

# Functions of Chikungunya Virus Nonstructural Proteins

Tero Ahola and Andres Merits

## Introduction

The nonstructural proteins (nsPs) of chikungunya virus (CHIKV) and other alphaviruses are encoded at the 5' region of the positive-strand viral RNA genome, and thus translated directly after the entry of the RNA into the cytoplasm, to enable the immediate start of the RNA replication process. Many of the replicative functions remain poorly studied for CHIKV nsPs, but have been earlier characterized for proteins from other alphaviruses, primarily Semliki Forest virus (SFV) and Sindbis virus (SINV; Kaariainen and Ahola 2002). It is expected that the basic biochemical functions are conserved in CHIKV proteins, the amino acid (aa) sequence of which are on average 71 % identical with the more closely related SFV nsPs. Therefore we review data from CHIKV whenever available, but also draw heavily on other alphaviruses, and point out the uncertainties regarding the functions of CHIKV proteins.

Secondly, the nsPs have many functions in host–virus interactions, including the evasion of antiviral responses. Clearly, these functions are more virus- (and/or cell-type) specific in nature, so care needs to be taken in making inferences regarding CHIKV. For instance, it has been shown that hnRNP K, a cellular protein identified as a pro-viral (essential) factor for CHIKV replication (Bourai et al. 2012) acts as a negative effector of replication for SFV (Varjak et al. 2013). A primary division within the mammalian/avian alphaviruses lies between the Old World viruses (e.g., CHIKV, SFV, and SINV) and the New World viruses (e.g., Venezuelan equine encephalitis virus, VEEV), and these two groups display quite different modes of

---

T. Ahola

Department of Food and Environmental Sciences, University of Helsinki, Helsinki, Finland

e-mail: [tero.ahola@helsinki.fi](mailto:tero.ahola@helsinki.fi)

A. Merits (✉)

Institute of Technology, University of Tartu, Tartu, Estonia

e-mail: [andres.merits@ut.ee](mailto:andres.merits@ut.ee)

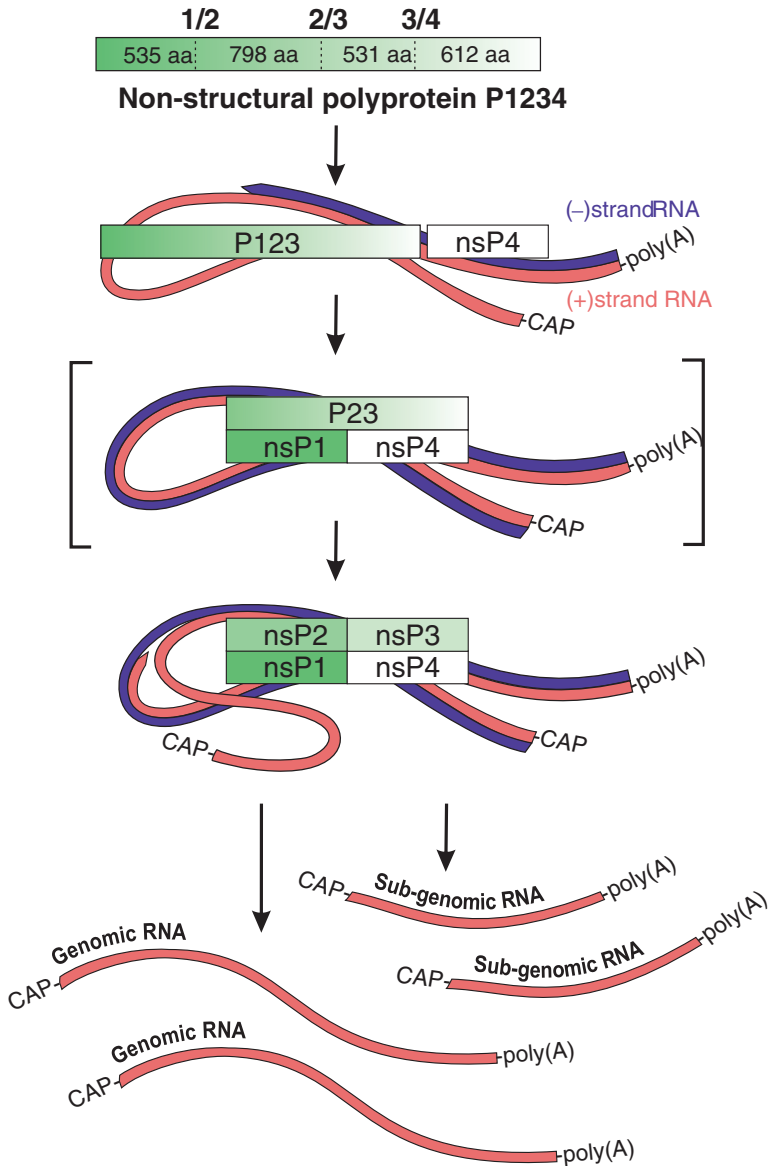
host interaction (Garmashova et al. 2007; Kim et al. 2016). Due to the multiple functions of the nsPs, a complex pattern of protein localization is observed within the cell. A sizable fraction of the nsPs stays together in the replication complexes, but a significant fraction also displays localizations specific to each of the proteins, engaging in various virus–host interactions.

For CHIKV itself, most of the data we have concerning the functions of the ns-proteins originates from studies of proteins encoded by viruses belonging to the East/Central/South African (ECSA) genotype (typically from the Indian Ocean outbreak of 2005–2007). However, two more genotypes of CHIKV, West African (WA) and Asian, are known. Ns-polyproteins, encoded by viruses belonging to different genotypes, have 95 % (WA compared to the other two genotypes) to 97 % (ECSA compared to Asian genotype) identity. The identity is highest in nsP4 and lowest for the C-terminal hypervariable region of nsP3. Three to five percent difference in aa sequence certainly creates a possibility that some of the nonenzymatic functions of nsPs encoded by different CHIKV genotypes may be substantially different. Thus far, no head-to-head comparison of functions of nsPs from different CHIKV genotypes has been performed. However, comparison of data available for WA and ECSA genotype viruses supports the possibility that the nsPs encoded by viruses from these genotypes have substantially different properties in specific host interactions (see below under “nsP2”).

## Expression and Processing of Nonstructural Polyproteins

In alphaviruses, the replication proteins are simply called nonstructural proteins 1–4 (nsP1–4). Similar to many other positive-stand RNA viruses, the replicase proteins of alphaviruses are expressed in the form of ns-polyprotein precursor(s) (Strauss and Strauss 1994). Most clinical isolates of CHIKV encode for two ns-polyproteins, P123 and P1234; the latter results from the read-through of an in-frame opal termination codon located close to the sequence encoding for the C-terminus of nsP3. However, in some CHIKV isolates and laboratory-adapted strains, the opal codon is replaced by a codon for arginine and the corresponding genomes express P1234 as the only ns-polyprotein. Interestingly, the presence or absence of the opal codon seems to have relatively minor effects on the virus, most likely because in natural isolates of the virus both variants (with the terminator and without it) of the genomes are present. As in the other alphaviruses, the expression of CHIKV ns-polyprotein(s) occurs only at the early stages of infection in mammalian cells and is shut down typically 6–8 h post infection (Scholte et al. 2013), due to the general inhibition of cellular translation during infection.

In order to give rise to functional proteins, the processing of ns-polyproteins must be precise and well regulated. CHIKV P1234 contains three cleavage sites hereafter termed 1/2 (site between nsP1 and nsP2), 2/3, and 3/4 sites (Fig. 1). All three sites are processed by the viral protease activity residing in nsP2 or in ns-polyproteins containing nsP2. The protease must specifically recognize each of the cleavage sites, and different affinity towards the three sites is partially responsible



**Fig. 1** Processing of CHIKV ns polyprotein P1234 and RNA synthesis. P1234 is shown at the top with the cleavage sites and sizes of the individual nsPs indicated. The intermediate P123 + nsP4 is responsible for negative-strand synthesis, the nsP1 + P23 + nsP4 complex is very short-lived (shown in brackets), and mature replication complexes producing positive-strand genomic and sub-genomic RNAs consist of fully processed nsPs

for the ordered pattern of cleavages (Lulla et al. 2006). However, recent studies of SFV P1234 processing reveal that a second mechanism, the presentation of cleavage sequences via long-range interactions between different domains of the polyprotein,

has a major role (Lulla et al. 2012, 2013). The three cleavage sites of CHIKV ns-polyproteins are more similar compared to each other than those of SFV, therefore one can assume that for CHIKV the presentation of cleavage sites has the same or even a larger role than for SFV. Current data already indicate that the cleavage patterns of P1234 in CHIKV and SFV are closely similar: in order to allow the formation of functional replicase complexes (see below), the cleavage of the 3/4 site must take place first, followed by cleavage of the 1/2 and finally by cleavage of the 2/3 site (Vasiljeva et al. 2003).

## Replication Cycle

After the cleavage of the 3/4 site and recruitment of the viral genome as template, the partially processed replicase P123 + nsP4 is responsible for negative-strand synthesis (Lemm et al. 1994; Shirako and Strauss 1994; Wang et al. 1994; Fig. 1). Replication complexes cannot be assembled from fully processed nsPs, but instead the polyprotein P123 is an essential intermediate. It is hypothesized that normally each replicase would make only a single negative strand and then switch to positive-strand synthesis, with the newly made negative strand as the template. This switch is connected to the processing of 1/2 and 2/3 sites, and replication complexes active in positive-strand synthesis thus contain a set of fully processed nsPs (Fig. 1). If these two latter cleavages are blocked, positive-strand synthesis still occurs but is very inefficient (Lemm et al. 1994; Shirako and Strauss 1994). The final cleavage of the 2/3 site may further facilitate the synthesis of subgenomic RNAs. CHIKV expresses a single subgenomic RNA, whose synthesis is initiated from the subgenomic promoter located on the negative strand. The subgenomic RNA corresponds to the final third of the genomic RNA, and it encodes the structural proteins of the virus. Both positive-strand RNAs are 5' capped and 3' polyadenylated. This model of RNA synthesis, although attractive and verified with other alphaviruses, has some unsolved conundrums. First, the stoichiometry of functional replication complexes remains unknown, and they may contain multiple copies of some or all of the nsPs or their precursors, as suggested, for example, by the *trans*-cleavage of the 2/3 site (see below under “nsP2”). Secondly, replication is connected with the formation of small membrane invaginations known as spherules as the site of replication (see the concluding section), and therefore the geometry of the replication complexes could be strictly constrained in a limiting space.

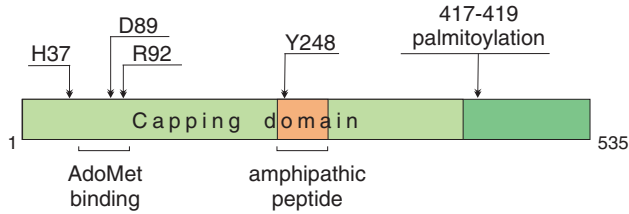
Recently, a robust *in vitro* replication system derived from CHIKV-infected cells has been described, which should be amenable to biochemical analysis, and it is also very useful in characterizing the mode of action of antiviral molecules (Albulescu et al. 2014). A *trans*-replication system for CHIKV, in which a template and the replication proteins are expressed from separate plasmids, following the model of SFV, has also been constructed (Spuul et al. 2011; Utt et al. 2016).

As a consequence of polyprotein processing, only nsP1 has an N-terminal Met residue whereas the N-terminal residues of nsP2, nsP3, and nsP4 of CHIKV are Gly,

Ala, and Tyr, respectively. The native N-termini represent important requirements for the functionality of nsP2, nsP4, and possibly nsP3 in different alphaviruses: the properties of these proteins are usually altered, and some of their functions may even be abolished by minor changes such as the addition or substitution of a single N-terminal aa-residue (Vasiljeva et al. 2003; Shirako and Strauss 1998). The same is clearly the case for CHIKV nsP2 (Das et al. 2014b) and presumably also for nsP3 and nsP4. Most likely this property originates from the fact that nsPs in their mature form have partially different functions and/or are involved in different interactions compared to the corresponding regions in ns-polyproteins. This has important implications for the analysis of the functions of alphavirus nsPs; in order to express the proteins, an initiation codon is often added to the expression constructs of nsP2, nsP3, and nsP4 or the proteins are tagged at their termini. Based on the considerations highlighted above, the data obtained by using native (not tagged or truncated) nsPs should most closely reflect the true functions of the proteins, whereas the use of modified proteins can support misleading conclusions. This does not mean that data obtained using such tools have no value; however, researchers and readers should be aware of the biological effects of such manipulations.

## nsP1

The enzymatic functions of alphavirus nsP1 are needed in the capping of the viral positive-strand genomic and subgenomic RNAs, which bear a cap0 structure  $m^7G(5')ppp(5')ApNp\dots$ . SFV nsP1 was characterized as a unique mRNA capping enzyme, which first transfers a methyl group from *S*-adenosyl-methionine (AdoMet) to position 7 of GTP, and subsequently forms a covalent complex with  $m^7GMP$ , releasing pyrophosphate (Ahola and Kaariainen 1995). Finally, the methylated guanylate residue needs to be transferred to the 5' end of the viral RNAs bearing diphosphate (ppRNA), to complete the cap structure. The first reaction of capping, the triphosphatase step removing the first phosphate of the nascent pppRNA, is carried out by nsP2, as explained below. The reactions catalyzed by nsP1 are unique, inasmuch as cellular and DNA poxvirus-encoded mRNA capping enzymes first transfer GMP to the RNA, and the methylation of the cap is carried out as a final step by a separate cap methyltransferase protein/domain (Decroly et al. 2012). The capping domain related to nsP1 is found in all members of the alphavirus-like superfamily, which encompasses the animal alpha-, rubi-, and hepeviruses as well as dozens of genera of plant viruses, but related proteins have not been found elsewhere (Rozanov et al. 1992). The capping functions of nsP1 are essential for alphavirus replication (Wang et al. 1996). Although the capping activities have been described for SFV and SINV enzymes, given the conservation of function within the entire superfamily, there is no doubt that they are also present in CHIKV nsP1. Due to its unique features, nsP1 is an appealing target for specific antivirals. However, the lack of structural information and the propensity for membrane association (see below) have impeded studies of the protein, and only recently have specific inhibitors been described (Delang et al. 2016).



**Fig. 2** Schematic of the domain structure of CHIKV nsP1. Conserved active site residues and regions important for membrane association have been marked. See the text for details

Recent sequence and secondary structure analyses have demonstrated that the conserved capping domain is a region of  $\geq 400$  aa residues (Ahola and Karlin 2015; Fig. 2). There are only four amino acid residues, which are (almost) universally conserved in all members of the superfamily. A conserved histidine (H37 in CHIKV nsP1) is thought to be the covalent binding site for  $m^7GMP$ , a conserved aspartate (D89) is involved in binding the methyl donor AdoMet, and a conserved arginine (R92) and tyrosine (Y248) are also essential for the methyltransferase function of nsP1. Several other residues that are partially conserved and important for function have also been identified (Ahola and Karlin 2015). It has been speculated, based on secondary structure and very limited sequence conservation, that at least the AdoMet-binding site of nsP1 could be structurally related to the large family of methyltransferases having a so-called Rossmann fold (Ahola et al. 1997). This hypothesis can only be verified by the high-resolution structure of the protein, which is still lacking. The evolutionary origins of the protein remain an interesting mystery, and could even be a case of virus-specific evolution, where an ancestral methyltransferase has also acquired the ability to catalyze the guanylyltransferase reaction.

nsP1 is the sole membrane anchor of the alphavirus replication complex. The protein does not possess transmembrane segments, but instead interacts with membranes in a monotopic fashion, that is, with only the cytoplasmic sheet of the membrane bilayer. The main interaction appears to be mediated by a membrane-binding peptide located in the central part of the protein (approximately residues 244–263 in CHIKV nsP1; Fig. 2), which forms an amphipathic  $\alpha$  helix, as characterized for the corresponding peptide from SFV (Ahola et al. 1999; Lampio et al. 2000). The hydrophobic facet of the helix interacts with the hydrophobic lipid chains, whereas the hydrophilic, predominantly positively charged facet interacts with the head groups of the phospholipids. The peptide and the entire nsP1 have a specific affinity towards membranes bearing a negative charge, carrying, for instance, a high density of phosphatidylserine (Ahola et al. 1999). Several residues of the peptide are essential for the membrane binding of the entire replication complex, and perturbation of membrane association is lethal for the virus (Spuul et al. 2007). Sequence comparisons within related proteins from other superfamily members have sug-

gested that there might be other membrane association sites, but these remain to be tested experimentally (Ahola and Karlin 2015).

nsP1 is also covalently modified by palmitoylation, which serves as a secondary mechanism further tightening the membrane association (Laakkonen et al. 1996). nsP1s from different alphaviruses have 1–3 cysteine residues around amino acid 420, CHIKV having three (aa residues 417–419). Although palmitoylation is important for the normal function of nsP1, it has been possible to select SFV variants that are devoid of palmitoylation, but bear compensatory mutations elsewhere in nsP1 (Zusinaite et al. 2007).

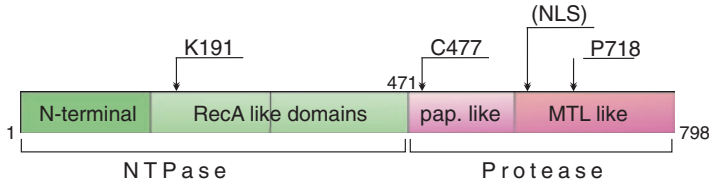
When expressed alone in mammalian cells, nsP1 is completely membrane associated and localizes to the cytoplasmic surface of the plasma membrane. In infected cells, nsP1 also directs the entire replication complex to the membrane (Spuul et al. 2007). In some alphavirus infections (e.g., SINV), the replication complexes remain predominantly at the plasma membrane, whereas in others (e.g., SFV) they are later efficiently internalized and found on the outer surfaces of enlarged endosomes and lysosomes (so-called cytopathic vacuoles type I; Frolova et al. 2010; Spuul et al. 2010). nsP1 does not appear to mediate the internalization process, and processed nsP1 molecules not associated with replication complexes remain on the inner surface of the plasma membrane. CHIKV seems to follow the model in which the replication complexes mostly remain on the plasma membrane in mammalian cells (Thaa et al. 2015).

In the replication complex, nsP1 must naturally interact with the other nsPs, but these interactions remain poorly understood. Nevertheless, there is some evidence that a major interaction is with nsP4 (Zusinaite et al. 2007; Fata et al. 2002; Shirako et al. 2000). Interactions of nsP1 with host proteins are also poorly characterized, but similar to the other nsPs, they are important for viral pathogenesis. Interestingly, CHIKV nsP1 downregulates the expression of and thus counteracts the function of tetherin (also known as BST-2), which is a virus restriction factor retaining virus particles on the cell surface (Jones et al. 2013).

## nsP2

nsP2 is the largest of the ns-proteins (798 aa residues in CHIKV) and is often said to consist of an N-terminal RNA helicase and C-terminal protease domains. Strictly speaking, this is not correct. The three-dimensional structure of the fragment corresponding to the C-terminal part of CHIKV nsP2 (aa residues 471–791) has been resolved at 2.4-Å resolution (Protein Data Bank code 3TRK), revealing a fold highly similar to that of nsP2 protease from VEEV (Russo et al. 2006) and SINV (Shin et al. 2012). This clearly shows that this fragment consists of two structural domains, a papain-like protease domain (aa residues ~471–605), and a remnant of an Ftsj methyltransferase-like (MTL) domain (aa residues ~606–791) that is non-functional as a methyltransferase because of the absence of a number of crucial





**Fig. 3** The domain structure of nsP2, showing the helicase motif I with a conserved lysine and the active site cysteine of the protease. The putative NLS and the conserved proline residue affecting cytotoxicity have been marked

structural elements. These domains function as one unit and they are essential for the protease activity of nsP2. The structure of the N-terminal region of nsP2 has not been resolved. Molecular modeling reveals that it consists of three structural domains, the second and third being RecA-like domains characteristic for RNA helicases. The first N-terminal domain has little homology outside of alphaviruses (Das et al. 2014b). Thus, it is more correct to say that CHIKV nsP2 is a multidomain protein (Fig. 3).

Similar to the other ns-proteins, nsP2 is a multitasking protein having several enzymatic activities and a large array of important nonenzymatic activities. As nsP2 from Old World alphaviruses is central in the suppression of cellular transcriptional and antiviral responses (Garmashova et al. 2006, 2007; Breakwell et al. 2007), CHIKV nsP2 has been studied in more detail than the other CHIKV ns-proteins, and it may already be the best-studied alphaviral nsP2.

*Enzymatic functions.* nsP2 has four known enzymatic activities, all of which have been revealed and analyzed for CHIKV nsP2.

The first function to be analyzed was the protease activity (nsP2 belongs to papain-like Cys-proteases). It was shown that a truncated recombinant protein, corresponding to aa residues 422–639 of nsP2, has no protease activity. In contrast, a protein spanning from aa residue 422 to the end of nsP2 was capable of cleaving short (3–8 aa residues) fluorogenic peptides corresponding to 1/2 and 3/4 sites (Pastorino et al. 2008). The peptides corresponding to the 3/4 site were cleaved more efficiently than peptides corresponding to the 1/2 site; however, this could be attributed to the differential solubility of substrates. Thus, the true cleavage requirements and efficiencies remained unknown. The actual protease activity of CHIKV nsP2 has been recently analyzed using the setup previously developed for the SFV protease (Lulla et al. 2006, 2012; Vasiljeva et al. 2001): using full-length nsP2 with the native N-terminus as an enzyme and recombinant proteins containing cleavage sites of different length, as substrates (Utt et al. 2015). This analysis revealed that substrates representing 1/2 and 3/4 sites (10 aa residues preceding and 5 aa residues following the cleavage position) were cleaved very efficiently but the substrate representing the 2/3 site was not cleaved at all. This site was, however, cleavable in an extended form (10 preceding and 170 following aa residues), similar to what had been previously shown for SFV. Thus, the cleavages of the 2/3 and 3/4 sites of



CHIKV are very similar to those of SFV. In contrast, whereas SFV protease poorly cleaves the 1/2 substrate in *trans*, CHIKV protease does it very efficiently. In cells at early stages of CHIKV infection both 3/4 and 1/2 sites can be (and likely are) cleaved in *cis*, whereas the 2/3 site is always cleaved in *trans*. Nevertheless, the data indicate that in order to maintain the correct sequence and timing of ns-polyprotein processing, CHIKV needs to rely more on high-order regulation mechanisms and less on sequence-specific recognition of cleavage sites (see above).

The NTPase and RNA 5'-triphosphatase activities of CHIKV nsP2 were reported few years later (Karpe et al. 2011). The assays were carried out using a fusion protein consisting of maltose-binding protein followed by a fragment of nsP2 (aa residues 166–630). Such recombinant protein efficiently hydrolyzed GTP, dGTP, ATP, dATP, and to some extent also CTP, dCTP, and UTP substrates; in contrast no hydrolysis of dTTP was detected. The enzyme also removed  $\gamma$ -phosphates from RNA substrates. As this reaction was significantly inhibited by the presence of ATP it was concluded that the NTPase and RNA 5'-triphosphatase activities have a common active site (Karpe et al. 2011). Interestingly, it was found that unlike many studied helicases belonging to superfamily 1 (SF1), the NTPase activity of the fusion protein was not stimulated by RNA or DNA oligonucleotides. This and other open questions were subsequently addressed in experiments using different forms of recombinant nsP2, including full-length nsP2 with the native N-terminus, full-length nsP2 with a hexahistidine tag at its N-terminus and a truncated N-terminal fragment of nsP2 (aa residues 1–470) (Das et al. 2014b). These recombinant proteins were capable of hydrolyzing all the canonical dNTPs (including dTTP) and NTPs with no clear preference for substrate. The NTPase activity was found to depend on the intactness of the N-terminus of the enzyme as well as on the presence of the C-terminal region: compared to modified enzymes, the native nsP2 was five- to sevenfold more active. Finally, it was demonstrated that the NTPase activity of nsP2 is clearly stimulated by both RNA- and DNA oligonucleotides. This stimulation was, however, dependent on the presence of the C-terminal region of nsP2, explaining why this effect was not detected using a fusion protein lacking aa residues 631–798.

Based on the sequence analysis of the N-terminal 470 aa residues, CHIKV nsP2 belongs to the SF1 group of helicases. Nonetheless, no helicase activity was detected for the recombinant protein used by Karpe and coworkers (Karpe et al. 2011) or for the N-terminal fragment of nsP2 (Das et al. 2014b). In the latter study it was found that only full-length nsP2 is capable of acting as a helicase. nsP2 was unable to unwind double-stranded (ds)DNA or RNA duplexes lacking a 5' single-stranded overhang. However, it demonstrated unwinding of dsRNA (containing 12 b or longer 5' single-stranded overhang) in a 5'–3' directionally biased manner. Furthermore, it was found that nsP2 also has RNA strand annealing activity. Both RNA-helicase and RNA-strand annealing activities were dependent on the presence of the C-terminal region of nsP2. The combined results indicate that functional cross-talk between different domains of nsP2 is essential for the enzymatic activities of the protein (Das et al. 2014b).

*Nuclear transport, cytotoxic effects, and interference with antiviral responses.* nsP2 is the only ns-protein, which at least in the Old World alphaviruses is localized both to the cytoplasm and the nucleus of infected cells. This phenomenon has been known for a long time (Peranen et al. 1990); however, the mechanism(s) of how nsP2 enters the nucleus is not fully understood. Neither of the two nuclear localization signals (NLS) predicted for SINV nsP2 affects the nuclear transport of the protein (Frolov et al. 2009). In SFV nsP2, a putative NLS has been identified at the beginning of the MTL domain. Mutation of this sequence indeed abolishes the nuclear transport of nsP2 (Rikkonen et al. 1992), however, this is the case only at 37°C but not at 28°C (Tamm et al. 2008). Therefore it is possible that the region is not acting as a canonical NLS, but instead mutations at this site may alter the conformation of the protein in a temperature-dependent manner. In addition, no NLS can be predicted in the corresponding region of CHIKV nsP2 (hereafter the segment of CHIKV nsP2 corresponding to the NLS of SFV nsP2 is referred to as “putative NLS”). Nevertheless, the nuclear localization of nsP2 is not only well documented but also biologically highly relevant, as it is required (but not sufficient) for the inhibition of cellular antiviral responses (Breakwell et al. 2007) and for the shutdown of host-cell transcription. In SINV and CHIKV it has been shown that the latter occurs due to the ability of nsP2 to induce degradation of Rpb1, a catalytic subunit of cellular RNA polymerase II. Unlike the proteases of picornaviruses, which inhibit cellular translation and transcription by cleavage of host cell proteins, the protease activity of nsP2 is not essential for the suppression of cellular transcription (Bourai et al. 2012; Akhrymuk et al. 2012). The ability to induce cellular transcription shutdown is reduced by mutations in the C-terminal region of CHIKV nsP2 (including Pro718 to Gly mutation; Bourai et al. 2012) and by the corresponding mutation in SINV nsP2 (Akhrymuk et al. 2012). However, the effects of mutations introduced into the NTPase/helicase active site were different for these viruses: in SINV such a mutation strongly reduced the nsP2-dependent degradation of Rpb1 whereas CHIKV nsP2 mostly retained its ability to block host gene expression.

The nuclear localization of CHIKV nsP2 is best studied for CHIKV belonging to the ECSA genotype. Immunofluorescence microscopy reveals that in infected cells nsP2 does not localize uniformly. Instead, a lot of variation is observed: in some cells nsP2 is found mostly in the nucleus whereas in others it is almost exclusively located in the cytoplasm. Most likely, this reflects the time-dependence of nuclear localization. Indeed, the fractionation of infected cells confirmed that at early stages (4 h) post infection the nuclear localization of nsP2 is prominent. In contrast, at later stages (12 h) post infection nsP2 is mostly located in the cytoplasm (Utt et al. 2015). The biological reasons, significance, and mechanisms responsible for this phenomenon are not known.

The ability of nsP2 to induce shutdown of cellular transcription, probably combined with other activities of the protein, makes it highly cytotoxic. The cytotoxic functions of nsP2 can be suppressed by different mutations introduced into the protein. Thus far no single mutation has been described that would completely eliminate the cytotoxic effects of CHIKV nsP2 and simultaneously allow virus replication. However, different groups have reported four combinations of mutations, each

allowing persistent growth of CHIKV replicon RNA in BHK-21 cells. For the ECSA genotype these are combinations of Pro718 to Gly substitution with 5 aa residue insertion between residues 647 and 648 of nsP2 (Pohjala et al. 2011) or with Glu117 to Lys substitution (Utt et al. 2015). For the WA genotype these are Pro718 to Ser substitution combined with double substitutions in the putative NLS region (Lys + Arg649 to Ala + Ala) or with substitution of Asp711 to Gly (Fros et al. 2013). For nsP2 of ECSA genotype, the above-mentioned mutations and their combinations compromise all the known enzymatic activities of nsP2 (Utt et al. 2015); similar data concerning the nsP2 of WA genotype are not available. Consistent with observations made for SINV and SFV (Tamm et al. 2008; Frolov et al. 1999), these mutations, in combination or alone, reduce the levels of viral RNA replication (Utt et al. 2015, 2016; Fros et al. 2013). The general tendency is that mutations having a bigger impact on cytotoxicity also have a more severe effect on RNA replication.

The effect of cytotoxicity reducing mutations on CHIKV replication and nuclear transport of nsP2 is one of the few occasions where data are available for viruses belonging to two different genotypes (ECSA and WA). The significant discrepancy of these results illustrates that information available for one CHIKV genotype does not necessarily apply for the other(s). Thus, Pro 718 to Ser mutation severely inhibits RNA synthesis of WA genotype replicons (Fros et al. 2013) and it has virtually no effect on the RNA synthesis of ECSA genotype replicons (Utt et al. 2015, 2016). Furthermore, in the WA genotype this mutation did not affect the nuclear localization of nsP2, whereas for the ECSA genotype virus the nuclear localization became more pronounced. The double substitution in the putative NLS region (Lys + Arg649 to Ala + Ala) of nsP2 of WA genotype virus prevents the nuclear transport of nsP2 (Fros et al. 2013) whereas insertion of 5 aa residues into the corresponding region of nsP2 of ECSA genotype also destroys the sequence of putative NLS but has absolutely no effect on the nuclear localization of nsP2 (Utt et al. 2015).

Infection of vertebrate cells with wild-type alphaviruses also results in rapid shut-off of cellular translation. Translation of viral genomic RNAs is also inhibited, but translation of subgenomic RNAs remains active (Strauss and Strauss 1994). The phenomenon of translation shut-off has been mostly studied using SINV as a model. It has been shown that inhibition of translation is achieved by both protein kinase R- (PKR-) dependent and PKR-independent mechanisms with the latter being the major pathway (Gorchakov et al. 2004). Importantly, it has been clearly demonstrated that transcriptional and translational shut-offs are independent events (Gorchakov et al. 2005). In order to achieve a noncytotoxic phenotype, the alphavirus replicons must lack the ability to inhibit both cellular transcription and translation. Thus, mutations in CHIKV nsP2 generating noncytotoxic replicons (Utt et al. 2015; Fros et al. 2013) eliminate both of these abilities. This would suggest that nsP2 is important for CHIKV-induced translational shut-off. To date, no nsP2 mutation, resulting in elimination of translational, but not transcriptional, shut-off has been reported for CHIKV. In contrast, such mutations have been described for SINV (Gorchakov et al. 2005; Mayuri et al. 2008). Thus far it is not known how nsP2 mediates translational shut-down, although pull-down experiments with SINV nsP2 have revealed that nsP2 interacts with several proteins involved in translation and

with almost the entire sets of ribosomal proteins composing both subunits (Atasheva et al. 2007). Similarly, association of nsP2 of New World alphavirus VEEV with the ribosomal protein RpS6 has been revealed. It has been demonstrated that VEEV infection results in rapid and dramatic decrease of RpS6 phosphorylation which may be important for alphavirus induced translational shut-off (Montgomery et al. 2006). However, as an alternative explanation, it may be that translational shut-off is mainly a host antiviral response, and the low levels of replication in the noncytotoxic mutants are unable to induce this reaction.

In addition to causing a general shut-off of cellular transcription and translation, CHIKV infection also specifically suppresses antiviral responses mediated by type I/type II interferons (IFN). Such a phenomenon is common for all alphaviruses; for Old World alphaviruses, it has been revealed that this is also caused by nsP2 (Breakwell et al. 2007; Frolova et al. 2002; Nikonov et al. 2013). Using transient expression of nsP2 and replicon vector of CHIKV WA genotype it has been shown that CHIKV is no exception to this rule, and its nsP2 is a potent inhibitor of IFN-induced JAK-STAT signaling. The Pro718 to Ser mutation in CHIKV-nsP2, when introduced into a replicon vector, significantly reduced JAK-STAT inhibition (Fros et al. 2010). However, because in transient expression experiments the same mutation failed to reduce the nuclear translocation of STAT1 (Fros et al. 2013), it is likely that the effect detected in the replicon experiment was mostly, if not exclusively, indirect (due to the reduced RNA replication and nsP2 synthesis). In contrast, nsP2 carrying a mutation in the putative NLS (Lys + Arg649 to Ala + Ala) completely lost the ability to inhibit the nuclear translocation of STAT1. Thus, regardless of the actual mechanism of its nuclear entry, the nuclear localization of nsP2 is strictly required for blocking of the JAK-STAT pathway (Fros et al. 2013). This conclusion is also supported by results obtained using ECSA genotype virus (Utt et al. unpublished). Thus, host shut-off-independent inhibition of IFN signaling is common for different CHIKV genotypes and is likely to have an important role in viral pathogenesis.

*Interaction with other viral and host proteins.* In replication complexes all the ns-proteins of alphaviruses interact with each other directly or indirectly. This has been confirmed by pull-down experiments performed using tagged nsP2, nsP3, or nsP4 (Atasheva et al. 2007; Cristea et al. 2006, 2010; Frolova et al. 2006). There are also internal domain-domain interactions within these large proteins as indicated above for nsP2 (Das et al. 2014b). There is convincing evidence showing a functional interaction between the N-terminal region of nsP2 and the N-terminal domains of nsP3 for SFV (Lulla et al. 2012). Finally, for SINV the 3D-structure of a polyprotein, representing the C-terminal part of nsP2 and the first two domains of nsP3, has been resolved. The revealed structure shows that nsP2 and nsP3 share an extensive interface with 3000 Å<sup>2</sup> of buried surface area. The interface between nsP2 and nsP3 is charged, with the nsP2 surface being mostly basic, whereas nsP3 is generally acidic. The significance of the interaction of these two proteins was highlighted by site-directed mutagenesis (Shin et al. 2012). Interestingly, the major aa residue affecting cytotoxicity (nsP2 Pro718) is located at the interface, and could thus affect the conformation of the polyprotein causing multiple functional alterations.

For CHIKV, the most complete study of viral–host protein interactions has been performed with a yeast two-hybrid system (Bourai et al. 2012). Using full-length nsP2 and seven different fragments of nsP2 as bait Bourai and colleagues identified in total 21 different hits; 15 of them were subsequently positively validated using biochemical assays. Importantly, using nsP2s of SFV and SINV in the same assay, 10 interactions were validated with all three viruses, and 6 more were validated with two out of three viruses. Hits validated for CHIKV included nuclear proteins (hnRNP-K, SRSF3), cytoskeleton components (VIM, TACC3, CEP55, KLC4), regulators of gene transcription (ASCC2, TRIM27, MRFAP1L1, EWSR1, IKZF1, ZBTB43), and host factors involved in protein degradation and/or autophagy (CALCOCO2/NDP52, UBQLN4, RCHY1, WWP1). It was proposed that CHIKV interacts through nsP2 with these cellular functions. However, it was observed that siRNA-mediated knockdown of most of these proteins has little effect on CHIKV replication. The exceptions from this were hnRNP-K and UBQLN4, as knockdown of their expression substantially inhibited CHIKV replication (Bourai et al. 2012).

Cellular processes other than the IFN response also affect CHIKV infection. For example, autophagy (a catabolic cellular process, which sequesters cytosolic components within double-membrane vesicles and targets them for degradation) has been shown to play different important roles in the outcome of CHIKV infection. Interestingly, autophagy plays an antiviral role in mouse cells where it delays caspase-dependent cell death (Joubert et al. 2012), but has a proviral role in CHIKV-infected human cells (Krejchich-Trotot et al. 2011). It was found that this difference can be explained by the fact that human autophagy receptor NDP52, but not its mouse orthologue, interacts with CHIKV nsP2 (and also with SFV and SINV nsP2). The interaction of nsP2 with human NDP52 can be detected by coimmunoprecipitation and yeast two-hybrid system. Furthermore, in infected human cells NDP52 localizes in the vicinity of CHIKV replicase complexes. It was proposed that NDP52–nsP2 interaction in the cytoplasm restricts the nuclear localization of highly cytotoxic nsP2 and delays cell death allowing prolonged replication. As mouse NDP52 does not interact with nsP2, autophagy is antiviral in mouse cells (Judith et al. 2013).

It has been proposed that alphaviruses are sensitive to the effects of the unfolded protein response (UPR), which is triggered by ER stress resulting from the expression of alphavirus glycoproteins (Barry et al. 2010). Therefore it is not surprising that CHIKV has developed a strategy to avoid UPR during infection. In cells infected with CHIKV, the splicing of XBP1 mRNA (a characteristic event in UPR) is incomplete and at a late stage of infection the cells become resistant to treatment with tunicamycin, a potent inducer of UPR. As transient expression of CHIKV glycoproteins triggers UPR, the ability to suppress UPR is associated with the non-structural region. Indeed, it was also shown that transient expression of N-terminally tagged nsP2 (from CHIKV of WA genotype) is sufficient for UPR suppression and that mutations, known to suppress cytotoxic properties of nsP2, make the protein unable to interfere with UPR (Fros et al. 2015).

*Antivirals targeting nsP2.* The proteases of human immunodeficiency virus type 1 and hepatitis C virus have been successfully used as targets for antiviral compounds.

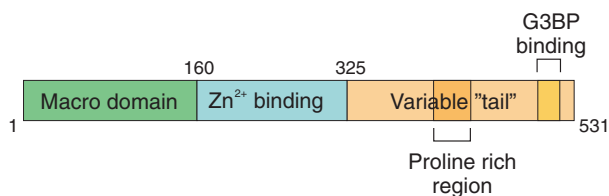
Therefore the highly multifunctional CHIKV nsP2, especially its protease region with known 3D structure, represents an attractive target for the design of antiviral compounds.

In their study, Bassetto and coworkers used the 3D structure of nsP2 protease, obtained by homology modeling, for virtual screening of a large library of commercially available compounds. The initial screen was followed by molecular docking and then by several rounds of screening on CHIKV-infected cells. One of the initial hits was shown to have an  $IC_{50}$  value of  $\sim 5 \mu M$  and some compounds, structurally related to it, had similar  $IC_{50}$  values. The most potent compound, resulting from this multistep screening process, had an  $IC_{50}$  value of  $\sim 3 \mu M$  and a selectivity index of 32 (Bassetto et al. 2013). In another study, carried out by Jadav and coworkers, series of arylalkylidene derivatives of 1,3-thiazolidin-4-one were synthesized and tested in a similar cell culture assay. In this study, the most active compound had an  $IC_{50}$  below  $1 \mu M$  and molecular docking simulation suggested protease inhibition as a possible mode of action (Jadav et al. 2015). Recently, for the first time, a new panel of predicted nsP2 inhibitors was shown to inhibit both the protease activity of nsP2 as well as CHIKV replication in cell culture (Das et al. 2016).

An alternative cell-based assay is taking advantage of the ability of nsP2 to suppress cellular transcription. A screen of 3040 compounds led to identification of a single nontoxic natural compound that inhibited nsP2-mediated cytotoxicity and was also a weak inhibitor ( $IC_{50} = 31 \mu M$ ) of CHIKV replication in cell culture. The mechanism of action of this compound and whether it targets nsP2 or a host component, involved in the development of nsP2-mediated cytotoxicity, is currently unknown (Lucas-Hourani et al. 2013).

## nsP3

Alphavirus nsP3 can be divided into three different domains (Fig. 4). The N-terminal domain of  $\sim 160$  aa is known as the macro domain, and the crystal structure of the CHIKV macro domain has been determined (Malet et al. 2009). The macro domains are named after macrohistones, unusual types of histones, which in addition to the canonical histone domain contain the macro domain. Macro domains are also found in other proteins, and the human genome expresses 11 proteins encoding 1–3 macro domains (Feijs et al. 2013). In RNA viruses, macro domains are expressed by



**Fig. 4** The domain structure of nsP3, and the motifs important for association with host proteins



alphaviruses, rubella virus, hepatitis E virus, and the otherwise unrelated coronaviruses. The biochemical function of most macro domains is to bind mono-ADP-ribose (ADPr) or poly-ADP-ribose (PAR). Mono- or poly-ADP-ribosylation is a post-translational modification of many cellular proteins, and it is catalyzed by a large family of mono-ADP-ribosyltransferases and poly-ADP-ribose polymerases (Feijs et al. 2013; Li et al. 2016), some of which have been implicated in antiviral responses against alphaviruses (Atasheva et al. 2014; Bick et al. 2003). The current consensus is that many macro domains, including the virus-encoded ones, could act as mono-ADP-ribosyl hydrolases, removing the ADP-ribosyl marks on proteins (Feijs et al. 2013; Li et al. 2016). In vitro, the CHIKV macro domain can bind both mono- and poly-ADP-ribose, as well as RNA, and it can also hydrolyze the small substrate analogue ADP-ribose-1''-phosphate (Malet et al. 2009), suggesting that it might also release ADP-ribose bound to proteins. This is an exciting area where more biochemical and proteomic studies are required.

The macro domain is followed by another small domain, found thus far only in alphaviruses, which was defined when the crystal structure of a fragment of SINV replicase was solved. This domain binds a structural Zn<sup>2+</sup> ion, coordinated by four conserved cysteines (Shin et al. 2012). Although this domain, like the macro domain, is essential for alphavirus replication, its precise functions remain to be determined. Both domains are involved as regulators of ns polyprotein processing, but they certainly have additional functions.

These two globular domains are followed by a "tail" region, which is variable in length and sequence between different alphaviruses. In CHIKV, the length of the tail is approximately 205 aa residues. The tail region is heavily phosphorylated on serines and threonines (Vihinen et al. 2001) and it is considered to be unstructured. As is typical for such protein segments, the tail can bind to multiple host proteins, and due to its variation, it could mediate several virus-specific and even cell-type-specific interactions with the host. The tail does tolerate large deletions and insertions, and in several viruses, including CHIKV, it has been used as an insertion site for marker proteins, including fluorescent proteins and *Renilla* luciferase (Pohjala et al. 2011; Kummerer et al. 2012). In spite of the variation, the tail also contains functional elements conserved in several or even all alphaviruses. A conserved proline-rich sequence stretch (PVAPRRRRR) in the tail binds the SH3-domain containing host proteins known as amphiphysin-1 and -2 in CHIKV, SFV, SINV, and most likely in all Old World alphaviruses (Neuvonen et al. 2011). For CHIKV, this interaction has been seen upon nsP3 transfection, but has not been verified in infected cells. For SFV and SINV, the amphiphysins are recruited to viral replication complexes (Neuvonen et al. 2011). They are proteins that can recognize or induce membrane curvature, and thus have been proposed to be structural components of the membranous replication spherules.

The nsP3 of CHIKV and other Old World alphaviruses also directly binds to G3BP1 and G3BP2, which are components of cellular stress granules, dynamic structures involved in arresting RNA translation. Recruitment of G3BPs by nsP3 prevents the formation of stress granules during infection. Instead, nsP3 recruits G3BPs to alternate, presumably virus-induced granules, which are devoid of many stress granule components (Fros et al. 2012; Panas et al. 2014). The binding is medi-

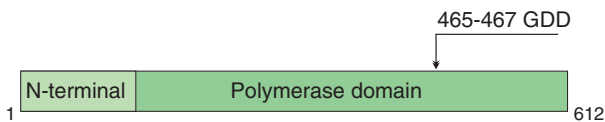


ated by two short, conserved repeat sequences close to the C-terminus of nsP3 (Panas et al. 2014). Therefore, in addition to the replication complexes, nsP3 is prominently localized in granular cytoplasmic structures (Neuvonen et al. 2011). G3BPs have a proviral function (presumably facilitating the switch from ns-polyprotein translation to RNA replication), as depletion of these proteins results in reduced CHIKV replication (Scholte et al. 2015; Kim et al. 2016; Schulte et al. 2016). Additionally, alphavirus nsP3 can interact with several other host proteins, but the significance of these interactions is as yet poorly understood.

Interestingly, recent results indicate that CHIKV nsP3 can be replaced by SFV nsP3 in a chimeric virus, which retains infectivity but becomes temperature-sensitive, only replicating well at 28°C (Merits, unpublished). This seems to reinforce the notion that the basic essential functions of the nsPs are conserved and can be replaced by homologues, but more specific (host) interactions may be altered in different alphaviruses and can be compromised in chimeras.

## nsP4

nsP4 is the catalytic subunit of the alphavirus replicase, the RNA-dependent RNA polymerase (RdRp) (Rubach et al. 2009). Unlike many viral RdRps, the 3D structure of nsP4 remains unknown; however, based on sequence conservation and secondary structure prediction, it has the “right-hand” fold typical for RNA and DNA polymerases. It is clear that nsP4 is assisted in RNA synthesis by the other ns-proteins, which are probably also needed for the correct folding of nsP4. In addition to RdRp activity, nsP4 has terminal adenylyltransferase activity (Tomar et al. 2006), which is likely used for the synthesis of poly(A) tails of positive-strand RNAs in a template-independent fashion. In infected cells, nsP4 is the least abundant ns-protein. This is due to the in-frame terminator codon located upstream of nsP4 coding region (present in most alphaviruses and CHIKV isolates) and to the rapid degradation of individual nsP4 by the proteasome (de Groot et al. 1991). The degradation is induced by the N-terminal Tyr residue of nsP4 which has been shown to be crucial for the enzymatic activity of the protein (Shirako and Strauss 1998). When nsP4 is included in replicase complexes, it becomes stable. The N-terminus of nsP4 consists of 100 aa that could form a partially unstructured region, which precedes the canonical polymerase fold, but this region is essential for the functions of nsP4 (Rupp et al. 2011; Fig. 5).



**Fig. 5** The domain structure of nsP4, with the central polymerase motif Gly-Asp-Asp (GDD) marked

nsP4 is the most conserved protein of alphaviruses and it is very likely that the basic properties of CHIKV nsP4 are similar to the nsP4s of other alphaviruses. Thus far there are no reports of successful expression and purification of functional CHIKV nsP4 as a recombinant protein. Such attempts are hampered by instability of individual nsP4 in eukaryotic cells and/or by insolubility of recombinant nsP4 expressed using the bacterial system. Nevertheless there are convincing, though indirect, pieces of evidence that CHIKV nsP4 indeed acts as RdRp. Mutations in the canonical polymerase active site motif Gly-Asp-Asp (aa residues 465–467) completely abolish CHIKV RNA synthesis (Utt et al. 2016). Furthermore, application of chemical compounds affecting RNA synthesis results in mutations mapping to nsP4. Thus, a mutation associated with resistance to a broad-spectrum viral polymerase inhibitor Favipiravir maps to aa residue 291 of nsP4 (Lys291 to Arg; Delang et al. 2014). Selection of CHIKV mutants in the presence of ribavirin and 5-fluorouracil led to identification of Cys483 to Tyr mutation in nsP4. This mutation confers resistance to the mutagenic effects of these compounds by increasing the replication fidelity of CHIKV. However, this mutation also reduces the genetic diversity of viral progeny and negatively affects CHIKV fitness both in invertebrate and vertebrate hosts (Coffey et al. 2011). Conversely, it has been reported that changing of Cys483 to Ala, Gly, or Trp residue has an opposite effect and results in increased mutation frequency (Rozen-Gagnon et al. 2014).

As indicated above, alphavirus infection results in ER stress and may trigger UPR. It has been reported that during CHIKV infection it is antagonized by nsP2 (see above; Fros et al. 2015). Interestingly, a study by Rathore and coworkers revealed that nsP4 may also have a role in this process. It was observed that the splicing of XBP-1 mRNA is affected in CHIKV infection. Furthermore, it was noticed that compared to SINV, which induces the rapid phosphorylation of translation initiation factor eIF2 $\alpha$  at Ser51, CHIKV possesses mechanism(s) that allow the avoidance of eIF2 $\alpha$  phosphorylation. Transient overexpression of EGFP-nsP4 fusion protein was found to be sufficient to prevent eIF2 $\alpha$  phosphorylation induced by subsequent tunicamycin treatment (Rathore et al. 2013). The mechanism behind this effect is unknown. Similarly, it remains unknown whether nsP4, expressed in natural CHIKV infection, has the same effect.

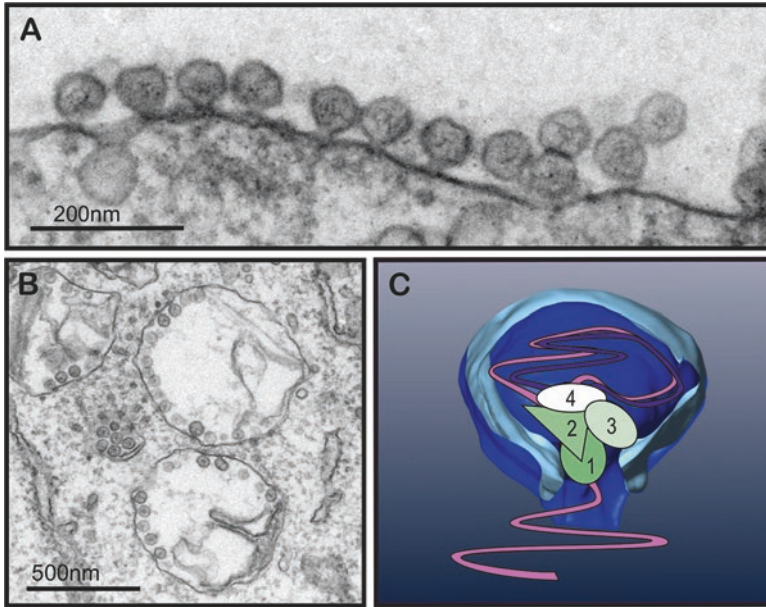
The N-terminal region of nsP4 has been shown to have functional linkage with all the other ns-proteins (Fata et al. 2002; Shirako et al. 2000; Rupp et al. 2011). In contrast, little information is available concerning the cellular proteins interacting with nsP4 in any alphavirus as such studies are hampered by the requirement of an aromatic N-terminal aa-residue, low abundance, and low stability of nsP4. Nevertheless, at least one important cellular interaction partner is known for CHIKV nsP4: heat shock protein 90 alpha subunit (HSP-90 $\alpha$ ). This interaction was first identified using coprecipitation with overexpressed EGFP-nsP4 protein and, importantly, then confirmed by reverse pull-downs using antibodies specific for HSP-90 $\alpha$ . Finally, the functional significance of this interaction was confirmed by siRNA knock-down experiments. Based on these data it was proposed that the interaction between CHIKV

nsP4 and HSP-90 $\alpha$  might be critical for the assembly of the viral replication complex (Rathore et al. 2014). In the same study HSP-90, more specifically HSP-90 $\beta$ , was shown to interact with CHIKV nsP3. Subsequently Das and coauthors reported that HSP90 affects the stability of nsP2 to a larger extent than it affects the stability of nsP3 or nsP4 (Das et al. 2014a). HSP90 was also coprecipitated with an antibody against nsP2 and vice versa. However, as in that study extracts of CHIKV-infected cells (rather than extracts of cells expressing individual nsP2) were used it remains possible that the interaction between nsP2 and HSP90 is mediated by nsP4 and/or nsP3.

## Replication Complex

The RNA replication of all positive-strand RNA viruses takes place in association with cellular membranes (Paul and Bartenschlager 2013; Salonen et al. 2005). The viruses actively modify various membranes to build platforms and protective environments for efficient replication. The alphaviruses generate thousands of small membrane invaginations termed spherules (Spuul et al. 2010). It is thought that one spherule corresponds to one replication complex and each spherule would therefore contain one negative strand. The spherules arise concomitantly with virus RNA synthesis. They are very stable structures, and after being formed, they are active for several hours in producing multiple positive-strand genomic and subgenomic RNAs (Sawicki et al. 2006; Fig. 6). The spherule interior is always connected to the cytoplasm by a narrow neck, allowing the exit of RNAs and the entry of nucleotides. In mammalian cells, the alphavirus spherules are initially formed on the plasma membrane, and in some alphavirus-infected cells they are later endocytosed and found on the cytoplasmic surfaces of endosomes and large, vacuolated lysosomes, termed cytopathic vacuoles type I (Frolova et al. 2010; Spuul et al. 2010; Chen et al. 2013).

Many other viruses also give rise to similar replication spherules, although depending on the virus they have different localizations. For example, the well-characterized spherules of brome mosaic virus are located on endoplasmic reticulum membranes, and the nodavirus spherules reside on the mitochondrial outer membrane (Kopek et al. 2007; Schwartz et al. 2002). The formation of spherules is still poorly understood. For brome mosaic virus, it has been proposed that one of the viral replication proteins would form an internal coating or shell, giving structural support for the spherule (Schwartz et al. 2002). However, in the case of alphaviruses there is no good evidence for the presence of a shell-like structure, and different hypothetical models have been proposed (Kallio et al. 2013). Host proteins must be involved in the formation of spherules. In cells, transient structures resembling spherules are generated by the ESCRT proteins, which are involved in making the internal vesicles on multivesicular endosomal structures. Although ESCRTs may be involved in spherule formation in some viruses (Barajas et al. 2009), there is no evidence that alphaviruses utilize the ESCRTs as host factors. Interestingly, during alphavirus replication the size of the replicating RNAs determines the size of the spherules, whereas in some other viruses the spherule size appears to be fixed by the proteins involved (Kallio et al. 2013).



**Fig. 6** (a, b) CHIKV replication spherules at the plasma membrane and on intracellular vacuoles in infected BHK cells, respectively. (c) Hypothetical model of a spherule, with the dsRNA replicative intermediate inside, and positive strand being released to the cytoplasm through the neck. The nsPs are not to scale, but are shown larger than predicted based on their molecular weight. Thus, multiple copies of some of the nsPs could be present. The location of the proteins (neck, or body of the spherules, or both) is also uncertain

As described above, many of the enzymatic and nonenzymatic activities of the nsPs have been characterized. However, it is still very poorly understood how they function together as a replication machine, coordinating at least the polymerase, helicase, and RNA capping activities. This issue is further complicated by the (likely) multimeric nature of the replicase and the restricted membranous environment in which it acts. Yet another unsolved facet is the interaction with and recruitment of the template to the replication complex. The sequences essential for alphavirus RNA replication are very well defined, and consist of the 5' untranslated region (UTR) of the genome, a structured RNA element (replication enhancer) within the coding sequence for nsP1, and the 3' end of the 3' UTR followed by a poly-A stretch. The subgenomic promoter is only needed for the transcription of the subgenomic RNA but not for genome replication (Hellström et al. 2016). Although the core polymerase nsP4 needs to initiate RNA synthesis, the current results indicate that the presence of the other nsPs in the complex may be required for the recognition of at least some of the promoter elements (Li and Stollar 2004).

In conclusion, the methods to study the individual alphavirus nsPs both as recombinant proteins and in transfected and infected cells have been slowly developed over the last 30 years, and can now be applied to CHIKV. Many further methodological advances are needed in the coming years before the formation and functioning of the replication complex itself can be understood. A second area that still has many

unknown features, are the multiple virus–host interactions involved in antiviral responses, in the viral countermeasures, and in causing the pathogenic outcomes of CHIKV infection.

**Acknowledgments** We warmly thank Katri Kallio for drawing the figures, and Finny Varghese for providing the EM images. This work has been supported by Academy of Finland grant 265997 to TA and by institutional research funding (IUT20-27) from the Estonian Ministry of Education and Research to AM.

## References

- Ahola T, Kaariainen L (1995) Reaction in alphavirus mRNA capping: formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP. *Proc Natl Acad Sci U S A* 92(2):507–511
- Ahola T, Karlin DG (2015) Sequence analysis reveals a conserved extension in the capping enzyme of the alphavirus supergroup, and a homologous domain in nodaviruses. *Biol Direct* 10:16
- Ahola T et al (1997) Critical residues of Semliki Forest virus RNA capping enzyme involved in methyltransferase and guanylyltransferase-like activities. *J Virol* 71(1):392–397
- Ahola T et al (1999) Semliki Forest virus mRNA capping enzyme requires association with anionic membrane phospholipids for activity. *EMBO J* 18(11):3164–3172
- Akhrymuk I, Kulemzin SV, Frolova EI (2012) Evasion of the innate immune response: the Old World alphavirus nsP2 protein induces rapid degradation of Rpb1, a catalytic subunit of RNA polymerase II. *J Virol* 86(13):7180–7191
- Albulescu IC et al (2014) An in vitro assay to study chikungunya virus RNA synthesis and the mode of action of inhibitors. *J Gen Virol* 95(Pt 12):2683–2692
- Atasheva S et al (2007) Development of Sindbis viruses encoding nsP2/GFP chimeric proteins and their application for studying nsP2 functioning. *J Virol* 81(10):5046–5057
- Atasheva S, Frolova EI, Frolov I (2014) Interferon-stimulated poly(ADP-Ribose) polymerases are potent inhibitors of cellular translation and virus replication. *J Virol* 88(4):2116–2130
- Barajas D, Jiang Y, Nagy PD (2009) A unique role for the host ESCRT proteins in replication of Tomato bushy stunt virus. *PLoS Pathog* 5(12), e1000705
- Barry G et al (2010) Semliki forest virus-induced endoplasmic reticulum stress accelerates apoptotic death of mammalian cells. *J Virol* 84(14):7369–7377
- Bassetto M et al (2013) Computer-aided identification, design and synthesis of a novel series of compounds with selective antiviral activity against chikungunya virus. *Antiviral Res* 98(1):12–18
- Bick MJ et al (2003) Expression of the zinc-finger antiviral protein inhibits alphavirus replication. *J Virol* 77(21):11555–11562
- Bourai M et al (2012) Mapping of chikungunya virus interactions with host proteins identified nsP2 as a highly connected viral component. *J Virol* 86(6):3121–3134
- Breakwell L et al (2007) Semliki Forest virus nonstructural protein 2 is involved in suppression of the type I interferon response. *J Virol* 81(16):8677–8684
- Chen KC et al (2013) Comparative analysis of the genome sequences and replication profiles of chikungunya virus isolates within the East, Central and South African (ECSA) lineage. *Virol J* 10:169
- Coffey LL et al (2011) Arbovirus high fidelity variant loses fitness in mosquitoes and mice. *Proc Natl Acad Sci U S A* 108(38):16038–16043
- Cristea IM et al (2006) Tracking and elucidating alphavirus-host protein interactions. *J Biol Chem* 281(40):30269–30278
- Cristea IM et al (2010) Host factors associated with the Sindbis virus RNA-dependent RNA polymerase: role for G3BP1 and G3BP2 in virus replication. *J Virol* 84(13):6720–6732

- Das et al (2014a) Heat shock protein 90 positively regulates Chikungunya virus replication by stabilizing viral non-structural protein nsP2 during infection. *PLoS One* 9(6):e100531
- Das et al (2014b) Functional cross-talk between distant domains of chikungunya virus non-structural protein 2 is decisive for its RNA-modulating activity. *J Biol Chem* 289(9):5635–5653
- Das et al (2016) Design and validation of novel chikungunya virus protease inhibitors. *Antimicrob Agents Chemother* in press
- de Groot RJ et al (1991) Sindbis virus RNA polymerase is degraded by the N-end rule pathway. *Proc Natl Acad Sci U S A* 88(20):8967–8971
- Decroly E et al (2012) Conventional and unconventional mechanisms for capping viral mRNA. *Nat Rev Microbiol* 10(1):51–65
- Delang L et al (2014) Mutations in the chikungunya virus non-structural proteins cause resistance to favipiravir (T-705), a broad-spectrum antiviral. *J Antimicrob Chemother* 69(10):2770–2784
- Delang L et al. (2016) The viral capping enzyme nsP1: a novel target for the inhibition of chikungunya virus infection. *Sci Rep* 6:31819
- Fata CL, Sawicki SG, Sawicki DL (2002) Modification of Asn374 of nsP1 suppresses a Sindbis virus nsP4 minus-strand polymerase mutant. *J Virol* 76(17):8641–8649
- Feijs KL et al (2013) Macrodomein-containing proteins: regulating new intracellular functions of mono(ADP-ribosyl)ation. *Nat Rev Mol Cell Biol* 14(7):443–451
- Frolov I et al (1999) Selection of RNA replicons capable of persistent noncytopathic replication in mammalian cells. *J Virol* 73(5):3854–3865
- Frolov I et al (2009) Random insertion mutagenesis of sindbis virus nonstructural protein 2 and selection of variants incapable of downregulating cellular transcription. *J Virol* 83(18):9031–9044
- Frolova EI et al (2002) Roles of nonstructural protein nsP2 and alpha/beta interferons in determining the outcome of Sindbis virus infection. *J Virol* 76(22):11254–11264
- Frolova E et al (2006) Formation of nsP3-specific protein complexes during Sindbis virus replication. *J Virol* 80(8):4122–4134
- Frolova EI et al (2010) Functional Sindbis virus replicative complexes are formed at the plasma membrane. *J Virol* 84(22):11679–11695
- Fros JJ et al (2010) Chikungunya virus nonstructural protein 2 inhibits type I/III interferon-stimulated JAK-STAT signaling. *J Virol* 84(20):10877–10887
- Fros JJ et al (2012) Chikungunya virus nsP3 blocks stress granule assembly by recruitment of G3BP into cytoplasmic foci. *J Virol* 86(19):10873–10879
- Fros JJ et al (2013) The C-terminal domain of chikungunya virus nsP2 independently governs viral RNA replication, cytopathicity, and inhibition of interferon signaling. *J Virol* 87(18):10394–10400
- Fros JJ et al (2015) Chikungunya virus nsP2-mediated host shut-off disables the unfolded protein response. *J Gen Virol* 96(Pt 3):580–589
- Garmashova N et al (2006) Sindbis virus nonstructural protein nsP2 is cytotoxic and inhibits cellular transcription. *J Virol* 80(12):5686–5696
- Garmashova N et al (2007) The Old World and New World alphaviruses use different virus-specific proteins for induction of transcriptional shutoff. *J Virol* 81(5):2472–2484
- Gorchakov R et al (2004) PKR-dependent and -independent mechanisms are involved in translational shutoff during Sindbis virus infection. *J Virol* 78(16):8455–8467
- Gorchakov R, Frolova E, Frolov I (2005) Inhibition of transcription and translation in Sindbis virus-infected cells. *J Virol* 79(15):9397–9409
- Hellström K et al (2016) Ability of minus strands and modified plus strands to act as templates in Semliki Forest virus RNA replication. *J Gen Virol* 97(6):1395–1407
- Jadav SS et al (2015) Thiazolidone derivatives as inhibitors of chikungunya virus. *Eur J Med Chem* 89:172–178
- Jones PH et al (2013) BST-2/tetherin-mediated restriction of chikungunya (CHIKV) VLP budding is counteracted by CHIKV non-structural protein 1 (nsP1). *Virology* 438(1):37–49



- Joubert PE et al (2012) Chikungunya virus-induced autophagy delays caspase-dependent cell death. *J Exp Med* 209(5):1029–1047
- Judith D et al (2013) Species-specific impact of the autophagy machinery on chikungunya virus infection. *EMBO Rep* 14(6):534–544
- Kaariainen L, Ahola T (2002) Functions of alphavirus nonstructural proteins in RNA replication. *Prog Nucleic Acid Res Mol Biol* 71:187–222
- Kallio K et al (2013) Template RNA length determines the size of replication complex spherules for Semliki Forest virus. *J Virol* 87(16):9125–9134
- Karpe YA, Aher PP, Lole KS (2011) NTPase and 5'-RNA triphosphatase activities of chikungunya virus nsP2 protein. *PLoS One* 6(7), e22336
- Kim DY et al (2016) New World and Old World alphaviruses have evolved to exploit different components of stress granules, FXR and G3BP proteins, for assembly of viral replication complexes. *PLoS Pathog* 12(8):e1005810
- Kopek BG et al (2007) Three-dimensional analysis of a viral RNA replication complex reveals a virus-induced mini-organelle. *PLoS Biol* 5(9), e220
- Krejchich-Trotot P et al (2011) Chikungunya triggers an autophagic process which promotes viral replication. *Virology* 43:432
- Kummerer BM et al (2012) Construction of an infectious chikungunya virus cDNA clone and stable insertion of mCherry reporter genes at two different sites. *J Gen Virol* 93(Pt 9): 1991–1995
- Laakkonen P, Ahola T, Kaariainen L (1996) The effects of palmitoylation on membrane association of Semliki forest virus RNA capping enzyme. *J Biol Chem* 271(45):28567–28571
- Lampio A et al (2000) Membrane binding mechanism of an RNA virus-capping enzyme. *J Biol Chem* 275(48):37853–37859
- Lemm JA et al (1994) Polypeptide requirements for assembly of functional Sindbis virus replication complexes: a model for the temporal regulation of minus- and plus-strand RNA synthesis. *EMBO J* 13(12):2925–2934
- Li C et al (2016) Viral macro domains reverse protein ADP-ribosylation. *J Virol* 90(19): 8478–8486
- Li ML, Stollar V (2004) Identification of the amino acid sequence in Sindbis virus nsP4 that binds to the promoter for the synthesis of the subgenomic RNA. *Proc Natl Acad Sci U S A* 101(25):9429–9434
- Lucas-Hourani M et al (2013) A phenotypic assay to identify chikungunya virus inhibitors targeting the nonstructural protein nsP2. *J Biomol Screen* 18(2):172–179
- Lulla A et al (2006) Molecular determinants of substrate specificity for Semliki Forest virus non-structural protease. *J Virol* 80(11):5413–5422
- Lulla A, Lulla V, Merits A (2012) Macromolecular assembly-driven processing of the 2/3 cleavage site in the alphavirus replicase polyprotein. *J Virol* 86(1):553–565
- Lulla V et al (2013) Presentation overrides specificity: probing the plasticity of alphaviral proteolytic activity through mutational analysis. *J Virol* 87(18):10207–10220
- Malet H et al (2009) The crystal structures of Chikungunya and Venezuelan equine encephalitis virus nsP3 macro domains define a conserved adenosine binding pocket. *J Virol* 83(13): 6534–6545
- Mayuri et al (2008) Role for conserved residues of Sindbis virus nonstructural protein 2 methyltransferase-like domain in regulation of minus-strand synthesis and development of cytopathic infection. *J Virol* 82(15):7284–7297
- Montgomery SA et al (2006) Ribosomal protein S6 associates with alphavirus nonstructural protein 2 and mediates expression from alphavirus messages. *J Virol* 80(15):7729–7739
- Neuvonen M et al (2011) SH3 domain-mediated recruitment of host cell amphiphysins by alphavirus nsP3 promotes viral RNA replication. *PLoS Pathog* 7(11), e1002383
- Nikonov A et al (2013) RIG-I and MDA-5 detection of viral RNA-dependent RNA polymerase activity restricts positive-strand RNA virus replication. *PLoS Pathog* 9(9), e1003610
- Panas MD, Ahola T, McInerney GM (2014) The C-terminal repeat domains of nsP3 from the Old World alphaviruses bind directly to G3BP. *J Virol* 88(10):5888–5893



- Pastorino BA et al (2008) Expression and biochemical characterization of nsP2 cysteine protease of chikungunya virus. *Virus Res* 131(2):293–298
- Paul D, Bartenschlager R (2013) Architecture and biogenesis of plus-strand RNA virus replication factories. *World J Virol* 2(2):32–48
- Peranen J et al (1990) Nuclear localization of Semliki Forest virus-specific nonstructural protein nsP2. *J Virol* 64(5):1888–1896
- Pohjala L et al (2011) Inhibitors of alphavirus entry and replication identified with a stable Chikungunya replicon cell line and virus-based assays. *PLoS One* 6(12), e28923
- Rathore AP, Ng ML, Vasudevan SG (2013) Differential unfolded protein response during Chikungunya and Sindbis virus infection: CHIKV nsP4 suppresses eIF2alpha phosphorylation. *Virol J* 10:36
- Rathore AP et al (2014) Chikungunya virus nsP3 & nsP4 interacts with HSP-90 to promote virus replication: HSP-90 inhibitors reduce CHIKV infection and inflammation in vivo. *Antiviral Res* 103:7–16
- Rikonen M, Peranen J, Kaariainen L (1992) Nuclear and nucleolar targeting signals of Semliki Forest virus nonstructural protein nsP2. *Virology* 189(2):462–473
- Rozanov MN, Koonin EV, Gorbalenya AE (1992) Conservation of the putative methyltransferase domain: a hallmark of the ‘Sindbis-like’ supergroup of positive-strand RNA viruses. *J Gen Virol* 73(Pt 8):2129–2134
- Rozen-Gagnon K et al (2014) Alphavirus mutator variants present host-specific defects and attenuation in mammalian and insect models. *PLoS Pathog* 10(1), e1003877
- Rubach JK et al (2009) Characterization of purified Sindbis virus nsP4 RNA-dependent RNA polymerase activity in vitro. *Virology* 384(1):201–208
- Rupp JC, Jundt N, Hardy RW (2011) Requirement for the amino-terminal domain of sindbis virus nsP4 during virus infection. *J Virol* 85(7):3449–3460
- Russo AT, White MA, Watowich SJ (2006) The crystal structure of the Venezuelan equine encephalitis alphavirus nsP2 protease. *Structure* 14(9):1449–1458
- Salonen A, Ahola T, Kaariainen L (2005) Viral RNA replication in association with cellular membranes. *Curr Top Microbiol Immunol* 285:139–173
- Sawicki DL et al (2006) Role for nsP2 proteins in the cessation of alphavirus minus-strand synthesis by host cells. *J Virol* 80(1):360–371
- Scholte FE et al (2013) Characterization of synthetic chikungunya viruses based on the consensus sequence of recent E1-226V isolates. *PLoS One* 8(8), e71047
- Scholte FE et al (2015) Stress granule components G3BP1 and G3BP2 play a proviral role early in chikungunya virus replication. *J Virol* 89(8):4457–4469
- Schulte T et al (2016) Combined structural, biochemical and cellular evidence demonstrates that both FGDF motifs in alphavirus nsP3 are required for efficient replication. *Open Biol* 6(7):160078
- Schwartz M et al (2002) A positive-strand RNA virus replication complex parallels form and function of retrovirus capsids. *Mol Cell* 9(3):505–514
- Shin G et al (2012) Structural and functional insights into alphavirus polyprotein processing and pathogenesis. *Proc Natl Acad Sci U S A* 109(41):16534–16539
- Shirako Y, Strauss JH (1994) Regulation of Sindbis virus RNA replication: uncleaved P123 and nsP4 function in minus-strand RNA synthesis, whereas cleaved products from P123 are required for efficient plus-strand RNA synthesis. *J Virol* 68(3):1874–1885
- Shirako Y, Strauss JH (1998) Requirement for an aromatic amino acid or histidine at the N terminus of Sindbis virus RNA polymerase. *J Virol* 72(3):2310–2315
- Shirako Y, Strauss EG, Strauss JH (2000) Suppressor mutations that allow sindbis virus RNA polymerase to function with nonaromatic amino acids at the N-terminus: evidence for interaction between nsP1 and nsP4 in minus-strand RNA synthesis. *Virology* 276(1):148–160
- Spuul P et al (2007) Role of the amphipathic peptide of Semliki forest virus replicase protein nsP1 in membrane association and virus replication. *J Virol* 81(2):872–883

- Spuul P et al (2010) Phosphatidylinositol 3-kinase-, actin-, and microtubule-dependent transport of Semliki Forest virus replication complexes from the plasma membrane to modified lysosomes. *J Virol* 84(15):7543–7557
- Spuul P et al (2011) Assembly of alphavirus replication complexes from RNA and protein components in a novel trans-replication system in mammalian cells. *J Virol* 85(10):4739–4751
- Strauss JH, Strauss EG (1994) The alphaviruses: gene expression, replication, and evolution. *Microbiol Rev* 58(3):491–562
- Tamm K, Merits A, Sarand I (2008) Mutations in the nuclear localization signal of nsP2 influencing RNA synthesis, protein expression and cytotoxicity of Semliki Forest virus. *J Gen Virol* 89(Pt 3):676–686
- Thaa B et al (2015) Differential phosphatidylinositol-3-Kinase-Akt-mTOR activation by Semliki Forest and chikungunya viruses is dependent on nsP3 and connected to replication complex internalization. *J Virol* 89(22):11420–11437
- Tomar S et al (2006) Catalytic core of alphavirus nonstructural protein nsP4 possesses terminal adenylyltransferase activity. *J Virol* 80(20):9962–9969
- Utt A et al (2015) Mutations conferring a non-cytotoxic phenotype on chikungunya virus replicons compromise enzymatic properties of non-structural protein 2. *J Virol* 89(6):3145–3162
- Utt A et al (2016) Versatile trans-replication systems for chikungunya virus allow functional analysis and tagging of every replicase protein. *PLoS One* 11(3):e0151616
- Varjak M et al (2013) Magnetic fractionation and proteomic dissection of cellular organelles occupied by the late replication complexes of Semliki Forest virus. *J Virol* 87(18):10295–10312
- Vasiljeva L et al (2001) Site-specific protease activity of the carboxyl-terminal domain of Semliki Forest virus replicase protein nsP2. *J Biol Chem* 276(33):30786–30793
- Vasiljeva L et al (2003) Regulation of the sequential processing of Semliki Forest virus replicase polyprotein. *J Biol Chem* 278(43):41636–41645
- Vihinen H et al (2001) Elimination of phosphorylation sites of Semliki Forest virus replicase protein nsP3. *J Biol Chem* 276(8):5745–5752
- Wang YF, Sawicki SG, Sawicki DL (1994) Alphavirus nsP3 functions to form replication complexes transcribing negative-strand RNA. *J Virol* 68(10):6466–6475
- Wang HL, O'Rear J, Stollar V (1996) Mutagenesis of the Sindbis virus nsP1 protein: effects on methyltransferase activity and viral infectivity. *Virology* 217(2):527–531
- Zusinaite E et al (2007) Mutations at the palmitoylation site of non-structural protein nsP1 of Semliki Forest virus attenuate virus replication and cause accumulation of compensatory mutations. *J Gen Virol* 88(Pt 7):1977–1985