



## Biochemical characterisation of the nucleocapsid protein from a highly pathogenic porcine reproductive and respiratory syndrome virus strain

Stefanie S. Jourdan<sup>a</sup>, Fernando A. Osorio<sup>b</sup>, Julian A. Hiscox<sup>a,\*</sup>

<sup>a</sup> Institute of Molecular and Cellular Biology, Faculty of Biological Sciences, Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 6JP, UK

<sup>b</sup> School of Veterinary and Biomedical Sciences, Nebraska Centre for Virology, University of Nebraska-Lincoln, NE 68583-0900, USA

### ARTICLE INFO

#### Article history:

Received 11 November 2011

Available online 28 January 2012

#### Keywords:

Porcine reproductive and respiratory syndrome virus

PRRSV

Nucleocapsid protein

Biophysical

RNA binding

### ABSTRACT

The arterivirus nucleocapsid (N) protein is a multifunctional protein that binds viral RNA for encapsidation and has potential roles in host cell processes. This study characterised the N protein from a highly virulent North American strain of porcine reproductive and respiratory syndrome virus (PRRSV). The association with viral RNA was mapped to defined motifs on the N protein. The results indicated that disulphide bridge formation played a key role in RNA binding, offering an explanation why infectious virus cannot be rescued if cysteine residues are mutated, and that multiple sites may promote RNA binding.

© 2011 Elsevier Inc. All rights reserved.

### 1. Introduction

Coronaviruses and arteriviruses are families of positive strand RNA viruses that replicate in the cytoplasm of an infected cell. Due to similar genome organisation, replication and transcription strategies these viruses are grouped together into the *Nidovirales*. Many of their members cause important respiratory, gastrointestinal and reproductive diseases. These include severe acute respiratory syndrome coronavirus (SARS-CoV) [1], a zoonotic infection of humans and animals, and the arterivirus porcine reproductive and respiratory syndrome virus (PRRSV) that causes reproductive failure in pregnant sows, high mortality in piglets and respiratory disease in pigs of any age [2,3]. Although vaccines to a number of coronaviruses and arteriviruses exist, the high mutation rates during virus replication leads to new viral threats that are constantly emerging [4]. A more molecular understanding of viral protein function would allow the development of new vaccines, identify targets for targeted antiviral therapy and provide robust diagnostics [5,6].

The PRRSV genome shows rapid change and strains can vary in virulence [7], with various genes contributing to pathogenesis, including the nucleocapsid (N) protein [8–10]. For both the coronaviruses and arteriviruses the N protein is one of the most abundant virally encoded proteins and plays essential roles in the virus

life cycle [11,12]. Its most prominent role is complexing with viral RNA, although the mechanisms for this for either type of viral N protein is not precisely characterised. For the coronavirus N protein phosphorylation has been shown to play a role in mediating binding to RNA [13,14] with the amino-terminal region facilitating long-range non-specific electrostatic interactions potentially allowing contacts to form via a 'lure' and 'lock' mechanism [15].

Sequence comparison initially suggested that coronavirus N proteins exhibited a three-domain structure. This includes a disordered unstructured region at the N-terminal end that is followed by the structured N-terminal domain (NTD). A disordered serine/arginine- (SR-) rich region is located in the central position of the coronavirus N protein. Similarly the extreme C-terminus is disordered with a proximally structured C-terminal domain (CTD). The arterivirus N proteins are smaller in size with regard to molecular weight (the coronavirus N protein is approximately 45 kDa whereas the arterivirus N protein is approximately 15 kDa) and consists of an N-terminal half that is mainly disordered and a structured C-terminal half which has a similar fold as the CTD of the coronaviruses [16]. Due to the abundance of disordered regions in combination with proteolytic instability of the full length N proteins [17] the detailed structure of the N protein has not been solved in entirety using any one single technique and our picture of this is based on multiple complementary analyses.

The X-ray structures of the CTD of the arterivirus equine arteritis virus (EAV) and PRRSV N proteins have been solved and show structural similarities to the CTD of the coronavirus N protein [16,18,19]. There is a very high overall similarity between the CTD of EAV and PRRSV N proteins. The main differences exist in

\* Corresponding author. Address: Room 8.58, Garstang Building, Institute of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK. Fax: +44 113 343 5638.

E-mail address: [j.a.hiscox@leeds.ac.uk](mailto:j.a.hiscox@leeds.ac.uk) (J.A. Hiscox).

their  $\beta$ -loops and their most C-terminal  $\alpha$ -helices that end about 6 residues before the C-terminus. Monomers of EAV and PRRSV N harbour two anti-parallel  $\beta$ -strands that are flanked by two  $\alpha$ -helices. Both proteins contain another short  $\alpha$ -helix towards the N-terminal of the CTD. Moreover, PRRSV as well as EAV N are predicted to contain an  $\alpha$ -helix within their otherwise unstructured N-terminal half. Similar to the coronavirus N proteins EAV and PRRSV N form dimers with a  $\beta$ -sheet platform that is surrounded on one side of the dimer by the  $\alpha$ -helices. With a view to potential RNA binding or other interactions taking advantage of electrostatic interactions, the electrostatic potential is more positive on the side of the dimer on which the helices reside. The proposed RNA binding region of the PRRSV N protein, based on the structure, is in the unstructured N-terminal half, located within a region harbouring positively charged residues. With a negative electrostatic potential found on the side of the  $\beta$ -sheet floor this could support correct positioning of the RNA in complex to N protein. PRRSV N is found in a head to head helical arrangement within the crystals, supporting the ribbon-like pattern observed for N using cryo-tomography [20].

To provide a deeper understanding of the role of PRRSV N protein in binding RNA and to delineate RNA binding regions a sequential alanine substitution approach was used in combination with various biophysical assays and selected further point mutation. This analysis highlighted that several sites on the PRRSV protein may act cooperatively to promote RNA binding.

## 2. Materials and methods

### 2.1. Bacterial strains and construction of plasmids

The N gene from PRRSV strain NVSL #97-7895 was cloned into the pTriEx1.1Neo (Novagen) vector. Alanine block substitutions were constructed by a two-step PCR approach creating products allowing for a triple ligation with the pTriEx1.1Neo vector. Plasmid pTriExN<sub>C23S</sub> coding for an N protein carrying a cysteine (cys) to serine mutation at residue 23 was created by site directed mutagenesis.

### 2.2. Protein expression and purification

Plasmids coding for N protein variants were transformed into Tuner (DE3) pLacI (Novagen) *E. coli* cells that allowed for efficient protein expression.

### 2.3. Gel filtration chromatography

Aliquots of purified protein were analysed by analytical gel filtration chromatography (GFC) using a BioSep-Sec-S3000 column (Phenomenex) at a flow rate of 0.8 ml/min of the mobile phase; 20 mM Tris HCl pH 7, 0.5 M NaCl. Protein standards were used to establish a standard curve; Cytochrome C (12.4 kDa), carbonic anhydrase (20 kDa), bovine serum albumin (66 kDa),  $\beta$ -amylase (200 kDa), catalase (210 kDa), Dextran blue (2000 kDa). The molecular mass of the contents of the peak eluting upon N protein injection was estimated from its retention time using the standard curve.

### 2.4. Circular dichroism (CD)

Proteins used in CD analysis were purified without the addition of reducing agents and treated with high RNase A concentrations so that no RNA was co-purified. Prior to CD analysis proteins were dialysed into buffer containing 20 mM sodium phosphate pH 7.0 and 0.3 M or 0.5 M sodium fluoride. CD data was analysed using

Dichroweb [21] (<http://dichroweb.cryst.bbk.ac.uk/html/home.shtml>) similar to that described for the coronavirus N protein [15].

## 3. Results and discussion

### 3.1. The N protein interacts with RNA

To investigate RNA binding and establish an *in vitro* RNA binding assay for the N protein, the N gene from an infectious cDNA of the highly lethal North American PRRSV strain NVSL #97-7895 [22] was cloned into the expression vector pTriEx1.1Neo (Novagen) and the recombinant protein purified from *E. coli* using metal affinity chromatography as described [13]. RNA co-purification was used to indicate the RNA binding capability of N protein. The amount of RNA that could be extracted from an equal amount of purified protein decreased when the protein was incubated with increased RNase A concentration (Fig. 1A). To control for protein stability, protein samples were taken at the end of the incubation and analysed by SDS-PAGE (Fig. 1A).

The size of the co-purified RNA was estimated by electrophoretic analysis and a major species found to be approximately 150 nucleotides. To investigate whether the 150 nucleotides were bound by one N protein molecule, analytical gel filtration chromatography (GFC) was performed (Fig. 1B). These molecular weights indicated that the two peaks contained multimers and possibly monomers, respectively. From the mass deduced from the elution time it is tempting to speculate that peak 1 represents a decamer, suggesting that each monomer protected an average of 15 nt from nucleolytic cleavage.

### 3.2. Global structure of the PRRSV N protein

No complete structure for the PRRSV N protein is available, therefore a combination of modelling using the sequence encompassing the CTD of the N protein from the VR-2332 strain was used to map the equivalent region on the NVSL #97-7895 strain N protein using the HHPred program (pdb file 1p65; <http://toolkit.tuebingen.mpg.de/hhpred>) and the secondary structure prediction program Jpred (<http://www.compbio.dundee.ac.uk/www-jpred/>) was employed to model potential folds within the N-terminal half (NTH) (Fig. 2A). Overall the structure was confirmed using CD (Fig. 2B). For example, these suggested 15% alpha helix, 29% beta sheet and 56% random coil by the K2d neural network programme.

### 3.3. Potential role of disulphide bridge formation in RNA binding

The North American PRRSV N protein contains three cys residues at positions 23, 75 and 90 and has been shown to be present as disulphide-linked dimers in virions [23,24]. However, only equivalents of Cys23 and Cys75 are present within N proteins from European PRRSV strains. Cys23 and Cys90 may play an important role for in the virus life cycle as substitution of these residues with serine in recombinant infectious virus led to abrogation of progeny virus production [25]. To determine whether the cys residues had possible roles in mediating RNA binding, N protein was purified under reducing conditions and resulted in less RNA being co-purified (Fig. 3A). This indicated that the formation of disulphide linkages may play a role in RNA binding. Examination of the predicted N protein structure (Fig. 2A) indicated that Cys75 and Cys90 are too far away from their partners in the partial crystal structure to form disulphide bridges [19]. The possible role of Cys23 in RNA-binding was investigated by substituting this residue with serine (mutant N<sub>C23S</sub>), and resulted in co-purification of no or very small amounts of RNA (Fig. 3B).

Download English Version:

<https://daneshyari.com/en/article/10762162>

Download Persian Version:

<https://daneshyari.com/article/10762162>

[Daneshyari.com](https://daneshyari.com)