, the aggregation was nevertheless sufficiently slow in all cases to allow significant amounts of soluble capsid protein to be present during the 2 h budding assay (Table 1). Thus, aggregation itself should not *a priori* have had such a profound effect on budding, although the possibility that some mutations may have resulted in incompletely folded proteins cannot be ruled out. As the Tyr184→Ala mutation showed that NC formation was not required for budding, it is highly likely that the different mutations affected the budding process directly. For Tyr184→Ala, a capsid protein resulted that slowly aggregated in the cell without forming any NCs. Nevertheless, infectious virus particles were formed with an efficiency of 20% compared with wild type, and budding intermediates assembling at the plasma membrane of infected cells were readily seen by EM. These observations suggest that alphavirus budding can occur without preformed NCs. This is reminiscent of C-type retrovirus budding in which gag core particles form at the plasma membrane during budding.

An aromatic network defines the spike–NC interaction

The model of E2–capsid protein binding predicts that several specific interactions are required. This was expected as earlier attempts to isolate capsid suppressors for E2 tail mutations had failed [23]. Based on our binding model, the capsid Asp138 residue was predicted to be in such close contact with the pentapeptide that its substitution for an arginine residue should severely affect the positioning of the peptide on the capsid protein surface and may completely abolish budding. Val136 is situated at the entrance of the predicted binding cavity and its substitution for an argine residue should severely affect the penetration of the Tyr399 side chain. Our biochemical data fully support these predictions. It was suggested that within the cavity the hydroxyl group of E2 Tyr399 interacts with the edge of the aromatic ring of the capsid Trp251; the hydroxyl oxygen would be attracted by the positively charged aromatic ring hydrogens and repelled by the electron rich π-cloud, to which it would donate a hydrogen bond [42]. Similar interactions have been found to be preferred in numerous cases involving phenylalanine residues [43], and for interactions between sulphur and aromatic rings [44]. Such an interaction is also energetically favorable according to our MD simulations. Our analyses suggest that the Tyr399 side chain also interacts with Tyr184 in the cavity. The side chains of these two residues form a parallel sandwich so that Tyr184 is located furthest from the cavity entrance. A majority of known aromatic interactions have angles within 30° to 90° [45], but near parallel organization, such as could be the case here, is also rather frequent [46]. Thus, we suggest that Tyr399 forms an energetically favorable aromatic network with residues Tyr184 and Trp251. Such networks are known to stabilize protein structures, and a buried aromatic interaction can contribute, by a free energy of –0.6 to –1.3 kcal mol^{-1}, to stability at the physiological temperature [45]. The overall binding of Tyr399 is also favored as this excludes water from the hydrophobic cavity.

We tested the suggested interactions by a nested set of mutations (Fig. 5). When Tyr184 was substituted for phenylalanine, budding was reduced to 26% of wild type, suggesting that the hydroxyl group of Tyr184 plays a role in the docking of Tyr399 (Fig. 5b[I]). When the whole aromatic side chain of Tyr184 was removed, the same result was obtained (Fig. 5c[I]). This can be interpreted in two ways. The most simple explanation is that Tyr184 acts via its hydroxyl group and hence there should be no difference between Figure 5b[I] and c[I]. However, when the Tyr184→Ala mutation (Fig. 5c[I]) was analyzed by molecular modelling and MD simulation, we observed that the Trp251 changes its position in the cavity by rotating nearly 90° counterclockwise (shown in Fig. 1d and Fig. 5c). Thus, removal of the complete side chain of Tyr184 seems to have a more dramatic effect than removal of the hydroxyl group alone, although this change may be compensated for by the altered orientation of Trp251. The fact that the Tyr184→Ala mutation resulted in a capsid protein structure unable to form stable NCs supports the notion of an altered orientation of Trp251 and that this has a major suppressive effect on budding. When the side chain of Trp251 was removed (Fig. 5d[I]), budding dropped dramatically, to 1%, suggesting that this residue is essential to the main binding interaction. Indeed, when both Tyr184 and Trp251 side chains were removed (Fig. 5e[I]), the same result was obtained, indicating that as Trp251 was the main binding partner, removal of Tyr184 did not have any additional effect. Tyr184 could thus function in guiding the incoming Tyr399 side chain into the correct position for interaction with Trp251.

The notion that the Tyr399 side chain interacts with Trp251 is further supported by the results using the spike mutant variant in which the tyrosine residue was changed to phenylalanine, thus removing the hydroxyl group while otherwise maintaining the overall structure of this residue. In this case (Fig. 5a[II]), budding also dropped dramatically. When the hydroxyl group of Tyr184 was also removed (Fig. 5b[II]), there was no further effect, which can be explained by the lack of interaction between Tyr399 and Trp251. Thus, a simple sandwich effect between Tyr399 and Tyr184 would result in 1–2% budding. The drop in budding efficiency that is illustated in Figure 5b[I,II] would suggest that the hydroxyl group of Tyr184 has a role in the guiding of the tail Tyr399 to Trp251 but no effect on sandwiching. That this hydroxyl group may have a minor role in budding is also suggested by the fact that this residue is phenylalanine in a closely related alphavirus, O'Nyong-nyong (ONN) [25]. It could be that when the hydroxyl group of Tyr184 is removed its ability to guide the tail Tyr399 is diminished due to loss of

intramolecular interactions (according to our model the hydroxyl group of Tyr184 could interact with either Val231 or Lys188). The drop from 20% budding (Fig. 5c[I]) to only 0.11% (Fig. 5c[II]) may be the result of two factors. The hydroxyl group of the tail residue is missing, which should result in 2% budding (compared with that for the Fig. 5a[II] structure), but as the sandwiching effect of Tyr184 is missing, the efficiency drops by another tenfold. The results for the structures of Figures 5d,e [II] are more difficult to interpret. Very low levels of budding were observed, and it could be that removal of the hydroxyl group from the spike tail makes sandwiching less stable resulting in the observed drop from d[I] to d[II] or e[I] to e[II]. It is also possible that as the size of the cavity increases upon removal of Trp251, there is now room for a water molecule(s) in the cavity, and that the drop from d[I] to d[II], or from e[I] to e[II], is a result of the increased hydrophobicity around the water molecule. With a tryptophan replacing the E2 Tyr399 (d[III] and e[III]) there is no room for the water molecule and budding is similar to that of c[III]. Finally, the results using the tryptophan variant of the spike component are in agreement with those of group II, as both set of experiments involve aromatic side chains lacking the hydrophilic group through which binding is affected.

The spike-binding cavity is involved in NC assembly

It is unlikely that all point mutations would abolish NC formation simply by changing the protein structure. In fact, when the mutations were analyzed by MD simulations, no significant conformational changes were observed. Moreover, the mutated residues of the capsid protein are not situated in interface regions between capsid monomers in the NC and should therefore not interfere with capsomere assembly. In the accompanying paper [47], results are presented which give a clue as to how the phenotypes of various capsid mutations may be explained. They found that two leucine residues in the N-terminal region of the SCP (Met113 and Ile115 in SFCP) bind to the E2-binding cavity of an adjacent molecule in the capsid protein crystal. Mutation of the leucine residues abolished NC formation but not virus release, although this decreased severely. Quite recently, K Forsell and H Garoff (personal communication) also found that short deletions spanning this very N-terminal region of the SFCP were able to bud quite efficiently, but were completely deficient in NC formation. Finally, in SCP, a Tyr180→Ser/Glu183→Gly double mutant (corresponding to Tyr184/Gly187 in SFCP) resulted in a thermosensitive phenotype in which a portion of isolated virions had a T=7 type organization and in which capsid–spike interactions were affected [48,49]. Altogether, these observations suggest that binding of the N-terminal portion of the capsid protein may be important for NC or capsomere assembly. If so, then mutation of the N-terminal region, or its binding site, will result in a changed structure that is less efficient in NC assembly leading to slow aggregation of the protein. The fact that the Tyr184→Ala mutant virus in the present study was unstable after spike removal is in line with this notion, and suggests that virus assembly could occur via a co-assembly mechanism using simultaneous capsid–capsid and capsid–spike interactions. This is also consistent with the observed organization of the mature virion as trimeric spikes interconnect separate capsomeres [30,31].

The results in the accompanying paper [47] also corroborate our results on the binding site for the E2 protein. They demonstrate that the binding of the E2 tail is dependent on aromatic interactions within the hydrophobic pocket. Although the placement of Tyr399 in the two studies is rotationally different, the predicted aromatic network remains unaffected. This paper also suggests that the Leu401 residue may play a role in binding. Our study did not address this question, but does not exclude such a possibility. However, as removal of the Tyr399 side chain almost completely abolishes budding [23], the role of Leu401 may be more related to presentation of the tyrosine motif rather than contributing to the binding force. More work is needed to fully characterize the interactions between the two proteins.

Both studies found a dramatic shift in the positioning of the Trp251 residue when the side chain of Tyr184 was removed, suggesting that some kind of switch may occur upon spike binding. Altogether, our results suggest that the interaction between the alphavirus spike cytoplasmic domain and the capsid protein is dependent on several factors. In this model, where the main function of the pentapeptide repeat is to allow proper penetration of the Tyr399 residue into the hydrophobic cavity of the capsid protein. This binding would serve to provide both the energy required to curve the membrane during envelopment (240 binding pairs) as well as the specificity for incorporation of virus specific proteins into the maturing virion.

Biological implications

Maturation of enveloped viruses by budding at cellular membranes is believed to be dependent on interactions between viral core structures and cytoplasmic domains of transmembrane spike proteins. This interaction would provide both the energy for curving the membrane and specificity for selective incorporation of the viral spike proteins into the viral membrane while excluding cellular proteins from it. Incorporation of viral spike proteins is a prerequisite for making the released virion infectious. Despite much effort, the molecular details of core–spike interactions have not been defined for any virus.

In this study we have shown that a tyrosine residue in the cytoplasmic domain of the transmembrane spike glycoprotein of an alphavirus is docked into a specific hydrophobic cavity of the capsid protein. We suggest

that the binding of the tyrosine residue is dependent on an interaction with two hydrophobic residues within the capsid protein pocket, Tyr184 and Trp251. In this model the three side chains form an energetically favorable aromatic network that stabilizes the interaction and thus provides the driving force for budding and the required specificity for incorporation of correct spike proteins into the maturing virion.

Certain mutations of the hydrophobic pocket also prohibited nucleocapsid assembly without affecting virus budding to any great extent. This suggests that alphaviruses may, contrary to earlier belief, bud in the absence of preformed nucleocapsids, reminiscent of C-type retrovirus budding. Results presented in the accompanying paper [47] support and extend our findings, suggesting that nucleocapsid assembly and virus budding (spike binding) may be connected processes regulated by a molecular switch mechanism.

In a broader sense, the characterization of the molecular details of virus budding may also provide a first model to explain how protein interactions may be utilized to curve a membrane leading to the formation of a lipid vesicle carrying a defined protein cargo. Vesicle protein traffic within the cell is a central theme in cell function and numerous studies have shown that tyrosine signals located in cytoplasmic domains of transmembrane proteins are important for their incorporation into such transport vesicles.

Materials and methods

Computer modelling

Sequence alignment was carried out using program packages GCG [50] and MULTICOMP [51]. The template for molecular modelling was crystallographically determined SCP at 2.8 Å resolution [25], taken from Brookhaven Protein Data Bank [52]. Modelling was performed with Insight II (Biosym Technologies Ltd., San Diego, CA) and the minimization and molecular dynamics with CHARMM [53]. Minimization was started with 500 steps, using the steepest descent method, and finished with an adopted basis Newton-Raphson method. The predicted model was tested with programs POLDIAG [54] and VERIFY-3D [55]. The stereochemistry was checked with program PROCHECK [56]. Although only part of the protein was studied (residues 119–267), this portion of the molecule is known to form a functional entity of its own, being a serine-protease-type enzyme [24,26]. Therefore, in practice, the analysis concerned a complete protein, the structure of which should not be affected by the N-terminal part the structure. Secondary structures of the SCP were analyzed with program DSSP [57]. Standard CHARMM with parameter set PARAM22 was used in the MD simulations of core protein–E2 interaction and for studying the conformational changes in the mutant Tyr184→Ala. Program QUANTA (Molecular Simulations, Inc., Burlington, MA) was used to search for sequences related to the E2 motif in the Brookhaven Protein Data Bank.

Nucleic acid manipulations

The plasmids expressing either the capsid protein alone or the spike proteins and the mutant derivatives have been described [4,23]. *In vitro* mutagenesis [58] and *in vitro* transcription [15] were carried out as described previously.

Cell manipulations

BHK-21 cells (ATCC CL13) were grown to late log phase in 75 cm^2 bottles with 10–15 ml complete BHK medium (G-MEM [Gibco], 5% fetal calf serum, 10% tryptose phosphate broth, 20 mM HEPES [N-2-hydroxyethyl-piperazine-N′-2-etanesulphonic acid], 2 mM glutamine, 0.1 U ml^{-1} penicillin, 0.1 μg ml^{-1} streptomycin). Transfections were performed as described previously [15].

For metabolic labelling, cells were washed twice with prewarmed PBS, overlaid with starvation medium (methionine free MEM [Gibco], 2 mM glutamine, 20 mM HEPES) and incubated for 30 min. The medium was replaced with the same containing 100 μCi ml^{-1} of ^{35}S-methionine (Amersham) and incubated for 15 min for stability studies or 30 min for NC and virus formation studies. The medium was aspirated and the cells washed twice with chase medium (E-MEM [Gibco/Life Technologies Ltd, Paisley, Scotland], 2 mM glutamine, 20 mM HEPES, 150 μg ml^{-1} unlabelled methionine), overlaid with the same and incubated for different chase time periods. After the chase the cells were washed with cold PBS before NP-40 lysis buffer (1% NP-40, 50 mM Tris-HCl pH 7.6, 150 mM NaCl, 2 mM EDTA, 1 μg ml^{-1} PMSF, 10 mM iodoacetamide) was added to the dishes which were then placed on ice for 10 min. The lysed cells were scraped off and the dish was rinsed with additional NP-40 lysis buffer. The lysates were centrifuged at 6000 rpm for 5 min. SDS lysis buffer (1% SDS, 50 mM Tris-HCl pH 7.6, 150 mM NaCl, 2 mM EDTA, 1μg ml^{-1} PMSF, 10 mM iodoacetamide) was added to the NP-40 insoluble pellet and the lysates was homogenized by pulling through a 23G needle. To assay for released virus particles, the 5 h chase medium was centrifuged using a Beckman JA 18.1 rotor (Palo Alto, CA) at 17 000 rpm, for 2 h. The pellet was resuspended in TNE (50 mM TRIS-HCl pH 7.5, 100 mM NaCl, 1 mM EDTA) or sample buffer and analyzed by SDS-PAGE.

For immunofluorescence studies the cells were washed extensively with PBS 7 h after electroporation and incubated in MEM-BSA (minimal essential medium, 1 % bovine albumin fraction V [Gibco]) for an additional 2 h. The medium was collected and centrifuged at 6000 rpm for 5 min before it was used to infect BHK cells grown on coverslips. After a 1 h infection the medium was changed to complete BHK medium and the cells were incubated overnight. Indirect immunofluorescence was then carried out as described previously [59].

Electron microscopy

BHK cell were co-transfected with RNAs encoding capsid or spike proteins, incubated for 7 h, and then fixed in 2% glutaraldehyde in 0.1 M sodium cacodylate buffer, pH 7.4 at room temperature. Fixed cells were scraped off the plate and pelleted at 14 000 rpm for 2 min in an Eppendorf centrifuge and washed in 0.15 M sodium cacodylate buffer, pH 7.4. Specimens were post-fixed in 1% osmium tetroxide in 0.15 M sodium cacodylate buffer for 1 h at 4°C, and dehydrated in ethanol and acetone, and finally embedded in LX-112 (Ladd, Burlington, Vermont, USA). Thin sections were cut with a LKB (Bromma, Sweden) ultramicrotome and examined in a Phillips 420 electron microscope (Eindhoven, The Netherlands) operating at 60 kV.

Assay of NC formation

NC formation was analyzed by incubating 100 μl lysate (with or without pretreatment of RNase A 100 μg ml^{-1} for 10 min on ice) for 15 min on ice with 25 mM EDTA, before loading on a 15–30% (w/w) sucrose gradient in TNE NP-40 buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1 mM EDTA, 0.1% NP-40). Centrifugation was performed in a Beckman SW 41 rotor at 40 000 rpm for 2 h at 4°C. The gradients were fractionated from the bottom using a Gilson microfraction collector. 40 μl from each fraction was added to 2 ml lsc-cocktail (Packard, Groningen, The Netherlands) and radioactivity was measured using a liquid scintillation counter (LKB Wallac 1214 Rackbeta, Sollentuna, Sweden). NC stability was analyzed by incubation of 9 ml medium from a 5 h chase with 1 ml 10% NP-40 to strip the spike proteins and membrane from the NCs. The medium was then concentrated by using a Centriprep 100 (Amicon, Beverly, MA) and incubated with 10 mM IAA (iodoacetamid) and 2 mM EDTA before loading it on top of the gradient and centrifuges, as above.

Acknowledgements

We thank Birgitta Lindqvist and Kjell Hultenby for expert technical assistance. We are indebted to Richard Kuhn, Michael Rossmann and Henrik Garoff for sharing their data prior to publication. This work was supported by the Swedish Natural Science Foundation Council to LN and by the Swedish Medical Research Council and Swedish Research Council for the Engineering Sciences to PL. US is the recipient of a Karolinska Institute graduate student scholarship.

References

1. Simons, K. & Garoff, H. (1980). The budding mechanism of enveloped animal viruses. *J. Gen. Virol.* 50, 1–21.
2. Delchambre, M., *et al.*, & Bex, F. (1989). The GAG precursor of Simian Immunodeficiency virus assembles into virus-like particles. *EMBO J.* 8, 2653–2660.
3. Gheysen, D., *et al.*, & De Wilde, M. (1989). Assembly and release of HIV-1 precursor Pr55^{gag} virus-like particles from recombinant Baculovirus-infected cells. *Cell* 59, 103–112.
4. Suomalainen, M., Liljeström, P. & Garoff, H. (1992). Spike protein–nucleocapsid interactions drive the budding of alphaviruses. *J. Virol.* 66, 4737–4747.
5. Bruss, V. & Ganem, D. (1991). The role of envelope proteins in hepatitis B virus assembly. *Proc. Natl. Acad. Sci. USA* 88, 1059–1063.
6. Liu, C., Eichelberger, M.C., Compans, R.W. & Air, G.M. (1995). Influenza type A virus neuraminidase does not play a role in viral entry, replication, assembly, or budding. *J. Virol.* 69, 1099–1106.
7. Jin, H., Leser, G.P. & Lamb, R.A. (1994). The influenza virus hemagglutinin cytoplasmic tail is not essential for virus assembly or infectivity. *EMBO J.* 13, 5504–5515.
8. Metsikkö, K. & Simons, K. (1986). The budding mechanism of spikeless vesicular stomatitis virus particles. *EMBO J.* 5, 1913–1920.
9. Whitt, M.A., Chong, L. & Rose, J.K. (1989). Glycoprotein cytoplasmic domain sequences required for rescue of a Vesicular Stomatitis virus glycoprotein mutant. *J. Virol.* 63, 3569–3578.
10. Chong, L.D. & Rose, J.K. (1994). Interactions of normal and mutant Vesicular Stomatitis virus matrix proteins with the plasma membrane and nucleocapsids. *J. Virol.* 68, 441–447.
11. Li, Y., Luo, L., Schubert, M., Wagner, R.R. & Kang, C.Y. (1993). Viral liposomes released from insect cells infected with recombinant baculovirus expressing the matrix protein of Vesicular Stomatitis virus. *J. Virol.* 67, 4415–4420.
12. Justice, P.A., Sun, W., Li, Y., Grigera, P.R. & Wagner, R.R. (1995). Membrane vesiculation function and exocytosis of wild type and mutant matrix proteins of Vesicular Stomatitis virus. *J. Virol.* 69, 3156–3160.
13. Strauss, J.H. & Strauss, E.G. (1994). The alphaviruses: gene expression, replication, and evolution. *Microbiol. Rev.* 58, 491–562.
14. Barth, B.U., Suomalainen, M., Liljeström, P. & Garoff, H. (1992). Alphavirus assembly and entry: the role of the cytoplasmic tail of the E1 spike subunit. *J. Virol.* 66, 7560–7564.
15. Liljeström, P., Lusa, S., Huylebroeck, D. & Garoff, H. (1991). *In vitro* mutagenesis of a full-length cDNA clone of Semliki Forest virus: the 6000-molecular-weight membrane protein modulates virus release. *J. Virol.* 65, 4107–4113.
16. Loewy, A., Smyth, J., von Bonsdorff, C.-H., Liljeström, P. & Schlesinger, M.J. (1995). The 6-kilodalton membrane protein of Semliki Forest virus is involved in the budding process. *J. Virol.* 69, 469–475.
17. Metsikkö, K. & Garoff, H. (1990). Oligomers of the cytoplasmic domain of the p62/E2 membrane protein of Semliki Forest virus bind to the nucleocapsid *in vitro*. *J. Virol.* 64, 4678–4683.
18. Collier, N.C., Adams, S.P., Weingarten, H. & Schlesinger, M.J. (1992). Inhibition of enveloped RNA virus formation by peptides corresponding to glycoprotein sequences. *Antiviral Chem. Chemother.* 3, 31–36.
19. Liljeström, P. & Garoff, H. (1991). Internally located cleavable signal sequences direct the formation of Semliki Forest virus membrane proteins from a polyprotein precursor. *J. Virol.* 65, 147–154.
20. Gaedigk-Nitschko, K. & Schlesinger, M.J. (1991). Site-directed mutations in Sindbis virus E2 glycoprotein's cytoplasmic domain and the 6K protein lead to similar defects in virus assembly and budding. *Virology* 183, 206–214.
21. Ivanova, L. & Schlesinger, M.J. (1993). Site-directed mutations in the Sindbis virus E2 glycoprotein identify palmitoylation sites and affect virus budding. *J. Virol.* 67, 2546–2551.
22. Lopez, S., Yao, J.-S., Kuhn, R.J., Strauss, E.G. & Strauss, J.H. (1994). Nucleocapsid–glycoprotein interactions required for assembly of alphaviruses. *J. Virol.* 68, 1316–1323.
23. Zhao, H., Lindqvist, B., Garoff, H., von Bondsdorff, C.-H. & Liljeström, P. (1994). A tyrosine-based motif in the cytoplasmic domain of the alphavirus envelope protein is essential for budding. *EMBO J.* 13, 4204–4211.
24. Choi, H.-K., *et al.*, & Wengler, G. (1991). The crystal structure of Sindbis virus core protein and a proposed structure of the capsid. *Nature* 354, 37–43.
25. Tong, L., Wengler, G. & Rossmann, M.G. (1993). Refined structure of Sindbis virus core protein and comparison with other chymotrypsin-like serine proteinase structures. *J. Mol. Biol.* 230, 228–247.
26. Forsell, K., Suomalainen, M. & Garoff, H. (1995). Structure-function relation of the NH_2-terminal domain of the Semliki Forest virus capsid protein. *J. Virol.* 69, 1556–1563.
27. Geigenmüller-Gnirke, U., Nitschko, H. & Schlesinger, S. (1993). Deletion analysis of the capsid protein of Sindbis virus: identification of the RNA binding region. *J. Virol.* 67, 1620–1626.
28. Weiss, B., Geigenmüller-Gnirke, U. & Schlesinger, S. (1994). Interaction between Sindbis virus RNAs and a 68 amino acid derivative of the viral capsid protein further defines the capsid binding site. *Nucleic Acids Res.* 22, 780–786.
29. Paredes, A., *et al.*, & Prasad, B.V.V. (1993). Three-dimensional structure of a membrane-containing virus. *Proc. Natl. Acad. Sci. USA* 90, 9095–9099.
30. Cheng, R.H., *et al.*, & Baker, T.S. (1995). Nucleocapsid and glycoprotein organization in an enveloped virus. *Cell* 80, 621–630.
31. Fuller, S.D., Berriman, J.A., Butcher, S.J. & Gowen, B.E. (1995). Low pH induces swiveling of the glycoprotein heterodimers in the Semliki Forest virus spike complex. *Cell* 81, 715–725.
32. Fuller, S.D. (1987). The *T*=4 envelope of Sindbis virus is organized by interactions with a complementary *T*=3 capsid. *Cell* 48, 923–934.
33. Vogel, R.H., Provencher, S.W., von Bonsdorff, C.-H., Adrian, M. & Dubochet, J. (1986). Envelope structure of Semliki Forest virus reconstructed from cryo-electron micrographs. *Nature* 320, 533–535.
34. Argos, P. (1987). Analysis of sequence-similar pentapeptides in unrelated protein tertiary structures. Strategies for protein folding and a guide for site-directed mutagenesis. *J. Mol. Biol.* 197, 331–348.
35. Burmeister, W.P., Ruigrok, R.W.H. & Cusack, S. (1992). A resolution crystal structure of influenza B neuraminidase and its complex with sialic acid. *EMBO J.* 11, 49–56.
36. Rini, J.M., Schulze-Gahmen, U. & Wilson, I.A. (1992). Structural evidence for induced fit as a mechanism for antibody–antigen recognition. *Science* 255, 959–965.
37. Krishnaswamy, S. & Rossmann, M.G. (1990). Structural refinement and analysis of Mengo virus. *J. Mol. Biol.* 211, 803–844.
38. Coombs, K. & Brown, D.T. (1987). Topological organization of Sindbis virus capsid protein in isolated nucleocapsids. *Virus Res.* 7, 131–149.
39. Ulmanen, I., Söderlund, H. & Kääriäinen, L. (1976). Semliki Forest virus capsid protein associates with the 60S ribosomal subunit in infected cells. *J. Virol.* 20, 203–210.
40. Wengler, G. & Wengler, G. (1984). Identification of transfer of viral core protein to cellular ribosomes during the early stages of alphavirus infection. *Virology* 134, 435–442.
41. Froshauer, S., Kartenbeck, J. & Helenius, A. (1988). Alphavirus RNA replicase is located on the cytoplasmic surface of endosomes and lysosomes. *J. Cell Biol.* 107, 2075–2086.
42. Levitt, M. & Perutz, M. F. (1988). Aromatic rings act as hydrogen bond acceptors. *J. Mol. Biol.* 201, 751–754.
43. Thomas, K.A., Smith, G.M., Thomas, T.B. & Feldmann, R.J. (1982). Electronic distributions within protein phenylalanine aromatic rings are reflected by the three-dimensional oxygen atom environments. *Proc. Natl. Acad. Sci. USA* 79, 4843–4847.
44. Reid, K.S.C., Lindley, P.F. & Thornton, J.M. (1985). Sulphur–aromatic interactions in proteins. *FEBS Lett.* 190, 209–213.
45. Burley, S.K. & Petsko, G.A. (1985). Aromatic–aromatic interactions: a mechanism of protein structure stabilization. *Science* 229, 23–28.
46. Singh, J. & Thornton, J.M. (1992). Atlas of Protein Side-Chain Interactions. IRL Press, Oxford, UK.
47. Lee, S., *et al.*, & Kuhn, R. (1996). Identification of a protein binding site on the surface of the alphavirus nucleocapsid and its implication in virus assembly. *Structure* 4, 531–541.
48. Lee, H. & Brown, D.T. (1994). Mutations in an exposed domain of Sindbis virus capsid protein results in the production of noninfectious virions and morphological variants. *Virology* 202, 390–400.
49. Lee, H., Ricker, P.D. & Brown, D.T. (1994). The configuration of Sindbis virus envelope proteins is stabilized by the nucleocapsid protein. *Virology* 204, 471–474.

50. Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**, 387–395.
51. Vihinen, M., Euranto, A., Luostarinen, P. & Nevalainen, O. (1992). MULTICOMP: a program package for multiple sequence comparison. *CABIOS* **8**, 35–38.
52. Bernstein, F.C., *et al.*, & Tasumi, M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.
53. Brooks, B.R., Bruccoleri, R.E., Olafsen, B.D., States, D.J., Swaminathan, S. & Karplus, M. (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217.
54. Baumann, G., Frömmel, C. & Sander, C. (1989). Polarity as a criterion in protein design. *Protein Eng.* **2**, 329–334.
55. Lüthy, R., Bowie, J.U. & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature* **356**, 83–85.
56. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291.
57. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometric features. *Biopolymers* **22**, 2577–2637.
58. Kunkel, T.A., Roberts, J.D. & Zakour, R.A. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* **154**, 367–382.
59. Salminen, A., Wahlberg, J.M., Lobigs, M., Liljeström, P. & Garoff, H. (1992). Membrane fusion process of Semliki Forest virus II: cleavage dependent reorganization of the spike protein complex controls virus entry. *J. Cell Biol.* **116**, 349–357.