Virus Research 234 (2017) 58–73

Contents lists available at ScienceDirect

Virus Research

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

Review

Nidovirus RNA polymerases: Complex enzymes handling exceptional RNA genomes

Clara C. Posthuma[a,*,1], Aartjan J.W. te Velthuis[b,c,1], Eric J. Snijder[a,1]

[a] Molecular Virology Laboratory, Department of Medical Microbiology, Leiden University Medical Center, Leiden, the Netherlands
[b] Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, United Kingdom
[c] Clarendon Laboratory, Department of Physics, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom

ARTICLE INFO

Article history:
Received 9 December 2016
Received in revised form 24 January 2017
Accepted 26 January 2017
Available online 6 February 2017

Keywords:
Coronavirus
Arterivirus
Replication and transcription complex
Polymerase fidelity
Processivity factors
Subgenomic mRNA synthesis

ABSTRACT

Coronaviruses and arteriviruses are distantly related human and animal pathogens that belong to the order *Nidovirales*. Nidoviruses are characterized by their polycistronic plus-stranded RNA genome, the production of subgenomic mRNAs and the conservation of a specific array of replicase domains, including key RNA-synthesizing enzymes. Coronaviruses (26–34 kilobases) have the largest known RNA genomes and their replication presumably requires a processive RNA-dependent RNA polymerase (RdRp) and enzymatic functions that suppress the consequences of the typically high error rate of viral RdRps. The arteriviruses have significantly smaller genomes and form an intriguing package with the coronaviruses to analyse viral RdRp evolution and function. The RdRp domain of nidoviruses resides in a cleavage product of the replicase polyprotein named non-structural protein (nsp) 12 in coronaviruses and nsp9 in arteriviruses. In all nidoviruses, the C-terminal RdRp domain is linked to a conserved N-terminal domain, which has been coined NiRAN (nidovirus RdRp-associated nucleotidyl transferase). Although no structural information is available, the functional characterization of the nidovirus RdRp and the larger enzyme complex of which it is part, has progressed significantly over the past decade. In coronaviruses several smaller, non-enzymatic nsps were characterized that direct RdRp function, while a 3′-to-5′ exoribonuclease activity in nsp14 was implicated in fidelity. In arteriviruses, the nsp1 subunit was found to maintain the balance between genome replication and subgenomic mRNA production. Understanding RdRp behaviour and interactions during RNA synthesis and subsequent processing will be key to rationalising the evolutionary success of nidoviruses and the development of antiviral strategies.

Contents

Abbreviations: 5-AZC, 5-azacytidine; 5-FU, 5-fluorouracil; AV, arterivirus; CBV3, coxsackievirus B3; CoV, coronavirus; EAV, equine arteritis virus; EM, electron microscopy; ExoN, exoribonuclease; FCoV, feline coronavirus; FMDV, foot and mouth disease virus; HCoV-229E, human coronavirus 229E; IMPDH, inosine-5′-monophosphate dehydrogenase; IVRA, *in vitro* RNA synthesis assay; kb, kilobases; M2H, mammalian 2-hybrid; MERS, Middle East respiratory syndrome; MHV, murine hepatitis coronavirus; NiRAN, nidovirus RdRp-associated nucleotidyltransferase; nsp, non-structural protein; ORF, open reading frame; PABP, Poly(A)-binding protein; PEDV, porcine epidemic diarrhea virus; PNS, post-nuclear supernatant; pp, polyprotein; PRRSV, porcine reproductive and respiratory syndrome virus; RBV, ribavirin; RdRp, RNA dependent RNA polymerase; RTC, replication and transcription complex; SARS, severe acute resporiatory syndrome; sg, subgenomic; ss, single-stranded; TGEV, transmissible gastroenteritis virus; TRS, transcription-regulatory sequence; ub, ubiquitin; wt, wild-type; Y2H, yeast 2-hybrid; ZBD, zinc-binding domain.

* Corresponding author.
E-mail address: c.c.posthuma@lumc.nl (C.C. Posthuma).
[1] These authors contributed equally.

http://dx.doi.org/10.1016/j.virusres.2017.01.023

1. Introduction

Positive-stranded RNA (+RNA) viruses that belong to the order *Nidovirales* infect a wide range of vertebrates (families *Arteriviridae* and *Coronaviridae*) or invertebrates (*Mesoniviridae* and *Roniviridae*) (de Groot et al., 2012; Lauber et al., 2012) and can have a significant economic and societal impact. For example, infections with the arterivirus (AV) porcine reproductive and respiratory syndrome virus (PRRSV) have severely affected the swine industry for almost three decades now (Holtkamp et al., 2013; Perez et al., 2015; Pileri and Mateu, 2016), whereas zoonotic coronaviruses (CoVs) have caused episodes of severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS) in humans (Graham et al., 2013; Perlman and Netland, 2009) and may do so again (Menachery et al., 2015). Animal CoVs continue to emerge and cause great economic losses, as exemplified by the recent outbreaks of porcine deltacoronavirus and the porcine epidemic diarrhea virus (PEDV) in China and the United States (Choudhury et al., 2016; Weng et al., 2016; Zhang, 2016). Genetically, the nidoviruses constitute a monophyletic group that is characterized by common ancestry of their key replicative enzymes and associated similarities in genome organization and expression (Fig. 1A) (de Groot et al., 2012; Lauber et al., 2013). However, nidovirus genome sizes vary significantly, with AV genomes ranging from 13 to 16 kilobases (kb), mesonivirus genomes from 20 to 21 kb, and CoV genomes from 26 to 34 kb (Lauber et al., 2013). It has been postulated that this size variation reflects a long history of gradual genome expansion, during which the different nidovirus lineages adapted to their specific niches by acquiring a range of novel functions, encoded by domains that were either incorporated as additional genes or integrated into the large nidovirus replicase gene (Lauber et al., 2013).

In all nidoviruses, at least two-thirds of the capacity of the polycistronic genome is occupied by the two large open reading frames (ORFs; 1a and 1b) that together constitute the replicase gene (Fig. 1A). Both ORFs are translated directly from the viral genome and briefly overlap where a −1 ribosomal frameshift directs the expression of ORF1b to facilitate the formation of an ORF1ab-encoded polyprotein (pp1ab). Cleavage of the pp1a and pp1ab polyproteins by multiple intrinsic protease activities, in combination with −1 and −2 frameshifting in the nsp2 coding region in most arteriviruses with the exception of equine arteritis virus (EAV), results in the production of 13 to 17 non-structural proteins (nsps) (Fang et al., 2012; Li et al., 2014b; Snijder et al., 2013; Ziebuhr et al., 2000). The common ancestry of the extremely diverged nidovirus lineages is primarily reflected in the conservation of an array of 'core replicase domains' (Gorbalenya et al., 2006; Lauber et al., 2013; Snijder et al., 2016), which is composed of two *trans*-membrane proteins, the viral main protease, and – encoded downstream of the ORF1a/1b frameshift site – the RNA-dependent RNA polymerase (RdRp)- and helicase-containing subunits (Fig. 1B), with the canonical RdRp domain residing in CoV nsp12 and AV nsp9. Nidovirus genomes are thought to have expanded gradually by gene duplication and the acquisition of novel domains (Lauber et al., 2013). During this process, specific innovations may have enabled them to explore an unprecedented evolutionary space and adapt to a wide variety of host organisms, including mammals, birds, fish, insects and crustaceans. In particular, the general genome size restrictions of RNA viruses, commonly attributed to the poor nucleotide incorporation fidelity of the viral RdRp domain, may have been mitigated by the acquisition of compensatory enzymatic functions (Deng et al., 2014; Eckerle et al., 2010; Eckerle et al., 2007; Gorbalenya et al., 2006; Snijder et al., 2003). Consequently, some nidovirus nsps contain a unique set of activities not seen in other +RNA viruses (discussed *e.g.* in Sections 8 and 11 in more detail). Imaging and biochemical characterization of nidovirus nsps have shown that they are targeted to specific virus-induced membrane structures (reviewed in (Hagemeijer et al., 2012; Neuman et al., 2014; van der Hoeven et al., 2016)) where they assemble into a so-called replication and transcription complex (RTC; see (Neuman et al., 2014; Snijder et al., 2016; Subissi et al., 2014a) for reviews).

The RNA-templated synthesis of new RNA by the viral RNA polymerase is arguably the key step in the infection cycle of all RNA viruses. In the case of nidoviruses and their polycistronic genomes, RNA synthesis entails not only genome amplification but also the synthesis of a nested set of subgenomic (sg) mRNAs (Fig. 1A; reviewed in (Pasternak et al., 2006; Sawicki et al., 2007; Sola et al., 2011). The sg mRNAs serve to make the genes downstream of the nidovirus replicase ORFs 1a and 1b accessible for translation. These ORFs encode structural and so-called 'accessory' proteins, which are often dispensable for replication *in vitro*, but important for *e.g.* immune evasion and pathogenesis *in vivo* (Liu et al., 2014; Weiss and Leibowitz, 2011). For the purpose of this review, we will refer to the process of sg mRNA synthesis as 'transcription'; the underlying mechanisms will be discussed in more detail in Section 10.

Each nidoviral sg mRNA is produced from a complementary subgenome-length template. Minus strand RNA synthesis can be either continuous (producing a full-length minus-strand template for genome replication) or discontinuous to produce the subgenome-length templates for the production of the sg mRNAs (Sawicki and Sawicki, 1995; Sethna et al., 1989). In addition to the overall RNA structure of the genome and transcription-specific protein factors, conserved transcription-regulatory sequences (TRSs) in the genomic template are thought to be the prime trigger and guiding elements that make the nidovirus RTC synthesize a subgenome-length rather than a full-length minus strand. Thus, the TRSs constitute one class of *cis*-acting RNA signals with which the nidovirus RTC needs to interact, although we note that this feature has mainly been addressed from the RNA side (*e.g.*, by site-directed mutagenesis of TRSs) (Dufour et al., 2011; Pasternak et al., 2001; Zheng et al., 2014; Zuniga et al., 2004). The same essentially applies to the major *cis*-acting RNA elements that constitute the initiation sites for minus and plus strand RNA synthesis near the 3′ end of the plus and minus strand, respectively. Indeed, in both CoV and AV genomes, a number of primary and higher-order structural features have been identified and studied, including sev-