Virus Research 201 (2015) 94–100

Contents lists available at ScienceDirect

Virus Research

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

Structural and biophysical analysis of sequence insertions in the Venezuelan Equine Encephalitis Virus macro domain

Jaime Guillén[a,b], Julie Lichière[a,b], Nadia Rabah[a,b], Brett F. Beitzel[c], Bruno Canard[a,b], Bruno Coutard[a,b,*]

[a] CNRS, AFMB UMR 7257, 13288, Marseille, France
[b] Aix-Marseille Université, AFMB UMR 7257, 13288, Marseille, France
[c] Center for Genome Sciences, US Army Medical Research Institute of Infectious Disease, 1425 Porter Street, Fort Detrick, MD 21702, United States

ARTICLE INFO

Article history:
Received 26 August 2014
Received in revised form 18 February 2015
Accepted 19 February 2015
Available online 25 February 2015

Keywords:
Alphavirus
Non-structural protein
Macro domain
Sequence insertion
Structure
Biophysical studies

ABSTRACT

Random transposon insertions in viral genomes can be used to reveal genomic regions important for virus replication. We used these genomic data to evaluate at the protein level the effect of such insertions on the Venezuelan Equine Encephalitis Virus nsP3 macro domain. The structural analysis showed that transposon insertions occur mainly in loops connecting the secondary structure elements. Some of the insertions leading to a temperature sensitive viral phenotype (*ts*) are close to the cleavage site between nsP2 and nsP3 or the ADP-ribose binding site, two important functions of the macro domain. Using four mutants mimicking the transposon insertions, we confirmed that these insertions can affect the macro domain properties without disrupting the overall structure of the protein.

1. Introduction

Venezuelan equine encephalitis virus (VEEV) is a New World alphavirus, present in the United States and Central and South America. VEEV affects all equine species, but humans also can be infected causing flu-like symptoms and in few cases, encephalitis. In particular, children and individuals with weakened immune systems can become severely ill, with fatal outcomes or neurological sequels (Rivas et al., 1997). No vaccines or drugs are licensed for treatment against alphaviruses.

The 5′ two-third of the alphavirus genome encodes two polyproteins P123 and P1234 which are processed by a viral protease into four non-structural proteins (nsP1, nsP2, nsP3 and nsP4). nsP1–4 are involved in the viral replication and transcription machinery, whereas the 3′ part of the genome encodes the structural proteins capsid, E3, E2, 6K, and E1. Only partial and incomplete data about enzymatic activities and functional roles of non-structural proteins are available. NsP1 is a membrane-associated protein that bears methyltransferase and guanylyltransferase activities, required for RNA capping (Ahola and Kaariainen, 1995; Mi et al., 1989), and is necessary for the synthesis of minus-strand RNA (Lulla et al., 2008). NsP2 contains five domains and possesses multiple functions in viral replication. The N-terminal part, gathering three domains (a N-terminal domain, Rec1-like 1A and 2A domains), harbours RNA triphosphatase, and NTPase activities (Gomez de Cedron et al., 1999; Rikkonen et al., 1994; Vasiljeva et al., 2000), as well as helicase activity in collaboration with the C-terminal part of nsP2 (Das et al., 2014). The C-terminal part is carrying a protease activity for polyprotein processing (Golubtsov et al., 2006; Strauss et al., 1992) with a papain-like domain containing the catalytic diad, and followed by a classical 2′-O-methyltransferase fold. The latter domain is suggested to be not functional as a methyltransferase (Sawicki et al., 2006) but rather a major component for the regulation of minus-strand synthesis (Mayuri et al., 2008). It has therefore been called a MT-like domain. NsP3 functions, roles and activities are the least characterized and less well understood. NsP3 is essential in the transcription process at an early stage of the infection (LaStarza et al., 1994b; Shirako and Strauss, 1994; Wang et al., 1994). It is organized in three domains: a N-terminal macro domain which has been shown to bind ADP-ribose (ADPR), Poly ADP ribose (PAR) and RNA (Gorbalenya et al., 1991), a central alphavirus-specific Zinc Binding domain (ZBD) (Shin et al., 2012) and a hypervariable, phosphorylated C-terminal domain (Vihinen and Saarinen, 2000). NsP4

* Corresponding author at: Aix-Marseille Université, AFMB UMR 7257, 13288 Marseille Cedex 09, France. Tel.: +33 4 91 82 55 71; fax: +33 4 91 26 67 20.
E-mail address: bruno.coutard@afmb.univ-mrs.fr (B. Coutard).

http://dx.doi.org/10.1016/j.virusres.2015.02.018
0168-1702/

contains the RNA-dependent RNA polymerase (RdRp), involved in (−) and (+) strand RNA synthesis together with other nsPs (Rubach et al., 2009).

The processing of the polyprotein between nsP2 and nsP3 is a crucial step for the balance between synthesis of (−) and (+) RNA strands. Indeed, when this cleavage site called P23 is altered by mutations, the replication complex is producing continuously subnormal quantities of (−) RNA together with (+) RNA, whereas after cleavage occurs in *wt* virus, only genomic and subgenomic RNA are synthesized (Gorchakov et al., 2008; Kim et al., 2004; Mai et al., 2009; Sawicki and Sawicki, 1994). The processing of P23 requires the nsP2 protease that acts in *trans* but the integrity of P23 environment made of nsP2 MT-like domain, nsP3 macro domain and ZBD is also mandatory for an efficient substrate recognition and cleavage (Lulla et al., 2011). The structural organization of the P23 region has been determined by crystallographic studies, revealing that the cleavage site is located at the interface between nsP2 MT-like domain and nsP3 macro domain (Shin et al., 2012).

Preceding, and complementary, to the structural study of nsP2/nsP3, heat-sensitive mutations, traditionally known as temperature-sensitive (*ts*) mutations, have been widely used to study viral protein functions. Numerous studies have led to the identification of *ts* mutants in the non-structural proteins of alphaviruses (Hahn et al., 1989; Lemm et al., 1990; Lulla et al., 2006, 2008; Sawicki et al., 1990; Sawicki and Sawicki, 1993; Suopanki et al., 1998). Several of these *ts* mutants in the nsP3 region have been mapped and demonstrated the role of nsP3 in the synthesis of viral minus-strand and subgenomic RNAs (Hahn et al., 1989; LaStarza et al., 1994b; Wang et al., 1994). More recently, developments in large scale sequencing coupled to random transposon insertions enabled the increase of the throughput and resolution for the mapping of VEEV genome (Beitzel et al., 2010). In particular, the work performed by Beitzel and colleagues revealed regions of nsP3 that are tolerant for insertional mutations and leading to a *ts* phenotype.

In this study, we focused on sequence insertions on the nsP3 macro domain in order to understand if the viral *ts* phenotype of selected mutants could be related to a role of the macro domain in the P2 cleavage site, in the ADP-ribose and nucleotide binding site, or if the insertions affect the structure and stability of the protein. By means of structural analysis with the existing structural models followed by biophysical and biochemical characterizations, we show that *ts* insertions are located in the vicinity of the P2 interface or close to the ADP-ribose binding site and have different effects on thermal and chemical stability of the macro domain as well as different properties in ADP-ribose binding.

2. Material and methods

2.1. Structural models

The sequence of VEEV P23 region was submitted to PHYRE 2 webserver (Kelley and Sternberg, 2009) and the structural model was obtained using SINV P23 structure (PDB code 4GUA) as a template. *ts* regions were mapped on the built model and named *ts3-n*, where *ts* refers to temperature sensitivity, *3* to nsP3 and *n* is corresponding to the region number.

2.2. Cloning, expression and purification of VEEV macro domain

The cDNA corresponding to the nsP3 macro domain of VEEV (amino acids 1 to 160 of nsP3) were cloned into pDest14 plasmid using the "Gateway" cloning procedure (Life Technologies). A hexa-histidine sequence tag was fused at the C-terminal-end. One insertion mutant from each of the four *ts* regions in the macro domain (Beitzel et al., 2010) was selected and named *ts3-1* to *ts3-4* (see table in Fig. 1C). Coding sequences for *ts3-2* and *ts3-4* were kindly provided by Dr. Beitzel and Dr. Schmaljohn (The United States Army Medical Research Institute of Infectious Diseases, Fort Detrick, Maryland, USA) whereas *ts3-1* and *ts3-3* were engineered by site directed mutagenesis to mimic the transposition. The *wt* and *ts* proteins were then produced in *E. coli* Rosetta (DE3) pLysS cells (Novagen) at 37 °C in 2YT medium for 5 h after induction with 500 μM IPTG. Proteins were then purified by IMAC and SEC chromatography steps as described previously, resulting in pure proteins stored in buffer Hepes 10 mM, NaCl 150 mM, pH 7.5 (Malet et al., 2009).

2.3. Isothermal titration calorimetry (ITC)

ADP-ribose binding to VEEV macro domains was studied by ITC using a Microcal ITC200 microcalorimeter (MicroCal/GE Healthcare). Experiments were carried out at a constant temperature of 25 °C. The protein concentration in the cell was 50 μM, whereas the ADP-ribose (Sigma Aldrich) concentration in the syringe was 500 μM. The ADP-ribose solution was loaded into the 50 μl syringe and this titration solution was injected in 18 injections of 2 μl with a stirring speed of 750 rpm at discrete intervals of 200 s. Binding energy was measured by injecting the ligand into the protein solution. The heat dilution curve was subtracted from the experimental curves prior to data analysis. One-set of sites model was used to fit the resulting titration data using the Origin for ITC software package supplied by Microcal to obtain the stoichiometry (N), the dissociation constant (K_D) calculated from the binding affinity (K_a), the enthalpy (ΔH) and the entropy (ΔS) changes of binding. The Gibbs energy contribution was derived using the relationship:

$$\Delta G = \Delta H - T\Delta S \quad (1)$$

2.4. Thermal and chemical denaturations monitored by intrinsic fluorescence

Thermal denaturation was monitored by a decrease in intrinsic fluorescence intensity in a Varian Cary Eclipse spectrofluorimeter with excitation and emission of 280 and 350 nm, respectively, and 4 nm spectral bandwidths. A 1 cm pathlength quartz cell (Hellma) was used. Protein samples at 2 μM concentration were excited at 280 nm, and spectra were recorded between 300 and 400 nm every 2 °C.

For chemical denaturation, fluorescence emission spectra were performed on microtiter plates using a microplate reader (TECAN Safire 2) each well containing a final volume of 170 μl and stabilized at 25 °C. Protein samples at 1.7 μM concentration were excited at 280 nm, and spectra were recorded between 300 and 400 nm. Subtraction of background emission contribution was derived from urea titration in the buffer.

Changes with increasing concentrations of urea were followed by variations in the maximum fluorescence intensity. A least-squares analysis of this data was used to characterize the thermodynamics of the transition (Pace et al., 1987) obtaining the values of ΔG_0, m and $[\text{urea}]_{1/2}$:

$$\Delta G = -RT \ln K = -RT \ln[(I_F - I_x)/(I_x - I_U)] \quad (2)$$

where K is the equilibrium constant between the initial and final states, I_F and I_U are the fluorescence intensities of the initial and final state, respectively, and I_x is the experimental intensity at a given urea concentration. The value of ΔG_0 was calculated assuming a linear dependence of ΔG variation with the urea concentration in the transition:

$$\Delta G = \Delta G_0 - m[\text{urea}] \quad (3)$$