Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Available online at www.sciencedirect.com
ScienceDirect
Virus Research 126 (2007) 139–148
Virus Research
www.elsevier.com/locate/virusres

Identification of mouse hepatitis coronavirus A59 nucleocapsid protein phosphorylation sites

Tiana C. White [a,b,c], Zhengping Yi [b,d], Brenda G. Hogue [a,b,*]

[a] *The Biodesign Institute, Center for Infectious Diseases and Vaccinology, Arizona State University, Tempe, AZ 85287-5401, United States*
[b] *School of Life Sciences, Arizona State University, Tempe, AZ 85287-5401, United States*
[c] *Molecular Cell Biology Graduate Program, Arizona State University, Tempe, AZ 85287-5401, United States*
[d] *Proteomics Laboratory in the Center for Metabolic Biology, Arizona State University, Tempe, AZ 85287-5401, United States*

Received 7 December 2006; received in revised form 4 February 2007; accepted 8 February 2007
Available online 23 March 2007

Abstract

The coronavirus nucleocapsid (N) is a multifunctional phosphoprotein that encapsidates the genomic RNA into a helical nucleocapsid within the mature virion. The protein also plays roles in viral RNA transcription and/or replication and possibly viral mRNA translation. Phosphorylation is one of the most common post-translation modifications that plays important regulatory roles in modulating protein functions. It has been speculated for sometime that phosphorylation could play an important role in regulation of coronavirus N protein functions. As a first step toward positioning to address this we have identified the amino acids that are phosphorylated on the mouse hepatitis coronavirus (MHV) A59 N protein. High performance liquid chromatography coupled with electrospray ionization tandem mass spectrometry (HPLC-ESI-MS/MS) was used to identify phosphorylated sites on the N protein from both infected cells and purified extracellular virions. A total of six phosphorylated sites (S162, S170, T177, S389, S424 and T428) were identified on the protein from infected cells. The same six sites were also phosphorylated on the extracellular mature virion N protein. This is the first identification of phosphorylated sites for a group II coronavirus N protein.

Keywords: Coronaviruses; Mouse hepatitis coronavirus; Nucleocapsid protein; Phosphorylation; Mass spectrometry

1. Introduction

The *Coronaviridae* is a large family of medically important viruses that cause primarily respiratory and enteric infections in humans and a wide range of animals. The viruses are enveloped and contain a single-stranded positive-sense RNA genome that ranges in size from 27 to 31 kb. All of the viruses have at least four structural proteins. The spike (S), membrane (M) and envelope (E) proteins are anchored in the virion envelope. The nucleocapsid (N) protein encapsidates the viral genome as a helical nucleocapsid inside the virion (Davies et al., 1981; Macnaughton et al., 1978).

The focus of this report is the multifunctional N protein and its phosphorylation. Through its interactions with the viral RNA, the M protein and itself, N plays important roles in virus assembly (Escors et al., 2001; Hurst et al., 2005; Narayanan et al., 2003; Narayanan and Makino, 2001; Verma et al., 2006). The protein is also involved in viral RNA transcription and/or replication (Baric et al., 1988; Chang and Brian, 1996; Compton et al., 1987; Denison et al., 1999; van der Meer et al., 1999). Recent studies with coronavirus infectious clones provided direct evidence for a role of the N protein in replication and/or transcription (Almazan et al., 2004; Casais et al., 2001; Yount et al., 2000). The protein may also play a role in viral mRNA translation (Tahara et al., 1994). Additionally, MHV A59 and severe acute respiratory syndrome coronavirus (SARS-CoV) N proteins are type I interferon antagonists (Kopecky-Bromberg et al., 2007; Ye et al., 2007).

All coronavirus N proteins are phosphorylated and highly basic with isoelectric points (p*I*) of 10.3 to 10.7 (Laude and Masters, 1995). The role of phosphorylation is not known and only very recently were phosphorylated sites identified for transmissible gastroenteritis virus (TGEV) and avian

* Corresponding author at: The Biodesign Institute, P.O. Box 875401, Arizona State University, Tempe, AZ 85287-5401, United States. Tel.: +1 480 965 9478; fax: +1 480 727 7615.
E-mail address: Brenda.Hogue@asu.edu (B.G. Hogue).

0168-1702/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.virusres.2007.02.008

infectious bronchitis virus (IBV) N proteins, group I and III viruses, respectively (Calvo et al., 2005; Chen et al., 2005).

In the study reported here we used high pressure liquid chromatography coupled with electrospray ionization tandem mass spectrometry (HPLC-ESI-MS/MS) to identify the phosphorylated sites on the MHV A59 N protein from both infected cells and purified extracellular virions. High accurate mass resolution was achieved by ion cyclotron resonance (ICR) and Fourier transformation. Use of the ICR cell provides the most accurate mass data available today. A total of six phosphorylated sites (S162, S170, T177, S389, S424, and T428) were identified on the N protein taken from both infected cells and extracellular virions. This is the first identification of phosphorylated sites for a group II coronavirus N protein.

2. Materials and methods

2.1. Isolation of N protein from intracellular and extracellular fractions

Mouse 17Cl1 cells were infected with 0.1 pfu/ml of MHV A59. Intracellular and extracellular fractions were harvested separately at 18 h p.i. Cell culture supernatant containing extracellular virus was clarified by centrifugation to remove cell debris. Infected cells were rinsed with phosphate buffered saline chilled to 4 °C and disrupted in ice cold lysis buffer (100 mM Tris–HCl [pH 7.5], 100 mM NaCl, 0.5% Triton X-100, 1 mM phenylmethylsulfonyl fluoride, 1:100 dilution of phosphatase inhibitor cocktails 1 and 2 (Sigma), 10 μg RNase A). Nuclei were pelleted and the cytoplasmic lysates were resolved immediately by SDS-PAGE or stored at −80 °C.

Virions were precipitated from the clarified extracellular supernatant by addition while mixing of polyethylene glycol (PEG) 8000 to a final concentration of ∼12.5 mM, followed by slow addition of NaCl to a final concentration of 400 mM. The precipitation was continued for 3 h at 4 °C with continuous stirring. Precipitates were collected by centrifugation at 4 °C for 20 min at 10,000 × *g*. The precipitate was resuspended in TMEN (50 mM Tris–HCl, 50 mM maleic acid, 1 mM EDTA, 100 mM NaCl) buffer [pH 6.0] and clarified at 10,000 × *g* for 5 min. Clarified precipitated virus was purified in continuous 20–60% (w/w) sucrose gradients that were centrifuged at 90,000 × *g* for 4.5 h at 4 °C. Virions from pooled gradient fractions in the 1.16–1.20 g/cm^3 density range were subsequently pelleted through a 30% sucrose cushion at 30,000 rpm in a Beckman SW 50.1 rotor. The purified virus was resuspended in TMEN buffer [pH 6.0] and analyzed immediately or stored at −80 °C. Virions were disrupted in ice cold lysis buffer as described above prior to resolution of proteins by SDS-PAGE.

2.2. Isolation and preparation of the N protein for mass spectrometry

Proteins from the intracellular and extracellular fractions were separated by SDS-PAGE in 8% gels and visualized by Coomassie staining. The N protein from both fractions was excised from the gel. Positions of the N protein were confirmed by Western blotting of proteins from parallel lanes. Gel pieces containing the N protein were destained twice with 300 μl of 50% acetonitrile (ACN) in 40 mM NH_4HCO_3 and dehydrated with 100% ACN for 15 min. ACN was removed by aspiration and gel pieces were dried in a vacuum centrifuge at 60 °C for 20 min. Dried gel pieces were rehydrated in 20 μl of 40 mM NH_4HCO_3 containing 250 ng trypsin (Sigma) at 4 °C for 15 min prior to the addition of 50 μl of 40 mM NH_4HCO_3 containing 10 fmol/μl angiotensin II. Trypsin digests were incubated overnight at 37 °C. Protease activity was terminated by addition of 10 μl of 5% formic acid (FA) and incubation at 37 °C for 30 min. Digests were clarified for 1 min in a microcentrifuge and supernatants were removed. The extraction procedure was repeated using 40 μl of 5% FA. The resulting peptide mixtures were purified by solid-phase extraction using C18 ZipTips (Millipore) after loading in 0.05% heptafluorobutyric acid: 5% FA (v/v) and elution with 50% ACN:1% FA (v/v). The samples were dried by vacuum centrifugation and dissolved in 4 μl of 0.1% FA:2%ACN (v/v).

2.3. Mass spectrometry

HPLC-ESI-MS/MS was performed on a Thermo Finnigan (San Jose, CA) LTQ-FTICR fitted with a PicoView^{TM} nanospray source (New Objective, Woburn, MA). On-line HPLC was performed using a Michrom BioResources Paradigm MS4 micro two-dimensional HPLC (Alburn, CA) with a PicoFrit^{TM} column (New Objective, Woburn, MA, 75 μm i.d., packed with ProteoPep^{TM} II C18 material, 300 Å); mobile phase, linear gradient of 2 to 27% ACN in 0.1% FA in 45 min, a hold of 5 min at 27% ACN, followed by a step to 50% ACN, a hold of 5 min and then a step to 80%; flow rate, 250 nl/min.

A "top-10" data-dependent MS/MS analysis was performed (acquisition of a full scan spectrum followed by collision-induced dissociation (CID) mass spectra of the 10 most abundant ions in the survey scan) to identify N protein peptides. The survey scan was acquired using the Fourier transform ion cyclotron resonance (FTICR) mass analyzer, which offers high mass accuracy and resolution. A list of potential phosphorylated peptides was generated based on detected serine/threonine-containing peptides from the N protein. For localization of phosphorylation sites a scan protocol of 1 survey scan (FTICR), followed by 7 targeted MS/MS scans (CID spectra of specified *m/z* values were acquired using the LTQ mass analyzer). All uninterpreted tandem MS data were searched using Mascot (Matrix Science, London, UK). Assignments of the phosphopeptides were confirmed by manual comparison of the tandem mass spectra with the predicted fragmentation generated in silico by the MS-Product component of ProteinProspector (http://prospector.ucsf.edu). In the case of phosphopeptides containing more than one serine or threonine, we localized phosphorylation to particular residues by assigning fragment ion masses from the mass spectra that were unique to the fragmentation of a peptide phosphorylated at said residue.

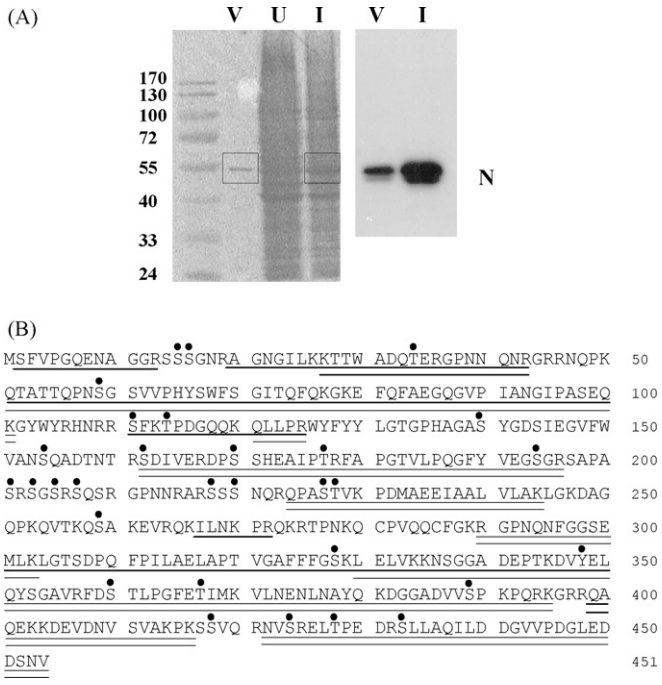


Fig. 1. Representative SDS-PAGE used to isolate MHV N protein and peptide map coverage. (A) Proteins were visualized by Coomassie staining (left panel). Boxes indicate the areas of gel excision. The positions of N from purified virus (V) and MHV infected cells (I) are indicated to the right of both panels. Control uninfected cells (U) are also shown. Positions of molecular weight standards in kDa are shown on the left. Western blot analysis of parallel lanes was used to verify the identity of the N protein (right panel). A rabbit polyclonal N antibody was used to identify the N protein (Cologna et al., 2000). (B) Sequence coverage of peptides that identified phosphorylated resides is shown. Double lines indicate coverage for both intracellular and extracellular N protein fractions. Single lines identify additional coverage of the intracellular fraction. Dots mark predicted phosphorylation sites.

3. Results

3.1. MHV A59N protein contains many potential phosphorylation sites

The MHV A59 N protein consists of 454 amino acids. The protein contains 41 serine, 22 threonine and 11 tyrosine residues, approximately 9, 5 and 2.4%, respectively, of the total amino acids. Of these, 24 serines, 5 threonines and 1 tyrosine are predicted by NetPhos 2.0 (http://www.cbs.dtu.dk/services/NetPhos/) (Blom et al., 1999) to be potential phosphorylation sites. To determine which of the predicted sites are indeed phosphorylated, the N proteins from both intracellular lysates and purified extracellular virions from MHV A59 infected mouse 17Cl1 cells were analyzed by HPLC-ESI/MS/MS.

3.2. Identification of phosphorylated sites on extracellular virion N protein by HPLC-ESI-MS/MS

To identify phosphorylated sites on the N protein in extracellular mature virions, virus was purified and proteins were separated by 1D SDS-PAGE and visualized by Coomassie staining. The N protein band was excised (Fig. 1) and in-gel digested with trypsin. Peptide pools were extracted, desalted and analyzed by on-line HPLC-ESI-MS/MS.

Mass spectra were acquired using a Thermo Finnigan LTQ-FTICR. FTICR mass analyzer has the ability to obtain high mass accuracy and resolution while the LTQ mass analyzer can perform tandem mass spectrometry (MS/MS and MS^3) to obtain sequence information of peptides by collision-induced dissociation (CID) of residues. It is one of the most sensitive instruments available for high mass resolution and mass accuracy for identification of protein phosphorylation and complex proteomic studies. We obtained 54% sequence coverage (by MASCOT) of the virion N protein (Fig. 1B).

The first phosphorylated site identified was serine 389 (pS389). The identification was made from analysis of MS/MS spectra of the peptide $N^{382–394}$ (DGGADVVpSPKPQR) which exhibited the most intense peak at mass-to-charge ratio (m/z) of 654.6, corresponding to $[M+2H-98]^{2+}$ (Fig. 2). The mass is representative of a doubly charged ion with a decrease in mass of 98 atomic mass units (amu), which is equal to the loss of one phosphoric acid group (H_3PO_4) from the precursor ion $m/z = 703.3^{2+}$. Since this peptide contained only one serine, MS^3 analysis was not necessary. Other intense peaks were assigned as b- and y-series sequence ions which result from the usual fragmentation of peptides at their amide bonds (Fig. 2). These results confirmed preliminary results from earlier MALDI analysis in our

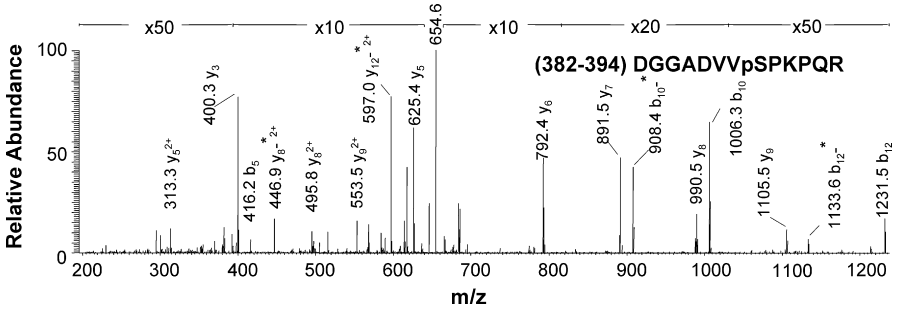


Fig. 2. CID MS/MS spectrum of phosphopeptide D_{382}–R_{394} from tryptic digest of MHV virion N protein. Peaks correspond to fragments of the peptide containing residues 382–394 as indicated by the sequence shown with the spectrum. The most intense peak ($m/z = 654.6$ $[M+2H-98]^{2+}$) corresponds to a precursor ion ($m/z = 703.3^{2+}$) with $a+2$ charge and minus H_3PO_4. The position of phosphoserine at position 389 (pS) is indicated in the peptide sequence. The b-amino and y-carboxy series ions are indicated above peptide fragment peaks. Symbol (*) indicates loss of H_3PO_4 (98 amu) from a fragment.

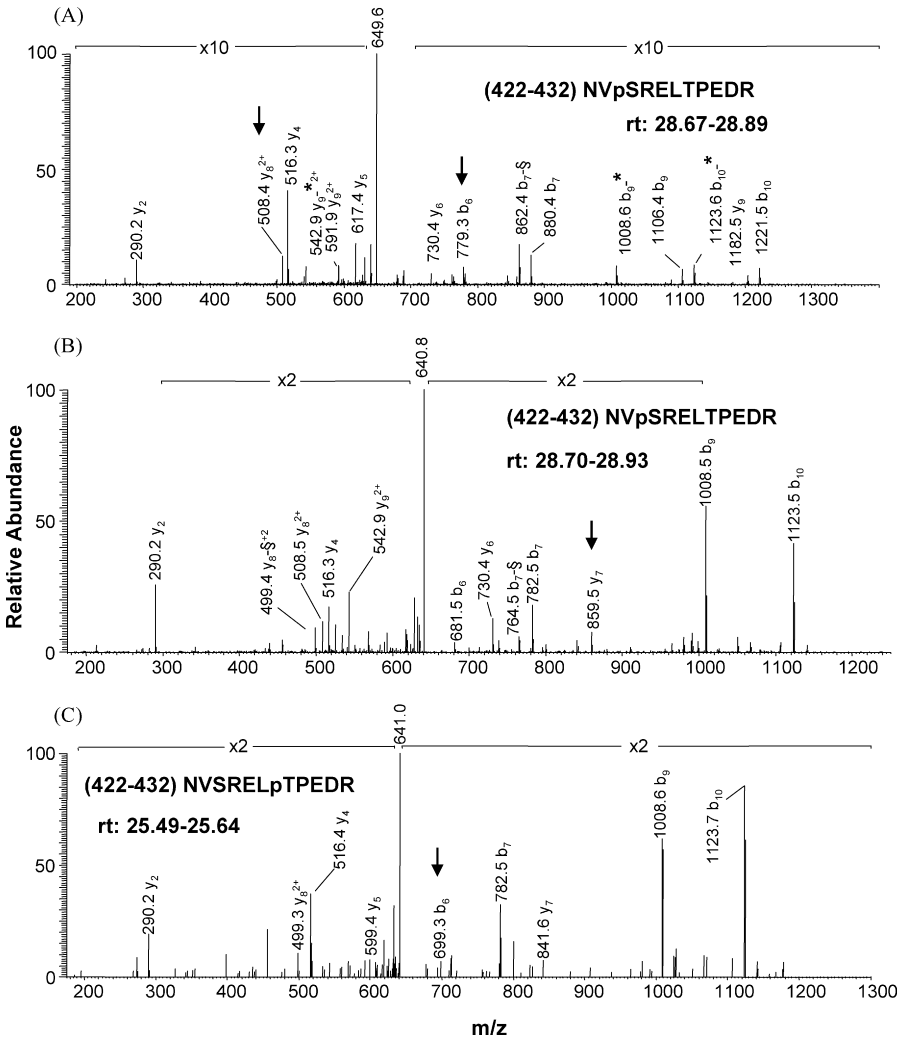


Fig. 3. CID MS/MS and MS^3 spectra of phosphopeptide N_{422}–R_{432} MHV virion N protein. Retention times (rt) for averaged spectra are indicated in each panel. (A) Precursor ion $m/z = 698.3$ was subjected to MS/MS. Peak $m/z = 649.6$ $[M+2H-98]^{2+}$ corresponds to the precursor minus H_3PO_4 and plus $a+2$ charge. Unique ions (arrows) $m/z = 508.4$ (y_8^{2+}) and 779.3 (b_6) localized phosphorylation at S424. (B) MS^3 spectrum of the same peptide eluted at 28.70–28.93 min. Unique ion 859.5 y_7 confirmed phosphorylation at S424. (C) MS^3 spectrum of precursor ion $m/z = 698.3^{2+}$ eluted at 25.49–25.64 min. A unique ion 699.3 b_6 confirmed phosphorylation at T428. Symbol (§) indicates loss of a H_2O molecule (18 amu). Symbol (*) indicates loss of H_3PO_4.

lab which indicated that S389 was phosphorylated (White and Hogue, 2006).

Two other phosphorylated amino acids, S424 and T428, were also identified in the carboxy end of the protein. Phosphorylated S424 was detected in phosphopeptide, $N^{422–432}$ (NVpSRELTPEDR). Precursor ion $m/z = 698.3$ was subjected to both MS/MS and MS^3 fragmentation. The MS/MS spectrum exhibited the most intense peak at m/z 649.6, corresponding to $[M+2H-98]^{2+}$ (Fig. 3A). Ions y_8^{2+} ($m/z = 508.4$) and b_6 ($m/z = 779.3$) are unique ions which localized phosphorylation to serine 424. Observation of y_7 ($m/z = 859.5$) in the MS^3 spectrum was also consistent with phosphorylation at S424 (Fig. 3B).

Two HPLC-MS/MS peaks corresponding to phosphopeptide $N^{422–432}$ were noted. Spectra generated from peptides that contained phosphorylated S424 had retention times between 28.67 and 28.93 min (Fig. 3B), whereas the retention time for the peptides in the second peak was 25.49–25.64 min, roughly 3 min prior to the phosphorylated S424-containing peptides. Spectra of the latter identified T428 as phosphorylated. The fragment ion b_6 ($m/z = 699.3$) in the MS^3 spectra was the unique identifier for phosphorylation at T428 (Fig. 3C).

Additional sites in the amino terminal half of the virion N protein were also identified. The precursor ion for peptide $N^{162–178}$ (pSDIVERDPSSHEAIPTR) was $m/z = 995.0^{2+}$. A peak resulting from the neutral loss of H_3PO_4 was observed as the most intense peak in the MS/MS spectra at $m/z = 946.2^{2+}$ (Fig. 4A). Observation of the appropriate y_{10} ion at $m/z = 1094.5$ in the CID spectrum indicated that S162 was phosphorylated. Further analysis showed that T177 in the 162–178 peptide was also phosphorylated. Peptide $N^{168–178}$ (DPpSSHEAIPTR) resulted from complete digestion of peptide 162–178 at arginine 167. The smaller peptide was fragmented and subjected to MS^3 analysis which allowed us to identify S170 as also phosphorylated (Fig. 4B).

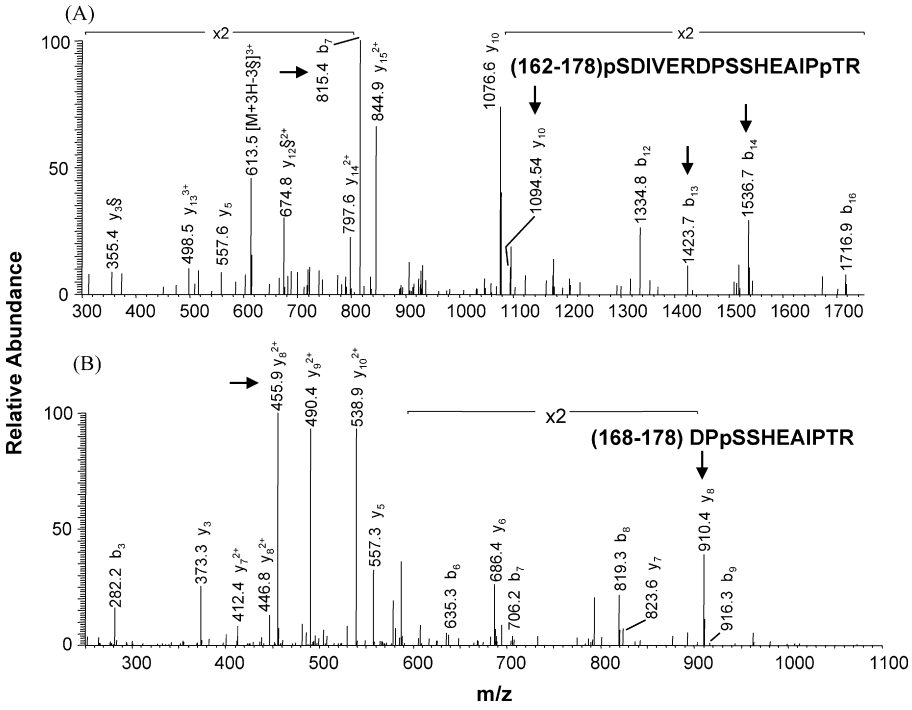


Fig. 4. CID spectra of phosphopeptides in the amino half of virion N protein. (A) MS^3 spectrum of phosphopeptide S_{162}–R_{178} resulting from a missed cleavage at R167 is shown. Unique ion 1094.5 y_{10} indicates phosphorylation at S162. Three unique ions 815.4 b_7, 1423.7 b_{13}, and 1536.7 b_{14} are all indicative of phosphorylation at T177. (B) MS^3 spectrum of the phosphopeptide consisting of residues 168–178 is shown with its signature ions indicated about the peaks. Unique ions (arrows) 455.9 y_8^{2+} and 910.4 y_8 indicate phosphorylation at S170.

3.3. Identification of N protein phosphorylation sites on MHV N protein from infectect cells by HPLC-ESI-MS/MS

To identify residues that are phosphorylated on the intracellular pool of the N protein from infected cells, cytoplasmic lysates were resolved by SDS-PAGE. A sample of purified virions was run on the gel and used as a guide to excise the N protein band from the whole cell lysates (Fig. 2). Mass spectra were again acquired using HPLC-ESI-MS/MS. We obtained 69% sequence coverage (by MASCOT) of the intracellular N protein (Fig. 1B).

In the intracellular N protein phosphopeptide $N^{162–178}$ was also detected. S162 and T177 were both phosphorylated. These phosphopeptides apparently coeluted and were subsequently fragmented together. Fig. 5A is an example a MS^3 spectrum for $N^{162–178}$. Detection of ions 797.5 y_{14}^{2+} and 1094.5 y_{10} is characteristic of phosphorylation at S162, while b_{12} and b_{14} ions (*m/z* 1352.6 and 1536.7, respectively) are consistent with phosphorylation at residue T177. Phosphopeptide $N^{162–178}$ contained one site that was not cleaved by trypsin, but the fully cleaved phosphopeptide, $N^{168–178}$ (DPpSSHEAIPTR), was also detected, which allowed us to localize phosphorylation at S170 as well (Fig. 5B and C).

Further analysis of peptides identified phosphorylation at S389, S424, and T428, as was also determined for the extracellular virion protein (Fig. 6). Identification of S389 was straight forward due to S389 being the only possible target for phosphorylation in peptide $N^{382–394}$ (Fig. 6A). Phosphopeptides containing phosphorylated S424 and T428 coeluted in the intracellular samples, in contrast to the previous separation of these peptides for the extracellular samples. However, MS^3 fragmentation of the phosphopeptide indicated phosphorylation at both S424 and T428 (Fig. 6B and C).

4. Discussion

We used HPLC-ESI/MS/MS to identify the phosphorylated sites on the N protein of MHV A59. This is the first identification of amino acids that are phosphorylated on the N protein for a group II coronavirus. Phosphorylated sites were identified on the protein from infected cells and mature extracellular virions. Six residues, S162, S170, T177, S389, S424 and T428, were found to be phosphorylated on both intracellular and extracellular virion N proteins (Fig. 7). Potential cellular kinases that may phosphorylate the identified sites were predicted using NetPhosK 1.0 (http://www.cbs.dtu.dk/services/NetPhosK/) which produces neural network predictions of kinases specific for eukaryotic phosphorylation sites (Fig. 7) (Blom et al., 2004). It remains to be determined which are responsible for phosphorylations of the N protein, but multiple potential kinases are predicted for all but one of the identified sites.

Our results demonstrate that both serine and threonine residues are modified by phosphorylation, in contrast to earlier studies which suggested that MHV A59, as well as closely related MHV JHM, N proteins are phosphorylated exclusively on serine residues (Siddell et al., 1981; Stohlman and Lai, 1979). The earlier conclusions were made based on phosphoamino acid analysis of [^{32}P]-labeled N peptides. Identification of the phosphorylated threonine residues in our study may reflect differences in the experimental approaches used to isolate and/or analyze the protein. Interestingly, based on comparative analysis

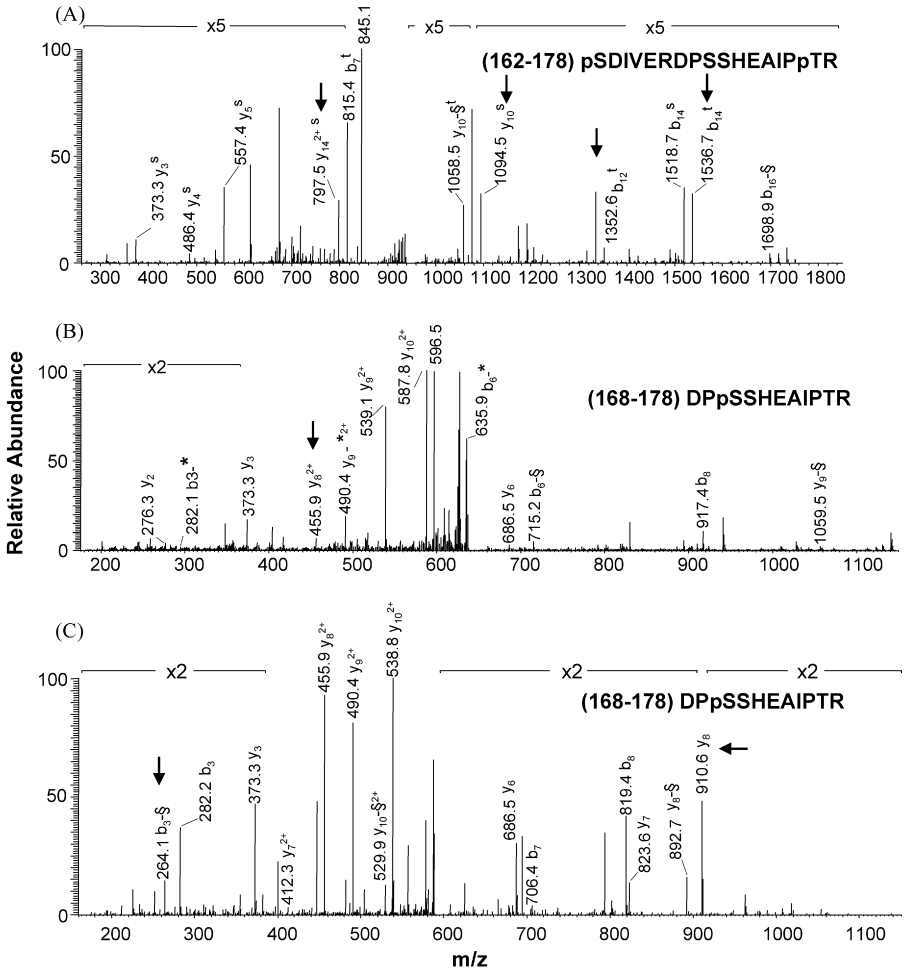


Fig. 5. CID fragmentation spectra of N phosphopeptides from the intracellular fraction of infected cells. (A) MS^3 spectrum of peptide S_{162}–R_{178}. Unique ions (arrows) 797.5 y_{14}^{2+}, and 1094.5 y_{10}, as well as 1352.6 b_{12}, and 1536.7 b_{14} localized phosphorylation to S162 and T177, respectively. Other ions resulting from fragmentation are indicated. (s) and (t) indicates ions derived from either S162 and T177 phosphopeptides, respectively. (B) MS/MS spectrum of peptide N168–178. Neutral loss of H_3PO_4 from the precursor ion mass of 645.28 corresponds to the most intense peak at $m/z = 596.5$. Identification of unique ion (arrows) 455.9 $y_8 + 2$ localized phosphorylation to S170. (C) MS^3 spectrum of phosphopeptide 168–178. Identification of additional unique ions 264.1 b_3-§ and 910.6 y_8 confirmed the localization of phosphorylation at S170. Symbol (§) indicates masses corresponding to a loss of a H_2O molecule. Symbol (*) indicates masses corresponding to a loss of H_3PO_4.

of earlier HPLC of tryptic peptides derived from the N proteins of MHV A59 and a plaque variant of the JHM virus, it was suggested that serine at positions 161 and 162, respectively, were phosphorylated (Wilbur et al., 1986). Thus, S161 identified in the earlier study likely corresponds to S162 identified in our study.

Alignment of the amino acid sequences for several members of the group II coronaviruses shows that, with the exception of S162, the sites are conserved in the bovine coronavirus (BCoV) and human coronavirus OC43 (HCoV OC43) N proteins (Fig. 8). While our work was in progress the phosphorylated sites on the N protein were identified for two other coronaviruses. Phosphoserines at positions 9, 156, 254 and 256 were identified for the group I TGEV protein from virus infected cells, whereas only sites S156 and S256 were identified on the N protein present in purified virions (Fig. 8) (Calvo et al., 2005). The sites that are phosphorylated on the group III IBV N protein expressed with baculovirus in insect cells were found to be identical to the ones on the N protein expressed alone in Vero cells (Chen et al., 2005). Data from the study suggests that S190, S192, T378 and S379 are phosphorylated on intracellular IBV N (Fig. 8). Comparison of the sites for the three viruses illustrates that the location is not conserved, but phosphorylated sites are near either the amino or carboxy side of the serine/arginine (SR) rich domain that is conserved in all coronavirus N proteins. The SR rich domain is a distinguishing feature of the previously identified MHV RNA binding domain (Figs. 7 and 8).

Recently structural information became available for IBV and SARS-CoV N proteins. The three-dimensional structures are based on NMR analysis of amino acids 45–181 (Huang et al., 2004) and X-ray crystal structures of the amino-terminal residues of IBV 19/29-160/162 (Fan et al., 2005; Jayaram et al., 2006) and SARS-CoV 47-175 (Saikatendu et al., 2007). Crystal structures have also been determined for carboxy-terminal residues 219–349 of IBV N (Jayaram et al., 2006) and residues 270–370 of SARS-CoV N (Yu et al., 2006). Based solely on amino acid alignments, this corresponds to the region encompassing roughly R45-N195 in the MHV A59 N protein

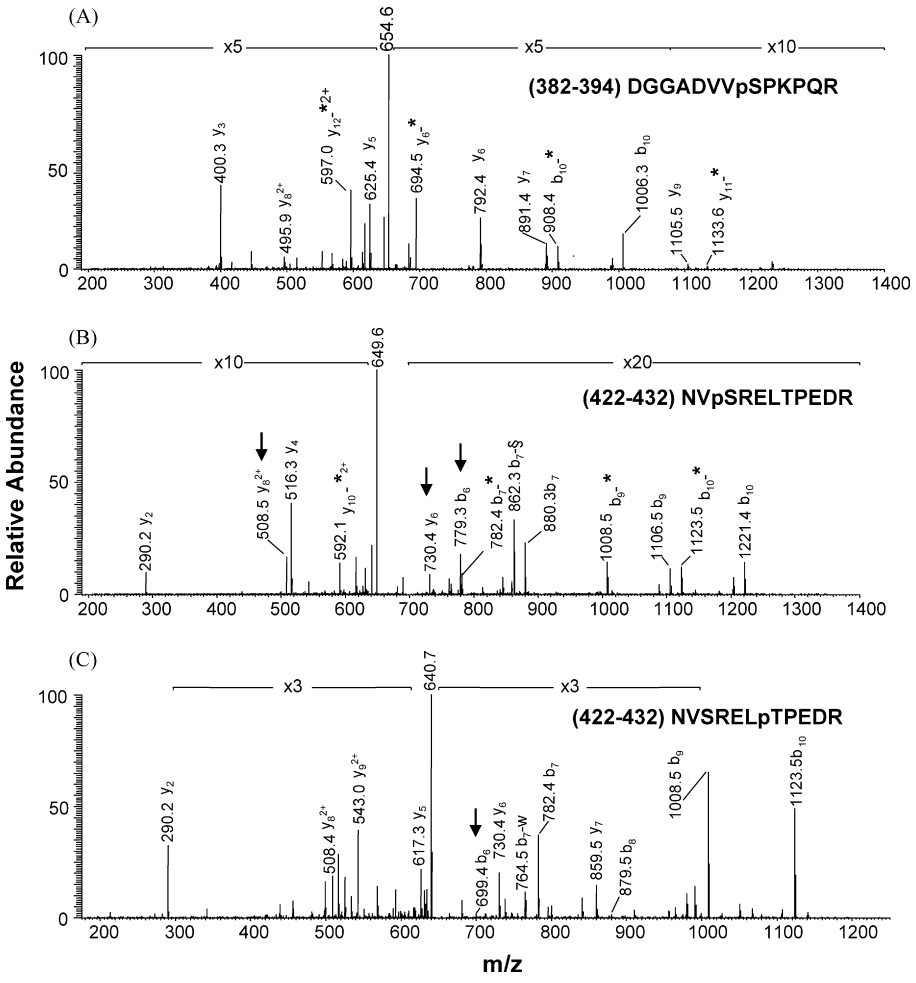
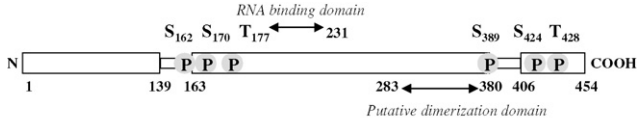


Fig. 6. CID induced MS/MS mass spectrum of peptides containing residues 382–394 and 422–432 from intracellular N. (A) MS/MS spectrum of residues 382–394. Loss of H_3PO_4 from the precursor ion ($m/z = 703.3^{2+}$) correlates with the most intense peak in the spectrum, $m/z = 654.6$ $[M+2H-98]^{2+}$. S389 is the only possible phosphorylation site in the peptide. (B) MS/MS spectrum of the peptide containing residues 422–432. Loss of H_3PO_4 from precursor ($m/z = 698.3^{2+}$) correlates with the most intense peak in the spectrum, $m/z = 649.6$ $[M+2H-98]^{2+}$. Unique ions 508.5 $y_8 + 2$, 730.4 y_6, and 779.3 b_6 identify S424 as phosphorylated. (C) MS^3 fragmentation of the same peptide containing residues 422–432 allowed detection of unique ion 699.4 b_6, indicative of phosphorylation at T428. Symbol (§) indicates loss of a H_2O molecule. Symbol (*) indicates loss of H_3PO_4.



Site	Sequence	Predicted Kinase[a]
S_{162}	TNTRpSDIVE	CKII, CaM-II
S_{170}	ERDPpSSHE	PKA, GSK3, PKG
S_{171}	ERDPSpSHE	PKC, GSK3, CKI, PKC
T_{177}	EAIPpTRFAP	PKC
S_{389}	ADVVpSPKPQ	GSK3, cdk5
S_{424}	QRNVpSRELT	PKA, CaM-II
T_{428}	SRELpTPEDR	GSK3, cdk5, p38MAPK

[a] CKI, II – casine kinase I, II, Ca2+/Calmodulin-dependent protein kinase II, PKC, A, G – protein kinase C, A, G, GSK3 – glycogen synthetase kinase 3, cdk5 – cyclin dependent kinase 5, p38MAPK, mitrogen-activated protein kinase 1

Fig. 7. Schematic illustration of MHV N three-domain model separated by the A and B spacer domains (Parker and Masters, 1990). The relative positions of the phosphorylated sites identified on intracellular and extracellular virion N are shown. The positions of the RNA binding domain (Nelson et al., 2000) and putative dimerization domain (Yu et al., 2006) are noted. Eukaryotic kinases that may phosphorylate the identified sites were predicted by NetPhosK 1.0 (http://www.cbs.dtu.dk/services/NetPhosK/) are shown in the table below the schematic.

(Fig. 8). No structural information is yet available for MHV N, but based on the sequence alignments we can speculate about the potential positions of the phosphorylated sites on MHV N. The sites in the amino terminal region (S162, S170, T177) potentially fall within a large loop between β-strands (β6 and β7 of SARS amino-terminus) (Saikatendu et al., 2007). These sites are adjacent to the previously identified MHV RNA binding domain that extends from T177 to P231 (Figs. 7 and 8) (Nelson et al., 2000). Localization of the phosphorylation sites near this domain could influence presentation of the RNA binding domain and in turn interaction with the RNA. Clearly, much remains to be understood about, not only the structure of the MHV N protein, but also its functional domains compared with other coronavirus nucleocapsids. RNA binding domains for IBV and SARS-CoV are located in the amino-terminal domain preceding the serine-rich domain (Fan et al., 2005; Huang et al., 2004; Saikatendu et al., 2007). Mutagenesis of a positively charged β-hairpin within the amino-terminus of IBV N identified residues involved in RNA-binding in vitro, and when introduced into an infectious clone virus replication

```
MHV      -MSFVPGQENAGGRSSSVNRAGNGILKKTTWADQTERGPNNQNRGRR-NQPKQTATTQ-PNSGSVVPHYSWFSGITQFQKGKEFQ  82
BCV      -MSFTPGKQSSS-RASFGNRSGNGILK---WADQSDQSRNVQTRGRR-AQPKQTATSQLPSGGNVVPYYSWFSGITQFQKGKEFE  79
OC43     -MSFTPGKQSSS-RASSGNRSGNGILK---WADQSDQFRNVQTRGRR-AQPKQTATSQQPSGGNVVPYYSWFSGITQFQKGKEFE  79
SARS     ---MSDNGPQSNQRSAPRITFGG----------PTDSTDNNQNGGRNGARPKQRRPQGLPNNT-----ASWFTALTQHGK-EELR  66
TGEV     MANQGQRVSWGD--------------------------ESTKTRGRSNSRGRKNNNIP----------LSFFNPITLQQGSKFWN  49
229E     ----MATVKWAD--------------------------ASEPQRGR-------QGRIP----------YSLYSPLLVD-SEQPWK  37
IBV      ----MASGKAAG-----------------------KTDAPAPVIKLGGPKPPKVGSSG---------NASWFQAIKAKKLNTPPP  49

MHV      FAEGQGVPIANGIPASEQKGYWYRHNRRSFKTPDGQQKQLLPRWYFYYLGTGPHAGASYGDSIEGVFWVANSQADTNTRSDIVER 167
BCV      FAEGQGVPIAPGVPATEAKGYWYRHNRRSFKTADGNQRQLLPRWYFYYLGTGPHAKDQYGTDIDGVFWVASNQADVNTPADILDR 164
OC43     FAEGQGVPIAPGVPATEAKGYWYRHNRRSFKTADGNQRQLLPRWYFYYLGTGPHAKDQYGTDIDGVYWVASNQADVNTPADIVDR 164
SARS     FPRGQGVPINTNSGPDDQIGYYRRATRR-VRGGDGKMKELSPRWYFYYLGTGPEASLPYGANKEGIVWVATEGALNTPKDHIGTR 150
TGEV     LCPRDFVPKGIG-NRDQQIGYWN--RQTRYRMVKGQRKELPERWFFYYLGTGPHADAKFKDKLDGVVWVAKDGAMNKPTTLGSRG 131
229E     VIPRNLVPINKK-DKNKLIGYWN--VQKRFRTRKGKRVDLSPKLHFYYLGTGPHKDAKFRERVEGVVWVAVDGAKTEPTGYGVRR 119
IBV      KFEGSGVPDNENIKPSQQHGYWR--RQARFKPGKGGRKPVPDAWYFYYTGTGPAADLNWGDTQDGIVWVAAKGADTKSRSNQGTR 132

MHV      DPSSHEAIPT--RFAPGTVLPQGFYVEGSGRSAPASRSGSRSQSRGP--NNRARSSSNQRQPASTVKPDMAEEIAALVLAKLGKD 248
BCV      DPSSDEAIPT--RFPPGTVLPQGYYIEGSGRSAPNSRSTSRASSRASSAGSRSRANSGNRTPTSGVTPDMADQIASLVLAKLGKD 247
OC43     DPSSDEAIPT--RFPPGTVLPQGYYIEGSGRSAPNSRSTSRTSSRASSAGSRSRANSGNRTPTSGVTPDMADQIASLVLAKLGKD 247
SARS     NPNNNAATVL--QLPQGTTLPKGFYAEGS-RGGSQASSRSSSRSRGNSRNSTPGSSRGN-SPARMASGGGETALALLLLDRLNQL 231
TGEV     ANNESKALKFDGKVP-GEFQLEVNQSRDNSRSRSQSRSRSRNRSQSRGRQ---QFNNKKDDSVEQAVLAALKKLGVDTEKQQQR- 211
229E     KNSEPEIPHFNQKLPNGVTVVEEPDSRAPSRSQSRSQSRGRGESKPQSRNPSSDRNHNSQDDIMKAVAAALKSLGFDKPQEKDKK 205
IBV      DPDKFDQYPL--RFSDGGPDGNFRWDFIPLNRGRSGRSTAASSAAASRAPSREGSRGRRSDSGDDLIARA AKIIQDQ------- 207

MHV      AGQP--------------------KQVTKQSAKKVRQKILNKPRQKRTPNK--QCPVQQCFGKRGPNQ---NFGGSEMLKLGTSD 308
BCV      ATKP--------------------QQVTKQTAKEIRQKILNKPRQKRSPNK--QCTVQQCFGKRGPNQ---NFGGGEMLKLGTSD 307
OC43     ATKP--------------------QQVTKHTAKEVRQKILNKPRQKRSPNK--QCTVQQCFGKRGPNQ---NFGGGEMLKLGTSD 307
SARS     ESKVS----------------GKGQQQQGQTVTKKSAAEASKKPRQKRTATK--QYNVTQAFGRRGPEQTQGNFGDQDLIRQGTDY 299
TGEV     ------------------SRSKSKERSNSKTRDTTPKNENKHTWKRTAG---KGDVTRFYGARSSSA---NFGDTDLVANGSSA 271
229E     SAKTGTPKPSRNQSPASSQTSAKSLARSQSSETKEQKHEMQKPRWKRQPNDDVTSNVTQCFGPRDLDH---NFGSAGVVANGVKA 286
IBV      ----------------------QKKGSRITKAKADEMAHRRYCKRTIPP--NYRVDQVFGPRTKGK-EGNFGDDKMNEEGIKD 265

MHV      PQFPILAELAPTVGAFFFGSKLELVKKN--SGGADEPTKDVYELQYSGAVRFDSTLPGFETIMKVLNENLNAYQKDGGADVVSPK 391
BCV      PQFPILAELAPTAGAFFFGSRLELAKVQNLSGNLDEPQKDVYELRYNGAIRFDSTLSGFETIMKVLNENLNAYQQQDGMMNMSPK 392
OC43     PQFPILAELAPTAGAFFFGSRLELAKVQNLSGNPDEPQKDVYELRYNGAIRFDSTLSGFETIMKVLNENLNAYQQQDGMMNMSPK 392
SARS     KHWPQIAQFAPSASAFFGMSRIGMEVTP----------SGTWLTYHGAIKLDDKDPQFKDNVILLNKHIDAYKTFPPTEPKKDK 373
TGEV     KHYPQLAECVPSVSSILFGSYWTSKEDG----------DQIEVTFTHKYHLPKDDP----KTGQFLQQINAYARPSEVAKEQRK 341
229E     KGYPQFAELVPSTAAMLFDSHIVSKESG----------NTVVLTFTTRVTVPKDHP----HLGKFLEELNAFTR------EMQQ 350
IBV      GRVTAMLNLVPSSHACLFGSRVTPKLQL----------DGLHLRFEFTTVVPCDDPQFDNYVKICDQCVDGVGTRPKDDEPKPK 339

MHV      PQRKGRRQAQEKKDEVDNVSVAKPKSSVQRNVSRELTPEDRSLLAQILDDGVVPDGLEDDSNV----------           454
BCV      PQRQ--RGQKNGQGENDNISVAAPKSRVQQNKSRELTAEDISLLKKMDEP-----YTEDTSEI----------           448
OC43     PQRQ--RGHKNGQGENDNISVAVPKSRVQQNKSRELTAEDISLLKKMDEP-----YTEDTSEI----------           448
SARS     KKKT-DEAQPLPQRQKKQPTVTLLPAADMDDFSRQLQNSMSGASADSTQA----------                          422
TGEV     RKSR--------SKSAERSEQDVVPDALIENYTDVFDDTQVEIIDEVTN----------                           382
229E     HPLL--------NPSALEFNPSQTSPATAEPVRDEVS-IETDIIDEVN----------                            391
IBV      SRSS-----SRPATRGNSPAPRQQRPKKEKKLKKQDDEADKALTSDEERNNAQLEFYDEPKVINWGDAALGENEL          409
```

Fig. 8. Alignment of representative coronavirus N protein amino acid sequences. Positions of the phosphorylated sites on MHV N, conserved resides in the group II viruses BCoV, HCoV OC43 and SARS CoV, and recently identified sites on IBV (Chen et al., 2005) and TGEV (Calvo et al., 2005) N proteins are highlighted. The conserved S/R rich domains are boxed in each sequence. Sequences for MHV A59 (P03416), BCoV (P10527), HCov OC43 (Q696P4), SARS CoV (P59595), TGEV (P04134), HCoV 229E (P15130) and IBV CoV (P69596) were obtained from Swiss-Prot. The alignment was generated using ClustalW (Thompson et al., 1994).

was attenuated (Tan et al., 2006). Kinetic analysis of the amino-terminus of IBV N showed that the domain binds viral RNA, but the full length protein exhibits stronger binding, suggesting that interactions with other regions of the protein are also important (Spencer and Hiscox, 2006). The conserved SR domain has been proposed to link the amino-terminal RNA binding and carboxy-terminal multimerization domains of IBV and SAR-CoV (Fan et al., 2005; Saikatendu et al., 2007). All of the identified phosphorylation sites on the IBV N protein are located on the carboxy side of its SR domain. The presence of phosphorylated residues in close proximity to the MHV SR domain may uniquely influence how the surrounding domains interact with viral RNA. Modeling of other coronavirus N proteins based on the crystal forms of SARS-CoV and IBV amino-terminal domains indicate that the proteins are similar in the overall organization of β-strands, but since the proteins differ in their surface charge distribution patterns, it was speculated that the residues involved in RNA binding and how they interact with the RNA is different (Saikatendu et al., 2007). In the case of MHV N, the presence of phosphorylated residues between the amino-end and the SR domain may contribute to this difference.

Only phosphorylated S389 is included in what can be speculated to mirror on the IBV carboxy end structure. Interestingly, this would place the S389 site only a few residues beyond the carboxy terminal α5 helix in the IBV structure (Jayaram et al., 2006). Phosphorylation at this site, as well as the downstream S424 and T428 sites, could be important for multimerization since they fall just beyond the region of MHV N that has been predicted to correspond to the carboxy terminal dimerization domain on the SARS-CoV N structure (Fig. 7) (Yu et al., 2006).

Recent data suggest that IBV N protein binding to viral RNA is influenced by phosphorylation since nonphosphorylated N

protein bound both viral and nonviral RNAs, whereas the phosphorylated N exhibited a higher binding affinity for viral RNAs (Chen et al., 2005). The results are consistent with earlier suggestions for a potential role of N phosphorylation (Laude and Masters, 1995; Nelson et al., 2000). Affinity for viral RNA could play a role in assembly or uncoating. Dephosphorylation of MHV N protein has been suggested to facilitate infection (Mohandas and Dales, 1991). Virion associated kinase activity has been reported for MHV JHM, but whether it plays any role during infection is not known (Siddell et al., 1981). Phosphorylation of N could be important in whatever role the protein plays in viral RNA transcription and/or replication. Phosphorylation may alter the structure of the N protein and in turn presentation of RNA binding domain(s) that is important for recognition of the packaging signal, transcription regulatory sequences or other signature sequences in the viral RNA(s). It was recently demonstrated that TGEV and SARS-CoV N proteins have RNA chaperone activity (Zuniga et al., 2007). We fully agree with the suggestion that phosphorylation may have a role in this activity.

Samples analyzed in this study reflect steady state levels of phosphorylation for the N protein present in virus infected cells at 18 h p.i. We did not detect any difference in phosphorylation of the protein from the infected cells and that in extracellular virions. It should be noted that coronaviruses assemble at intracellular membranes in the region of the endoplasmic reticulum Golgi complex (ERGIC) (Krijnse-Locker et al., 1994). The details of virus release are not fully understood, but virions are thought to transport through the constitutive secretory pathway after budding. Thus, the intracellular fraction in our study no doubt contained some assembled virions that had not been secreted, which would contribute to the profile of sites that were identified. Our identification of the same sites on both intracellular and virion N differs from the results from recent analysis of TGEV N as discussed above. Four sites were found to be phosphorylated on the protein from infected cells, while only two of these were identified on extracellular virion N (Calvo et al., 2005). What accounts for the difference is not known. It may reflect a basic difference between MHV and TGEV. Calvo et al. (2005) suggested the possibility that phosphorylation/dephosphorylation could play a role in TGEV assembly. We previously suggested that a more highly phosphorylated isoform of the N protein exists in BCoV (Hogue, 1995) and MHV infected cells (Hogue, unpublished data). We also hypothesized that nucleocapsids might undergo dephosphorylation during assembly into virions. We did not attempt to identify phosphorylation sites on different isoforms in the present study. Excised gel pieces included both forms of the protein that we previously suggested might be differentially phosphorylated based solely on phosphatase digestion results. Others have reported differences in the phosphorylation status of intracellular and virions N proteins. IBV virion N protein was reported to be more phosphorylated than the protein in virus infected cells (Jayaram et al., 2005).

Our identification of phosphorylation sites for MHV N protein is an important step toward deciphering the functional role(s) of the modification during the virus life cycle. Knowing where phosphorylated sites are located on the protein will help direct future molecular studies to address how the modification contributes to or modulates functions of the multifunctional N protein.

Acknowledgments

The work was supported by Public Health Service AI53704, from the National Institute of Allergy and Infectious Diseases to B.G.H. T.C.W. was supported by National Science Foundation Louis Stokes Alliance for Minority Participation (LSAMP) Bridge to the Doctorate Fellowship as part of Western Alliance to Expand Student Opportunities. Z.Y. was supported by R01 DK47936 to Lawrence J. Mandarino. We thank Dr. Dan Brune in the ASU Proteomics Lab and Drs. Randall W. Nelson and Eric E. Niederkofler at Intrinsic Bioprobes in Tempe, AZ for their help during the initial phase of this work and members of the Hogue Lab for helpful discussions throughout the study.

References

Almazan, F., Galan, C., Enjuanes, L., 2004. The nucleoprotein is required for efficient coronavirus genome replication. J. Virol. 78 (22), 12683–12688.

Baric, R.S., Nelson, G.W., Fleming, J.O., Deans, R.J., Keck, J.G., Casteel, N., Stohlman, S.A., 1988. Interactions between coronavirus nucleocapsid protein and viral RNAs: implications for viral transcription. J. Virol. 62 (11), 4280–4287.

Blom, N., Gammeltoft, S., Brunak, S., 1999. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J. Mol. Biol. 294 (5), 1351–1362.

Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S., Brunak, S., 2004. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics 4 (6), 1633–1649.

Calvo, E., Escors, D., Lopez, J.A., Gonzalez, J.M., Alvarez, A., Arza, E., Enjuanes, L., 2005. Phosphorylation and subcellular localization of transmissible gastroenteritis virus nucleocapsid protein in infected cells. J. Gen. Virol. 86 (Pt 8), 2255–2267.

Casais, R., Thiel, V., Siddell, S.G., Cavanagh, D., Britton, P., 2001. Reverse genetics system for the avian coronavirus infectious bronchitis virus. J. Virol. 75 (24), 12359–12369.

Chang, R.Y., Brian, D.A., 1996. cis Requirement for N-specific protein sequence in bovine coronavirus defective interfering RNA replication. J. Virol. 70 (4), 2201–2207.

Chen, H., Gill, A., Dove, B.K., Emmett, S.R., Kemp, C.F., Ritchie, M.A., Dee, M., Hiscox, J.A., 2005. Mass spectroscopic characterization of the coronavirus infectious bronchitis virus nucleoprotein and elucidation of the role of phosphorylation in RNA binding by using surface plasmon resonance. J. Virol. 79 (2), 1164–1179.

Cologna, R., Spagnolo, J.F., Hogue, B.G., 2000. Identification of nucleocapsid binding sites within coronavirus-defective genomes. Virology 277 (2), 235–249.

Compton, S.R., Rogers, D.B., Holmes, K.V., Fertsch, D., Remenick, J., McGowan, J.J., 1987. In vitro replication of mouse hepatitis virus strain A59. J. Virol. 61 (6), 1814–1820.

Davies, H.A., Dourmashkin, R.R., Macnaughton, M.R., 1981. Ribonucleoprotein of avian infectious bronchitis virus. J. Gen. Virol. 53 (Pt 1), 67–74.

Denison, M.R., Spaan, W.J., van der, M.Y., Gibson, C.A., Sims, A.C., Prentice, E., Lu, X.T., 1999. The putative helicase of the coronavirus mouse hepatitis virus is processed from the replicase gene polyprotein and localizes in complexes that are active in viral RNA synthesis. J. Virol. 73 (8), 6862–6871.

Escors, D., Ortego, J., Laude, H., Enjuanes, L., 2001. The membrane M protein carboxy terminus binds to transmissible gastroenteritis coronavirus core and contributes to core stability. J. Virol. 75 (3), 1312–1324.

Fan, H., Ooi, A., Tan, Y.W., Wang, S., Fang, S., Liu, D.X., Lescar, J., 2005. The nucleocapsid protein of coronavirus infectious bronchitis virus: crystal

structure of its N-terminal domain and multimerization properties. Structure 13 (12), 1859–1868.

Hogue, B.G., 1995. Bovine coronavirus nucleocapsid protein processing and assembly. Adv. Exp. Med. Biol. 380, 259–263.

Huang, Q., Yu, L., Petros, A.M., Gunasekera, A., Liu, Z., Xu, N., Hajduk, P., Mack, J., Fesik, S.W., Olejniczak, E.T., 2004. Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. Biochemistry 43 (20), 6059–6063.

Hurst, K.R., Kuo, L., Koetzner, C.A., Ye, R., Hsue, B., Masters, P.S., 2005. A major determinant for membrane protein interaction localizes to the carboxy-terminal domain of the mouse coronavirus nucleocapsid protein. J. Virol. 79 (21), 13285–13297.

Jayaram, H., Fan, H., Bowman, B.R., Ooi, A., Jayaram, J., Collisson, E.W., Lescar, J., Prasad, B.V., 2006. X-ray structures of the N- and C-terminal domains of a coronavirus nucleocapsid protein: implications for nucleocapsid formation. J. Virol. 80 (13), 6612–6620.

Jayaram, J., Youn, S., Collisson, E.W., 2005. The virion N protein of infectious bronchitis virus is more phosphorylated than the N protein from infected cell lysates. Virology 339 (1), 127–135.

Kopecky-Bromberg, S.A., Martinez-Sobrido, L., Frieman, M., Baric, R.A., Palese, P., 2007. SARS coronavirus proteins ORF 3B, ORF 6, and nucleocapsid function as interferon antagonists. J. Virol. 81 (2), 548–557.

Krijnse-Locker, J., Ericsson, M., Rottier, P.J., Griffiths, G., 1994. Characterization of the budding compartment of mouse hepatitis virus: evidence that transport from the RER to the Golgi complex requires only one vesicular transport step. J. Cell Biol. 124 (1–2), 55–70.

Laude, H., Masters, P.S., 1995. The Coronavirus Nucleocapsid Protein, vol. 7, pp. 141–163.

Macnaughton, M.R., Davies, H.A., Nermut, M.V., 1978. Ribonucleoprotein-like structures from coronavirus particles. J. Gen. Virol. 39 (3), 545–549.

Mohandas, D.V., Dales, S., 1991. Endosomal association of a protein phosphatase with high dephosphorylating activity against a coronavirus nucleocapsid protein. FEBS Lett. 282 (2), 419–424.

Narayanan, K., Kim, K.H., Makino, S., 2003. Characterization of N protein self-association in coronavirus ribonucleoprotein complexes. Virus Res. 98 (2), 131–140.

Narayanan, K., Makino, S., 2001. Characterization of nucleocapsid-M protein interaction in murine coronavirus. Adv. Exp. Med. Biol. 494, 577–582.

Nelson, G.W., Stohlman, S.A., Tahara, S.M., 2000. High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA. J. Gen. Virol. 81 (Pt 1), 181–188.

Parker, M.M., Masters, P.S., 1990. Sequence comparison of the N genes of five strains of the coronavirus mouse hepatitis virus suggests a three domain structure for the nucleocapsid protein. Virology 179 (1), 463–468.

Saikatendu, K.S., Joseph, J.S., Subramanian, V., Neuman, B.W., Buchmeier, M.J., Stevens, R.C., Kuhn, P., 2007. Ribonucleocapsid formation of SARS-CoV through molecular action of the N-terminal domain of N protein. J. Virol. 81, 3913–3921.

Siddell, S.G., Barthel, A., ter, M.V., 1981. Coronavirus JHM: a virion-associated protein kinase. J. Gen. Virol. 52 (Pt 2), 235–243.

Spencer, K.A., Hiscox, J.A., 2006. Characterisation of the RNA binding properties of the coronavirus infectious bronchitis virus nucleocapsid protein amino-terminal region. FEBS Lett. 580 (25), 5993–5998.

Stohlman, S.A., Lai, M.M., 1979. Phosphoproteins of murine hepatitis viruses. J. Virol. 32 (2), 672–675.

Tahara, S.M., Dietlin, T.A., Bergmann, C.C., Nelson, G.W., Kyuwa, S., Anthony, R.P., Stohlman, S.A., 1994. Coronavirus translational regulation: leader affects mRNA efficiency. Virology 202 (2), 621–630.

Tan, Y.W., Fang, S., Fan, H., Lescar, J., Liu, D.X., 2006. Amino acid residues critical for RNA-binding in the N-terminal domain of the nucleocapsid protein are essential determinants for the infectivity of coronavirus in cultured cells. Nucleic Acids Res. 34 (17), 4816–4825.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22 (22), 4673–4680.

van der Meer, M.Y., Snijder, E.J., Dobbe, J.C., Schleich, S., Denison, M.R., Spaan, W.J., Locker, J.K., 1999. Localization of mouse hepatitis virus nonstructural proteins and RNA synthesis indicates a role for late endosomes in viral replication. J. Virol. 73 (9), 7641–7657.

Verma, S., Bednar, V., Blount, A., Hogue, B.G., 2006. Identification of functionally important negatively charged residues in the carboxy end of mouse hepatitis coronavirus A59 nucleocapsid protein. J. Virol. 80 (9), 4344–4355.

White, T.C., Hogue, B.G., 2006. Mouse hepatitis coronavirus nucleocapsid phosphorylation. Adv. Exp. Med. Biol. 581, 157–160.

Wilbur, S.M., Nelson, G.W., Lai, M.M., McMillan, M., Stohlman, S.A., 1986. Phosphorylation of the mouse hepatitis virus nucleocapsid protein [published erratum appears in Biochem Biophys Res Commun 1986 December 15, 141(2), 884]. Biochem. Biophys. Res. Commun. 141 (1), 7–12.

Ye, Y., Hauns, K., Langland, J.O., Jacobs, B.L., Hogue, B.G., 2007. Mouse Hepatitis Coronavirus A59 Nucleocapsid Protein is a Type I Interferon Antagonist. J. Virol. 81 (6), 2554–2563.

Yount, B., Curtis, K.M., Baric, R.S., 2000. Strategy for systematic assembly of large RNA and DNA genomes: transmissible gastroenteritis virus model. J. Virol. 74 (22), 10600–10611.

Yu, I.M., Oldham, M.L., Zhang, J., Chen, J., 2006. Crystal structure of the severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein dimerization domain reveals evolutionary linkage between corona- and arteriviridae. J. Biol. Chem. 281 (25), 17134–17139.

Zuniga, S., Sola, I., Moreno, J.L., Sabella, P., Plana-Duran, J., Enjuanes, L., 2007. Coronavirus nucleocapsid protein is an RNA chaperone. Virology 357 (2), 215–227.