Available online at www.sciencedirect.com
ScienceDirect
Virology 370 (2008) 373–381
VIROLOGY
www.elsevier.com/locate/yviro

Role of phosphorylation clusters in the biology of the coronavirus infectious bronchitis virus nucleocapsid protein

Kelly-Anne Spencer [a], Michael Dee [b], Paul Britton [b], Julian A. Hiscox [a,c,*]

[a] Institute of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds, LS2 9JT, UK
[b] Division of Microbiology, Institute for Animal Health, Compton, Newbury, Berkshire, RG20 7NN, UK
[c] Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, LS2 9JT, UK

Received 16 July 2007; returned to author for revision 11 August 2007; accepted 16 August 2007
Available online 1 November 2007

Abstract

The coronavirus infectious bronchitis virus (IBV) nucleocapsid (N) protein is an RNA binding protein which is phosphorylated at two conserved clusters. Kinetic analysis of RNA binding indicated that the C-terminal phosphorylation cluster was involved in the recognition of viral RNA from non-viral RNA. The IBV N protein has been found to be essential for the successful recovery of IBV using reverse genetics systems. Rescue experiments indicated that phosphorylated N protein recovered infectious IBV more efficiently when compared to modified N proteins either partially or non-phosphorylated. Our data indicate that the phosphorylated form of the IBV N protein plays a role in virus biology.

Keywords: Coronavirus; RNA binding; Nucleocapsid protein; IBV; Surface plasmon resonance; Phosphorylation

Introduction

One of the fundamental stages in the life cycle of coronaviruses and other RNA viruses is the recognition of viral genomic RNA by the virus encoded nucleocapsid protein, for roles in RNA synthesis, encapsidation, packaging and in the correct folding of the RNA molecule. Phosphorylation of the nucleocapsid protein is thought to regulate many of these events although precise roles remain unclear. Coronaviruses are a group of positive strand RNA viruses which cause principally respiratory and gastro-intestinal infections and include the severe acute respiratory syndrome coronavirus (SARS-CoV) (Peiris et al., 2004) and avian infectious bronchitis virus (IBV) (Cavanagh, 2005).

Coronavirus genomes are approximately 30 kb in length, making them the largest RNA genomes so far discovered. Both RNA replication and mRNA transcription occur in virus infected cells and a series of subgenomic mRNA molecules are synthesized by the viral replicase via a discontinuous transcription mechanism (Enjuanes et al., 2006a; Pasternak et al., 2006; Sawicki et al., 2007), in which a viral leader sequence, derived from the 5′ end of the genome, is added to the 5′ end of each subgenomic mRNA (Masters, 2006). Adjacent to the leader sequence is a transcription regulatory sequence (TRS) which also precedes each open reading frame along the genome. Genetic analysis suggests that the TRS and flanking sequences are involved in the modulation of transcription (Alonso et al., 2002; Hiscox et al., 1995; Sola et al., 2005; Van Marle et al., 1995; Zuniga et al., 2004). The TRS has been shown to act as a high affinity binding site for the coronavirus nucleocapsid (N) protein suggesting a role for N protein in transcription (Baric et al., 1988; Nelson et al., 2000; Spencer and Hiscox, 2006a; Stohlman et al., 1988). Although replication can take place without N protein it does so at reduced efficiency (Almazan et al., 2004; Schelle et al., 2005). Furthermore N protein has been shown to co-localize with coronavirus replication complexes in the cytoplasm at early points post-infection (Denison et al., 1999; Van der Meer et al., 1999). Moreover, N protein is required for the efficient rescue of coronavirus genomic RNA from cDNA clones (Almazan et al., 2004; Coley et al., 2005;

* Corresponding author. Institute of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds, LS2 9JT, UK.
E-mail address: j.a.hiscox@leeds.ac.uk (J.A. Hiscox).

0042-6822/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2007.08.016

Schelle et al., 2005; Yount et al., 2000, 2003) and has been shown to be essential for the rescue of IBV using at least two different types of reverse genetics systems (Britton et al., 2005; Casais et al., 2001; Youn et al., 2005a,b).

Coronavirus N proteins have roles in both virus RNA synthesis and modulating host cell processes and phosphorylation may regulate these processes by exposing various functional motifs (You et al., 2005, 2007). By amino acid sequence comparison coronavirus N proteins can be separated into three regions of conserved homology (Parker and Masters, 1990). Situated within these regions are motifs which play a role in both aspects of virus biology. For example, IBV N protein contains an RNA binding site (Tan et al., 2006), a nucleolar localization signal (Reed et al., 2006) and a nuclear export signal (Reed et al., 2007). These latter two signals modulate its dynamic trafficking (Cawood et al., 2007) between the cytoplasm (the site of viral RNA synthesis) and the nucleolus, suggesting a nucleolar aspect to the virus life cycle (Hiscox, 2007).

Recent studies have shown that the N-terminal region RNA binding site of IBV N protein has the potential to form an electrostatic interaction with RNA (Fan et al., 2005). This was subsequently confirmed by measuring dynamic interactions (using surface plasmon resonance (SPR)) (Spencer and Hiscox, 2006a) and raised the possibility that N protein initially bound viral RNA via a weak electrostatic interaction (a 'lure' step) and that a 'lock' step was facilitated by specific high affinity binding to structures on the viral RNA such as the TRS (Spencer and Hiscox, 2006a). This latter stage being similar to the recognition of the hantavirus vRNA panhandle by its nucleocapsid protein (Mir et al., 2006) or the interaction of HIV-1 Rev protein with loop A (Greatorex et al., 2006). Again, similar to the ability of the hantavirus nucleocapsid protein (Mir and Panganiban, 2006), the coronavirus N protein has also been shown to have RNA chaperone activity (Zuniga et al., 2007).

IBV N protein has a predicted molecular weight of 45 kDa, although migrates under SDS–PAGE with a mobility of ~50 kDa, suggestive of post-translational modifications (Chen et al., 2003), which was confirmed by the incorporation of ^{32}P-orthophosphate into N protein during infection and was also found in virions (Jayaram et al., 2005). Although coronavirus N proteins were predicted to be phosphorylated at multiple serine residues, mass spectroscopic analyses identified two regions of phosphorylation on the IBV N protein (when over-expressed) which are located at two conserved amino acid clusters, Ser^{190} and Ser^{192} and Thr^{378} and Ser^{379} (Chen et al., 2005). The protein was heterogeneously phosphorylated with the predominant species having all four sites occupied (Chen et al., 2005), which reflects the mobility of IBV N protein (detected using western blot) during infection, in which a single N protein species is observed (Dove et al., 2006). Similarly the porcine coronavirus, transmissible gastroenteritis virus (TGEV), N protein was also found to be phosphorylated at two clusters, amino acids Ser^{156}, Ser^{254} and Ser^{256} (Calvo et al., 2005). The murine coronavirus, mouse hepatitis virus N protein is phosphorylated at Ser^{170}, Thr^{177}, Ser^{389}, Ser^{424} and Thr^{428} (White et al., 2007). Although the precise sites have not been mapped, the SARS-CoV N protein is thought to be phosphorylated at serine residues (Surjit et al., 2005). Thus for IBV and MHV N proteins these phosphate groups are located in the middle and C-terminal regions of the protein, and overall the IBV, TGEV and MHV analyses demonstrate that coronavirus N proteins are phosphorylated but not at multiple (<10) serine residues which was once accepted dogma based upon prediction analysis.

Kinetic analysis indicated that phosphorylation of the IBV N protein played a role in discriminating between non-viral (random RNA) and IBV RNA (the leader sequence); the phosphorylated N protein having a high binding affinity for viral RNA compared to non-viral RNA (Chen et al., 2005). In contrast, non-phosphorylated N protein bound to either viral or non-viral RNA with similar affinities (Chen et al., 2005).

However, given the conserved nature of N protein phosphorylation, the role of whether both or single phosphorylation clusters mediated RNA binding and the general role of phosphorylation of N protein in virus biology has not previously been elucidated and is addressed in this study.

Results and discussion

Generation and characterization of differentially phosphorylated N proteins

To investigate the role of the phosphorylation clusters in the binding of IBV N protein to RNA and in the biology of virus replication, site directed mutagenesis was used to replace the appropriate amino acid coding sequences in the IBV N gene sequence in the expression plasmid pTriExIBVN, generating plasmids pTriExIBVN$_{SSAA}$, pTriExIBVN$_{AATS}$ and pTriExIBVN$_{AAAA}$; all with a C-terminal his-tag coding sequence for protein purification purposes. Plasmid pTriExIBVN had previously been used to make a recombinant baculovirus for expressing phosphorylated IBV N protein in Sf9 cells (referred here as N$_{SSTS}$) (Chen et al., 2003, 2002). Previous mass spectroscopic analysis demonstrated that N protein expressed in insect cells had identical post-translational modifications to that produced in a cell line (Vero) permissive to virus infection (Chen et al., 2005). A similar approach was used in this study to generate recombinant baculoviruses for the expression of modified IBV N proteins. The baculoviruses were used to express modified N proteins in which either Ser^{190} and Ser^{192} (N$_{AATS}$) or Thr^{378} and Ser^{379} (N$_{SSAA}$) were substituted for the amino acid alanine resulting in a partially phosphorylated IBV N protein. A third recombinant baculovirus was produced that expressed a modified N protein in which all four amino acids were substituted for alanine resulting in a non-phosphorylated (N$_{AAAA}$) form of the IBV N protein. The N$_{AAAA}$ protein was used in place of *E. coli* expressed N protein (non-phosphorylated protein) to ensure that any additional post-translational modifications, such as acetylation, were present.

The modified IBV N proteins were purified by nickel affinity chromatography and analysis of the recombinant purified proteins using SDS–PAGE indicated that N$_{SSAA}$ and N$_{AATS}$ had identical mobility, whereas N$_{AAAA}$ had the fastest mobility