Virus Research 234 (2017) 44–57

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Review

Alphavirus polymerase and RNA replication

Maija K. Pietilä [1], Kirsi Hellström [1], Tero Ahola *

Department of Food and Environmental Sciences, University of Helsinki, Viikinkaari 9 (PO Box 56), 00014 Helsinki, Finland

ARTICLE INFO

Article history:
Received 21 October 2016
Received in revised form 5 January 2017
Accepted 9 January 2017
Available online 16 January 2017

Keywords:
Chikungunya virus
Replication complex
Membrane
Antiviral
Host factor
In vitro RNA synthesis

ABSTRACT

Alphaviruses are typically arthropod-borne, and many are important pathogens such as chikungunya virus. Alphaviruses encode four nonstructural proteins (nsP1–4), initially produced as a polyprotein P1234. nsP4 is the core RNA-dependent RNA polymerase but all four nsPs are required for RNA synthesis. The early replication complex (RC) formed by the polyprotein P123 and nsP4 synthesizes minus RNA strands, and the late RC composed of fully processed nsP1–nsP4 is responsible for the production of genomic and subgenomic plus strands. Different parts of nsP4 recognize the promoters for minus and plus strands but the binding also requires the other nsPs. The alphavirus polymerase has been purified and is capable of *de novo* RNA synthesis only in the presence of the other nsPs. The purified nsP4 also has terminal adenylyltransferase activity, which may generate the poly(A) tail at the 3′ end of the genome. Membrane association of the nsPs is vital for replication, and alphaviruses induce membrane invaginations called spherules, which form a microenvironment for RNA synthesis by concentrating replication components and protecting double-stranded RNA intermediates. The RCs isolated as crude membrane preparations are active in RNA synthesis *in vitro*, but high-resolution structure of the RC has not been achieved, and thus the arrangement of viral and possible host components remains unknown. For some alphaviruses, Ras-GTPase-activating protein (Src-homology 3 (SH3) domain)-binding proteins (G3BPs) and amphiphysins have been shown to be essential for RNA replication and are present in the RCs. Host factors offer an additional target for antivirals, as only few alphavirus polymerase inhibitors have been described.

Contents

Abbreviations: CPC, cyclopentenylcytosine; CPV, cytopathic vacuole; CSE, conserved sequence element; DI, defective interfering; dsRNA, double-stranded RNA; FXR, fragile X syndrome family proteins; G3BP, Ras-GTPase-activating protein (SH3 domain)-binding protein; hnRNP, heterogeneous nuclear ribonucleoprotein; ns, nonstructural; nsP, nonstructural protein; PABP, poly(A)-binding protein; PFZ, pyrazofurin; PI3K, phosphatidylinositol-3-kinase; RC, replication complex; RdRp, RNA-dependent RNA polymerase; RF, replicative form; RI, replicative intermediate; SH3, Src-homology 3; ssRNA, single-stranded RNA; TATase, terminal adenylyltransferase; ts, temperature-sensitive; UTR, untranslated region.

* Corresponding author.
E-mail address: tero.ahola@helsinki.fi (T. Ahola).
[1] Equal contribution.

http://dx.doi.org/10.1016/j.virusres.2017.01.007
0168-1702/

1. Alphavirus genome structure and replicase proteins

Alphaviruses belong to the family *Togaviridae* together with the genus *Rubivirus*, which contains only one member, rubella virus (Table 1). Most alphaviruses are arthropod-borne viruses, and several are important human and/or animal pathogens, causing either fever, rash and arthritis (Old World alphaviruses, including chikungunya virus (CHIKV), Ross River virus (RRV) and Sindbis virus (SINV)), or encephalitis (New World alphaviruses, *e.g.* Venezuelan equine encephalitis virus (VEEV)). The alphavirus genome is a single positive-strand RNA of ~11–12 kb with a 5′ cap0 structure and 3′ poly(A). The two open reading frames encode the nonstructural (ns) or replicase polyprotein, and the structural polyprotein expressed via a subgenomic RNA (Fig. 1) (Strauss and Strauss, 1994). The ns polyprotein is processed in a highly regulated sequence to four final products, the nonstructural proteins (nsPs) 1–4 (Kääriäinen and Ahola, 2002). nsP1 is the viral capping enzyme and membrane anchor of the replication complex (RC) (Ahola and Kääriäinen, 1995; Spuul et al., 2007), while nsP2 is an RNA helicase and the protease responsible for the ns polyprotein processing (Das et al., 2014; Hardy and Strauss, 1989; Vasiljeva et al., 2003). nsP3 interacts with several host proteins and may modulate protein poly- and mono-ADP-ribosylation (Kim et al., 2016; Li et al., 2016), and nsP4 is the core viral RNA-dependent RNA polymerase (RdRp) (Rubach et al., 2009). Structural information is only available for the protease region of nsP2 and the folded N-terminal portions of nsP3 (Shin et al., 2012).

The nsP4 polymerase is the most highly conserved protein in alphaviruses. Even nsP4s of the most diverged alphaviruses, the salmonid alphaviruses, are ≥50% identical in amino acid sequence with the other alphaviral nsP4s (Forrester et al., 2012; Weston et al., 2002). Within the mammalian/avian alphaviruses, the identities are even higher, *e.g* CHIKV nsP4 identity varying from 71% (with Barmah Forest virus, BFV) to 91% (with O'nyong-nyong virus, ONNV) (Khan et al., 2002). Alphavirus nsP4 is ~610 amino acids in length, containing a large C-terminal domain related to other viral RdRps, and an alphavirus-specific ~150 amino acid N-terminal domain. The N-terminal domain is crucial for virus replication. It may be partially disordered structurally and it appears to interact with the other nsPs in the RCs (Rupp et al., 2011). Beyond the family *Togaviridae*, the alphaviruses belong to the large alphavirus-like superfamily that contains the animal hepeviruses and numerous genera of plant viruses (Koonin et al., 2015). All of these possess within their replicase proteins domains related to the capping enzyme nsP1, the helicase domain of nsP2 and the polymerase nsP4. When the predicted secondary structures of the RdRps from bromo- and tobamoviruses were compared to the partial crystal structure of poliovirus polymerase, these alphavirus superfamily members showed the typical RdRp structure with fingers, palm containing the GDD motif, and thumb domains (O'Reilly and Kao, 1998). In addition, comparison revealed a region preceding the fingers domain, which is unique to the RdRps and might be essential for the oligomerization of the polymerase. There are no structures available for the alphavirus nsP4, nor for any of the polymerases within the alphavirus-like superfamily. Now that structures have been solved for some of the negative-strand RNA virus polymerases (Pflug et al., 2017), the alphavirus superfamily may be the most significant branch of RdRps entirely lacking structural information. The biochemical characterization of nsP4 has also been challenging, as discussed in the next section. Therefore, much of this review will focus on the activities and properties of the alphavirus RC as a whole.

2. Biochemical characteristics of nsP4: RNA synthesis and polyadenylation

Most of the work on the replication and RNA synthesis of alphaviruses has been done with SINV and Semliki Forest virus (SFV) (Rupp et al., 2015). Together with the distant sequence relationship to other RdRps, analysis of temperature-sensitive (ts) mutants defective in RNA synthesis indicated that nsP4 is the catalytically active core polymerase subunit (Barton et al., 1988; Hahn et al., 1989a; Sawicki et al., 1990).

SINV mutants ts6 and ts110 each have a single base substitution in nsP4 causing glycine to glutamic acid change at position 153 or 324, respectively. These substitutions are located within highly conserved regions of nsP4 (Hahn et al., 1989a). *In vitro* RNA synthesis of ts6 shows that the RCs are stable at nonpermissive temperature but fail to elongate RNA strands indicating that the elongation capacity of the polymerase is inactivated (Barton et al., 1988). The *in vitro* RNA synthesis of the RCs is reactivated when they are returned to the permissive temperature.

SINV nsP4 has been expressed in *E. coli* and purified resulting in the full-length polymerase with an authentic N-terminal tyrosine and *de novo* RNA-synthesis activity but only when supplied with the polyprotein P123 (Rubach et al., 2009). Remarkably, the purified nsP4 is capable of forming the RCs with P123 resulting in the synthesis of discrete template-length minus strands from the provided plus-strand templates. Furthermore, nsP4 produced in bacteria has the same template requirements as the mammalian nsP4. The core catalytic domain (Δ97nsP4, in which the N-terminal 97 amino acids are deleted) has also been expressed in *E. coli* and purified as a monomer (Tomar et al., 2006). Interestingly, Δ97nsP4 lacks *de novo* copying activity, even when combined with the polyprotein P123 (Rubach et al., 2009; Tomar et al., 2006). Thus, the 97 N-terminal residues seem to be crucial for the RdRp activity. The polyprotein P123 might be required in the template recognition or it may activate nsP4 through protein-protein interactions.

It is intriguing that nsP4 has also been shown to synthesize RNA *in vitro* in the absence of the other viral nsPs (Thal et al., 2007). SINV nsP4 was purified using detergent solubilisation of the membrane fraction from cells expressing uncleavable P123 and nsP4.

Table 1
Summary of the viruses discussed in this review.

Genus/Family	Virus species	Abbreviation
Alphavirus/Togaviridae	Barmah Forest virus	BFV
	Chikungunya virus	CHIKV
	Eastern equine encephalitis virus	EEEV
	O'nyong-nyong virus	ONNV
	Ross River virus	RRV
	Semliki Forest virus	SFV
	Sindbis virus	SINV
	Venezuelan equine encephalitis virus	VEEV
	Western equine encephalitis virus	WEEV
Rubivirus/Togaviridae	Rubella virus	RUBV
Bromovirus/Bromoviridae	Brome mosaic virus	BMV
Alphanodavirus/Nodaviridae	Flock House virus	FHV