Structure *14*, 75–86, January 2006 DOI 10.1016/j.str.2005.09.014

Structure and Interactions at the Viral Surface of the Envelope Protein E1 of Semliki Forest Virus

Alain Roussel,[1,3,4] Julien Lescar,[1,3,5] Marie-Christine Vaney,[1] Gisela Wengler,[2] Gerd Wengler,[2] and Félix A. Rey[1,6,*]
[1] Laboratoire de Virologie Moléculaire and Structurale
UMR 2472/1157 CNRS-INRA and IFR 115
Avenue de la Terrasse
91198 Gif-sur-Yvette Cedex
France
[2] Institut of Virology
University of Giessen
35390 Giessen
Germany

Summary

Semliki Forest virus (SFV) is enveloped by a lipid bilayer enclosed within a glycoprotein cage made by glycoproteins E1 and E2. E1 is responsible for inducing membrane fusion, triggered by exposure to the acidic environment of the endosomes. Acidic pH induces E1/E2 dissociation, allowing E1 to interact with the target membrane, and, at the same time, to rearrange into E1 homotrimers that drive the membrane fusion reaction. We previously reported a preliminary Cα trace of the monomeric E1 glycoprotein ectodomain and its organization on the virus particle. We also reported the 3.3 Å structure of the trimeric, fusogenic conformation of E1. Here, we report the crystal structure of monomeric E1 refined to 3 Å resolution and describe the amino acids involved in contacts in the virion. These results identify the major determinants for the E1/E2 icosahedral shell formation and open the way to rational mutagenesis approaches to shed light on SFV assembly.

Introduction

Semliki Forest virus (SFV) belongs to the alphavirus genus in the *Togaviridae* family of small, enveloped, messenger-sense single-stranded RNA viruses (Weaver et al., 2000). It has been used as a model to unveil the endocytotic entry route used by many pathogenic microorganisms (Helenius et al., 1980). The alphavirus morphology has been studied extensively by using biochemical, genetic, and structural approaches. The virions are icosahedrally symmetric particles with a glycoprotein cage enclosing a roughly spherical viral membrane (Mancini et al., 2000; Paredes et al., 2001; Zhang et al., 2002). Inside the membrane, the nucleocapsid (NC), a second icosahedrally ordered structure matching the symmetry of the outer cage, packs the roughly 12 kb long genomic RNA. The protein components of the virion display specific interactions to ensure proper particle assembly, budding, and maturation during exit from the infected cell, as well as receptor binding, endosomal uptake, and triggering of membrane fusion during the invasion of a new cell (see Schlesinger and Schlesinger, 2001 and references therein for a comprehensive review on alphavirus biology).

The alphavirus structural proteins are translated from a subgenomic messenger RNA with a single open reading frame coding for the polyprotein precursor C-p62-6K-E1, which undergoes cotranslational cleavage to release its four constituent polypeptides. C forms the icosahedral NC core structure underneath the viral membrane, encasing the viral genome; p62 is a precursor of the mature E2 protein, which is responsible for receptor recognition when infecting a new cell; E1 is responsible for triggering fusion of the viral and target cell membranes during entry; and, finally, the small intervening 6K protein is necessary for efficient particle budding (Liljestrom et al., 1991) and is present at low levels in virions (Gaedigk-Nitschko and Schlesinger, 1990; Lusa et al., 1991).

Both p62 and E1 are type I proteins containing a single transmembrane (TM) segment, which, in the case of p62, is followed by a small, cytosolic C-terminal tail. Protein p62 binds to E1—most likely cotranslationally—in the ER and acts as a chaperone for folding E1 in a membrane fusion-competent conformation (Andersson et al., 1997; Garoff et al., 1994). The p62/E1 heterodimer is then transported to the plasma membrane, and, during this process, p62 is cleaved by furin (Zhang et al., 2003), a TGN resident protease, to yield mature E2 and E3 glycoproteins (de Curtis and Simons, 1988; Watson et al., 1991). E3 is derived from the N-terminal end of p62 and remains bound to E2 in virions only for some strains of SFV. Mature E1/E2 heterodimers accumulate at the plasma membrane (Salminen et al., 1992) and interact with the NC cores present in the cytoplasm via the cytosolic tail of E2. This interaction leads to morphogenesis and budding of infectious virions (Skoging et al., 1996).

Cleavage of p62 activates the alphavirus fusogenic potential, yielding a metastable E1/E2 heterodimer that dissociates upon exposure to mildly acidic pH, allowing E1 to bind to the target membrane via its fusion peptide (reviewed by Bron et al., 1993; Kielian et al., 2000). Heterodimer dissociation is followed by a fusogenic conformational change of E1 to form a homotrimer in which the E1 subunits adopt a hairpin conformation, bringing together the fusion peptide and the TM segment at one end of the molecule (Gibbons et al., 2004). The membrane fusion reaction occurs as a cooperative process involving lateral interactions between E1 homotrimers that distort target and viral membranes in such a way that lipid merging is induced (Gibbons et al., 2003, 2004).

E1/E2 heterodimers form trimers that are the building blocks of the glycoprotein cage of the virion (von Bonsdorf and Harrison, 1975, 1978) (also, see the review by Garoff et al., 1994). These $(E1/E2)_3$ complexes make 80

*Correspondence: rey@vms.cnrs-gif.fr
[3] These authors contributed equally to this work.
[4] Present address: Centre de Biophysique Moléculaire, F45071 Orléans Cedex 2, France.
[5] Present address: School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Singapore.
[6] Present address: Virology Department, Institut Pasteur, 25 rue du Dr. Roux, 75015 Paris, France.

trimeric spikes at the virus surface that project from a protein scaffold of icosahedral architecture with triangulation T = 4 (Caspar and Klug, 1962). The TM segments of E1 and E2 cross the membrane together, roughly at the vertices of the triangular base of the spikes. The cytoplasmic tail of E2 makes one-to-one interactions with the NC, which also exhibits a T = 4 icosahedral organization, resulting in a total of 240 identical C/E2 contact sites (Zhang et al., 2002, Lee et al., 1996).

We previously reported the Cα trace of the E1 polypeptide built from a 3.5 Å resolution electron density map calculated from X-ray diffraction data (Lescar et al., 2001). The crystals were anisotropic, and the diffraction was of poor quality and did not allow further refinement of the polypeptide trace at the time. This preliminary model was fitted into a 3D reconstruction at 9 Å resolution of the SFV particle calculated by cryo-EM and image processing (Mancini et al., 2000). The overall arrangement of the E1 molecules in the virus particle was interpreted with the help of this model, and the remainder of the density was attributed to E2 (Lescar et al., 2001). The resulting assignment was independently confirmed by 3D reconstructions of glycosylation-deletion mutants of a related alphavirus, Sindbis virus (Pletnev et al., 2001; Zhang et al., 2002). However, the identity of the amino acid side chains involved in the contacts was not provided because the model was only preliminary. Here, we present the results of the crystallographic refinement of the E1 structure, by using diffraction data partially extending to 3 Å resolution. This model contains all of the atoms of the 384 N-terminal amino acids of the E1 ectodomain (out of 413 residues upstream of the TM segment—i.e., lacking most of the "stem" region of the ectodomain, residues 382–413). Placement of the refined model into the cryo-EM reconstruction now allows for visualization of the E1 residues involved in lateral E1/E1 and E1/E2 contacts that form the external glycoprotein cage of the particle. This model was also used to interpret the higher-resolution 3D reconstruction of Sindbis virus described in the accompanying manuscript (Mukhopadhyay et al., 2006). These results provide a rational basis for site-directed mutagenesis experiments to validate the importance of the interactions described for assembly and particle formation.

Results and Discussion

Structure Determination

The monomeric, soluble ectodomain of E1 (termed "E1-ΔS," containing amino acids 1–390 out of 438 in the intact protein) was obtained by controlled subtilisin treatment of the SFV glycoproteins detergent solubilized from purified virus preparations; it crystallizes in the hexagonal space group $P6_422$ (Wengler et al., 1999). The diffraction pattern is very anisotropic, with some crystals diffracting to about 2.4 Å resolution along the direction of the 6_4 hexagonal axis, but only to 3.7 Å in perpendicular directions. The crystals contain a single E1-ΔS monomer in the asymmetric unit and are generally nonisomorphous to each other. The crystal structure was determined by MIR and MAD methods, followed by multicrystal averaging, as described in Lescar et al. (2001), by using diffraction data from several nonisomorphous crystals with cell parameters differing by 2%–3%. For the crystallographic refinement reported here (Table 1), we chose one crystal in which the anisotropy was less marked, diffracting to 3 Å in the good direction and to about 3.3–3.5 Å in the perpendicular plane. The anisotropy is reflected in the drop in completeness of the diffraction data represented in Table 2. Several controls to independently identify the disulfide bonds of the molecule and the side chains containing a carboxylate group (aspartate and glutamate residues) were carried out (see Figures S1–S3 in the Supplemental Data available with this article online). Finally, the trace was further confirmed by the independently determined 3.3 Å resolution

Table 1. Crystallographic and Refinement Statistics

Data Collection Statistics	
Space group	$P6_422$
Cell parameters (Å, °)	a = 79.38, b = 79.38, c = 335.91, γ = 120
Resolution (Å)	39.0–3.0 (3.21–3.00)[a]
Measured reflections	82,854
Unique reflections	10,666
Redundancy	7.8 (5.5)[a]
Completeness (%)	87.6 (76.1)[a]
I/σ(I)	8.0 (5.4)[a]
R_{merge} (%)	7.5 (12.4)[a]
Refinement Statistics	
R factor (%)	26.6 (38.0)[a]
R_{free} (%)	31.8 (53.7)[a]
Number of amino acid residues	384
Number of water molecules	88
Number of protein non-hydrogen atoms	3023
Rms deviations from ideal value	
Bond lengths (Å)	0.006
Bond angles (°)	1.028
Chirality	0.061
Ramachandran statistics	
Most favored region (%)	78.6
Additional allowed region (%)	20.5
Generously allowed region (%)	0.6
Disallowed region (%)	0.3
$<B\ average>_{all\ atoms}$	
All domains	63.7
DI + DII	54.7
DIII	101.6

R factor = $\Sigma|F_o - F_c|/\Sigma|F_o|$. $R_{free} = \Sigma|F_o - F_c|/\Sigma|F_o|$, calculated using 10% of the data selected randomly from the observed reflections. F_o and F_c are, respectively, the observed and calculated structure factor amplitudes.
[a] Values in parentheses refer to the highest-resolution shell.

Table 2. Statistics per Resolution Shell

Resolution Range (Å)[a]	Completeness (%)	R_{free} Factor (%)
12.90	91.85	0.30
7.66	100.00	0.28
5.97	99.91	0.31
5.06	100.00	0.23
4.47	99.04	0.31
4.05	92.89	0.29
3.72	83.60	0.34
3.47	79.54	0.39
3.24	76.46	0.38
3.07	76.78	0.49

[a] Completeness at the highest-resolution shell reflects the anisotropy of the crystals. The overall resolution range is 39–3 Å, and it was divided in ten resolution shells spanning 1/39 to 1/3 $Å^{-1}$. The value quoted is $1/d_i$, where d_i is the center of the i^{th} resolution shell.