Antiviral Research 143 (2017) 38–47

Contents lists available at ScienceDirect

Antiviral Research

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

Chikungunya virus nsP4 RNA-dependent RNA polymerase core domain displays detergent-sensitive primer extension and terminal adenylyltransferase activities

Ming Wei Chen [a, 1], Yaw Bia Tan [a], Jie Zheng [b, 2], Yongqian Zhao [a], Bee Ting Lim [b, c], Tobias Cornvik [b, c], Julien Lescar [b], Lisa Fong Poh Ng [d], Dahai Luo [a, *]

[a] Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore
[b] School of Biological Sciences, Nanyang Technological University, Singapore
[c] Protein Production Platform, Nanyang Technological University, Singapore
[d] Singapore Immunology Network (SIgN), Agency for Science, Technology and Research (A*STAR), Singapore

ARTICLE INFO

Article history:
Received 17 November 2016
Accepted 4 April 2017
Available online 5 April 2017

Keywords:
Chikungunya
RdRP
HDX-MS
Fluorescence anisotropy
Homology modelling

ABSTRACT

Chikungunya virus (CHIKV) is an important arboviral infectious agent in tropical and subtropical regions, often causing persistent and debilitating disease. The viral enzyme non-structural protein 4 (nsP4), as RNA-dependent RNA polymerase (RdRP), catalyzes the formation of negative-sense, genomic and subgenomic viral RNAs. Here we report a truncated nsP4 construct that is soluble, stable and purified recombinantly from *Escherichia coli*. Sequence analyses and homology modelling indicate that all necessary RdRP elements are included. Hydrogen/deuterium exchange with mass spectrometry was used to analyze solvent accessibility and flexibility of subdomains. Fluorophore-conjugated RNA ligands were designed and screened by using fluorescence anisotropy to select a suitable substrate for RdRP assays. Assay trials revealed that nsP4 core domain is conditionally active upon choice of detergent species, and carries out both primed extension and terminal adenylyltransferase activities. The polymerization assay can be further developed to screen for antiviral compounds *in vitro*.

1. Introduction

Chikungunya virus is one of the most important emerging RNA viruses today (Burt et al., 2012). CHIKV belongs to the Alphavirus genus under the family Togaviridae (Schwartz and Albert, 2010; Strauss and Strauss, 1994); the mosquito species *Aedes aegypti* and *Aedes albopictus* serve as its vector while human is the ultimate host. CHIKV infection results in a febrile disease known as chikungunya fever (Borgherini et al., 2008; Mavalankar et al., 2007). Though rarely fatal in healthy adults, the infection often causes long-lasting joint pain and arthritis, possibly because of latent infection (Borgherini et al., 2008; Hoarau et al., 2010; Labadie et al., 2010; Levine et al., 1994). The increased prevalence of disease, the current lack of specific treatment and vaccine, and the severity of associated long-term sequelae make CHIKV a threat to global healthcare, with high socioeconomic costs (Couturier et al., 2012; Mavalankar et al., 2007; Oon and Ng, 2014; Queyriaux et al., 2008). Co-circulation of dengue, Zika and Chikungunya viruses in endemic areas resulting in co-infection (Villamil-Gómez et al., 2016) and adaptation of CHIKV to new mosquito hosts (Ledermann et al., 2014) may further elevate the crisis.

CHIKV contains one copy of single-chain RNA genome, just below 12 kb (Rupp et al., 2015). Upon entry into host cell, the RNA genome is released into the cytoplasm and acts as the messenger RNA for the non-structural polyprotein P1234, which is the precursor to the CHIKV replication complex (RC). The RC matures as the polyprotein is cleaved sequentially, eventually becoming four non-structural proteins (nsP1, nsP2, nsP3 and nsP4). Within the RC, nsP4 encodes the RNA-dependent RNA polymerase (RdRP), which is responsible for replicating viral RNA (Tomar et al., 2006). An intriguing feature of alphaviral nsP4 proteins is the presence of an N-terminal region (approximately 100 residues) before the core

* Corresponding author. 59 Nanyang Drive, Experimental Medicine Building, 03-07-01, Singapore 636921, Singapore.
E-mail address: luodahai@ntu.edu.sg (D. Luo).
[1] Present address: Protein Production Platform, Nanyang Technological University, Singapore.
[2] Present address: Department of Molecular Medicine, The Scripps Research Institute, Jupiter, FL, USA.

http://dx.doi.org/10.1016/j.antiviral.2017.04.001
0166-3542/

RdRP domain; the N-terminal "domain" is predicted to be mostly disordered and has little homology to known sequences. The absolutely conserved Tyr-1 residue was found to be important for the synthesis of minus-sense RNA (Rupp et al., 2011), and can only be mutated to other aromatic residues for activity to be preserved (Shirako and Strauss, 1998). Studies on the Sindbis virus nsP4 indicate that it is able to synthesize the negative sense anti-genome *de novo* and has terminal adenylyl transferase (TATase) activity (Rubach et al., 2009). However, the related full-length Sindbis virus (SINV) nsP4 has been shown to be insoluble when expressed in *E. coli* without additional solubilizing tags and optimized lysis buffers (Rubach et al., 2009). Therefore, it remains of interest to determine whether the C-terminal RdRP domain of CHIKV nsP4 may function on its own as the minimal catalytic core.

As viral RdRP proteins are the central component to RNA virus replication, they are an attractive target for antiviral development (Velkov et al., 2014; Waheed et al., 2013). Nucleoside and non-nucleoside inhibitors which specifically target the enzymes would be highly desirable as an infection cannot be established without viral RNA synthesis. In this work, we report a truncated nsP4 RdRP domain construct (nsP4-Δ118) from CHIKV that can be solubly expressed in *Escherichia coli* and purified to high homogeneity. Homology modelling suggests that all RdRP catalytic elements are present in the construct. The protein performs primed extension and TATase activity on short RNA oligonucleotides in the presence of a specific detergent N,N-dimethyldodecylamine N-oxide (LDAO). To our knowledge, this is the first reported truncated and active CHIKV nsP4 protein. The protein construct can be further utilized for structural studies and development of RdRP-inhibitor assays.

2. Results

2.1. CHIKV nsP4 RdRP domain is a soluble protein

Full-length CHIKV nsP4 (nsP4-FL) expressed recombinantly in *E. coli* could not be solubilized. We then tried to design core RdRP domain constructs that are devoid of the extraneous N-terminal region. Protein BLAST (Boratyn et al., 2012) suggested that residues 151–599 constitute the core RdRP domain, placing nsP4 in protein family pfam00978 that includes RdRPs from bromoviruses, tobamoviruses and togaviruses (Fig. 1A). However, the N-terminal region does not have significant homology to known protein families. Next, GlobPlot (Linding et al., 2003b) was used to identify subdomains and potentially disordered regions within nsP4. Residues 12–95 in the N-terminus were predicted to be globular in solution, but large portions of it may be intrinsically disordered as predicted by DisEMBL (Linding et al., 2003a). Residues 115–611 were predicted to form a globular domain and cover the entire core RdRP region. We then cloned, expressed and attempted to purify multiple constructs that are based on the suggested domain boundary. Other constructs were also made based on secondary structure prediction; a list of insoluble constructs is given in Table S1. Only one construct with the first 118 residues removed (nsP4-Δ118) (Fig. 1B) was soluble upon cell lysis. The tagged protein could be purified by using a three-step protocol to >95% purity by SDS-PAGE analysis (Fig. 1C).

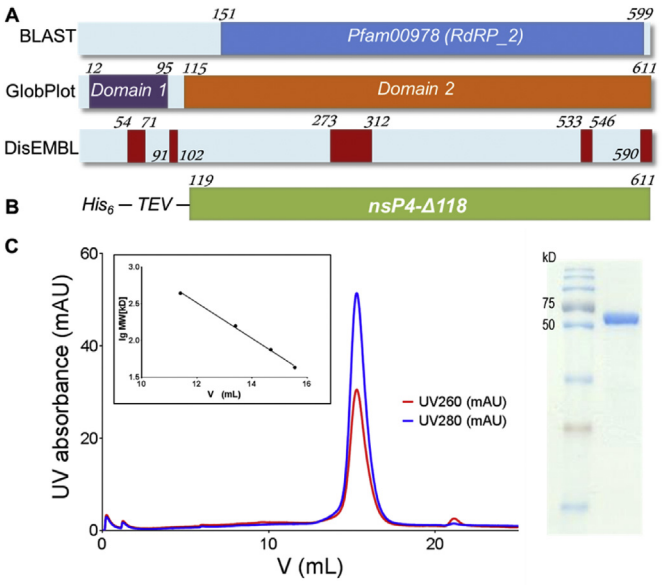

Fig. 1. Construct design and purification of a soluble, truncated CHIKV nsP4 protein. The numbering of amino acid residues in this paper shall be based on post-processing CHIKV nsP4 alone, i.e. starting from Y1 and ending with K611. **(A)** The 611-residue sequence was analysed to elucidate component domains within nsP4. Protein BLAST (Altschul et al., 1997) aligned residues 151–599 to the RdRP_2 group in the Conserved Domains Database, while the N-terminal region has no clear domain assignment. GlobPlot (Linding et al., 2003b) suggested the presence of two globular domains in nsP4; Domain 2 (residues 115–611) is longer than the RdRP core region and formed the basis of our construct design. DisEMBL (Linding et al., 2003a) suggested five major disordered regions based on the prediction of "hot loops", indicating that a large part of the N-terminal Domain 1, and the subsequent linker region, to be intrinsically disordered. **(B)** Construct nsP4-Δ118, with the first 118 residues removed and fused to an N-terminal hexahistidine tag and TEV protease cleavage site, could be solubly expressed in *E. coli*. **(C)** Analytical size-exclusion chromatography of purified nsP4-Δ118. Purified protein was injected onto a Superdex 200 10/300 column and eluted as a single symmetrical peak at an expected elution volume. Left inset: Calibration of the analytical column. Ovalbumin (43 kD), conalbumin (75 kD), aldolase (158 kD) and ferritin (440 kD) were used as reference. The protein size is estimated to be 60 kD (calculated: 57.2 kD) indicating its monomeric nature. Right inset: SDS-PAGE analysis of the purified protein showing >95% purity at an expected molecular size.

2.2. Homology modelling reveals an RdRP fold similar to picornaviral polymerases

Top hits that surfaced on BLAST queries against the PDB belonged to RdRPs of the Picornaviridae family, which are also positive-sense ssRNA viruses. The hits included RdRPs of Rhinovirus B14 (PDB ID: 1xr5)(Love et al., 2004), Enterovirus 71 (EV71; PDB ID: 4ika)(Chen et al., 2013), Coxsackie virus (PDB ID: 3ddk)(Campagnola et al., 2008), Encephalomyocarditis virus (EMCV; PDB ID: 4nz0)(Vives-Adrian et al., 2014), and Norwalk virus (PDB ID: 2b43)(Högbom et al., 2009) of the Caliciviridae. The RdRP amino acid sequences aligned with nsP4 sequence only within the predicted core RdRP region, resulting in sequence identity below 20% (Fig. 2A). nsP4 RdRP is more similar to the abovementioned "small" RdRPs as opposed to "large" flaviviral RdRPs that have C-terminal extensions containing extra features such as priming loops (Caillet-Saguy et al., 2014; Yap et al., 2007) (Fig. S1). In the absence of a crystal structure of nsP4, a structural model of CHIKV RdRP domain was generated by homology modelling using I-TASSER server (Yang et al., 2015). The C-score for the best-ranked model was −1.20 falling within the range for correct global topology.

The calculated nsP4-Δ118 model showed a classical, basic RdRP architecture with well-defined fingers, thumb and palm domains (Fig. 2B). On top of the palm domain and between the fingers and thumb domains was a tunnel across the structure, corresponding to the RNA-binding channel and RdRP active site. To compare the model to the crystal structures mentioned above, each structure was separately superimposed to the calculated nsP4-Δ118 model by using the secondary-structure matching algorithm in WinCoot (Emsley et al., 2010), with root-mean-squared deviations of