EXPRESSION AND STRUCTURAL ANALYSIS OF INFECTIOUS BRONCHITIS VIRUS NUCLEOPROTEIN

Kelly-Anne Spencer and Julian A. Hiscox*

1. INTRODUCTION

The coronavirus nucleoprotein (N protein) is one of the most abundantly expressed viral proteins in an infected cell, with the principal function of binding the viral RNA genome to form the ribonucleocapsid structure (RNP) and forming the viral core. N protein also has roles in viral replication, transcription, and translation as well as modulating cellular processes. Although coronavirus N proteins have the potential to be phosphorylated at multiple serine residues, mass spectroscopic analysis of both the avian infectious bronchitis virus (IBV) and porcine transmissible gastroenteritis virus (TGEV) have shown that N protein is phosphorylated at only three or four residues.[1,2] In the case of IBV N protein, these map to predicted casein kinase II sites.[2] Based on amino acid sequence comparisons, three conserved regions have been identified in the murine coronavirus, mouse hepatitis virus (MHV) N protein.[3] In general, other coronavirus N protein would appear to follow this pattern. Of the three regions, for MHV, region two has been shown to bind both coronavirus and non-coronavirus RNA sequences, whereas regions one and three for IBV N protein have been shown to bind to RNA.[4]

We hypothesize that phosphorylation of N protein controls the RNA binding activity and differential phosphorylation of the protein alters its structure to expose necessary binding motif(s). In support of this hypothesis, antibodies studies suggested that phosphorylation of N protein led to substantial conformation change,[5] and using surface plasmon resonance we have recently shown that phosphorylation of IBV N protein was involved in the recognition of viral RNA from nonviral RNA.[2]

To investigate this hypothesis, we have developed methodologies to determine whether the overall conformation of N protein changes upon binding viral RNA and what regions may be involved in this process. Our technique is to express either wild-type IBV N protein or its three subregions (termed NI, NII, and NIII) and measure changes in structure upon RNA binding using circular dichroism (CD). CD spectroscopy is a form of light adsorption spectroscopy that measures the difference in absorbance of right- and

*University of Leeds, Leeds, LS2 9JT, United Kingdom.

left-circularly polarized light. In proteins, the major chromophores are the amide bonds of the peptide backbone and the aromatic side chains. Polypeptides and proteins have regions where the peptide chromophores are in highly ordered arrays; as a consequence, many common secondary structure motifs, such as the α-helix, ß-pleated sheet, ß-turn and random coil, have very characteristic CD spectra.[6] CD also allows the detection of gross protein conformational changes; it can therefore be utilized to monitor changes in secondary structure upon ligand binding, multimer formation, and to analyze the protein in a variety of environments. CD was used to analyze structural changes that may occur when IBV N protein binds to models of viral RNA and to map any conformational changes to specific regions of the protein. Here we report our initial studies optimizing this system and our preliminary results.

2. MATERIALS AND METHODS

2.1. IBV N Protein Expression and Purification

IBV N protein and the three regions were expressed in Tuner (DE3) pLacI IPTG inducible *E. coli* to produce recombinant protein possessing an N-terminal hexa-histidine tag. Cell lysate containing recombinant protein was purified using nickel chelating chromatography; purified proteins were eluted using increasing concentrations of imidazole. Further purification was performed by separating proteins according to their size using gel filtration chromatography.

2.2. SDS-PAGE and Western Blot Analysis

Protein purity was analyzed by separation on NuPage Bis-Tris 10% pre-cast SDS-PAGE gels (Invitrogen), and proteins were visualized by staining with Coomassie. Proteins were transferred onto PVDF membrane for Western blot analysis and detected using ECL (Amersham/Pharmacia) according to the manufacturer's instructions.

2.3. Circular Dichroism Spectroscopy

CD experiments were performed on a Jasco J715 spectrophotometer. Measurements were taken in the far-UV (190–260 nm) and the CD signal recorded in a 1-mm path-length cell using a protein concentration of 0.2 mg/ml. Protein samples to be analyzed were dialyzed into sodium phosphate buffer, pH7.2. RNA was added at a 1:1 molar ratio of protein to RNA.

3. RESULTS AND DISCUSSION

3.1. Extraction and Purification of IBV N Protein Regions

IBV N regions were purified as described above. However, due to the differing properties of each protein, some subtle changes were made. NI is readily expressed and soluble in large quantities, with recoveries of the order of 10 mg obtained per 500 ml of

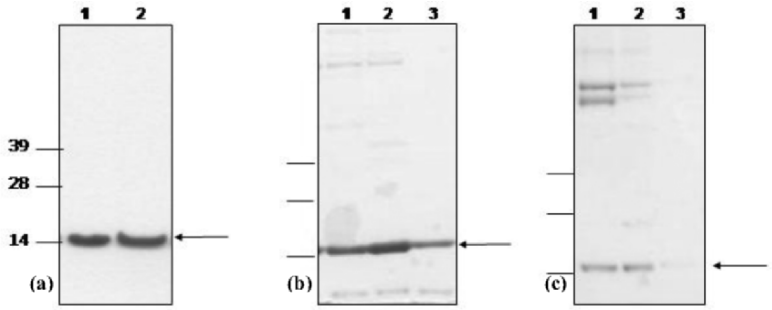

Figure 1. Coomassie-stained SDS-PAGE analysis of IBV N protein regions. (a) Purified region NI obtained by eluting protein from Ni^{2+} column using 150nm imidazole (lane 1) or 200 mM imidazole (lane 2). (b) Purified region NII obtained by eluting protein from Ni^{2+} column using 150nM imidazole (lane 1), 200 mM imidazole (lane 2), and by gel filtration chromatography (lane 3). (c) Purified region NIII obtained by eluting protein from Ni^{2+} column using 150 nm imidazole (lane 1), 200 mM imidazole (lane 2), and by gel filtration chromatography (lane 3). Apparent molecular weight markers (kDa) are shown to the left. Each IBV N protein region is indicated by an arrow.

culture. When purifying NI, high percentage purity was achieved from nickel chelation chromatography alone (Fig. 1a) eliminating the need for gel filtration chromatography. However, both NII and NIII could only be expressed at low levels. Relatively pure (approximately 80%) NII could be produced at lower concentrations (up to 800 µg/ml) after a two-step chromatography strategy, although some low molecular weight contaminants remained after gel filtration (Fig. 1b). NIII was purified using nickel chelating chromatography, after which the protein readily precipitates out of solution making it unsuitable for subsequent purification and dialysis steps (Fig. 1c). Western blotting was performed with antibody specific for IBV to verify the identity of the purified proteins (data not shown).

3.2. Effect of RNA Binding on IBV N Protein Structure

Using CD, we compared the structural changes of IBV N upon binding two models of the IBV genome. The first was RNA synthesized by runoff transcription from pCD-61 (generously provided by Dr. Paul Britton and Dr. Dave Cavanagh) to generate a 6.1-kb analogue of the IBV genome, and the second a synthetic RNAmer that was identical to the 5' end of IBV mRNA 3 up to, and including, the translation initiation codon for gene 3a; both of these targets were used in previous binding studies.[2] Various concentrations of IBV N protein were analyzed in the far UV and averaged. Due to the nature of CD it is difficult to assign definitive structures, however, when delineating the spectrum of IBV N protein the peak between 190 and 200 nm is indicative of ß-sheet and the dip at 220 nm can be due to α-helical secondary structures (Fig. 2). Other features of the spectrum, including the trough at around 210 nm, are consistent with random coil motifs. Upon the addition of CD61 or leader RNA, subtle changes can be seen in the overall shape of the spectra, in particular the negative signal at 210 nm becomes weaker and a stronger negative trough can be seen between 230 and 250 nm (Fig. 2). These changes occur upon binding both RNA's, therefore the 5' leader RNA was used for all further CD analysis.

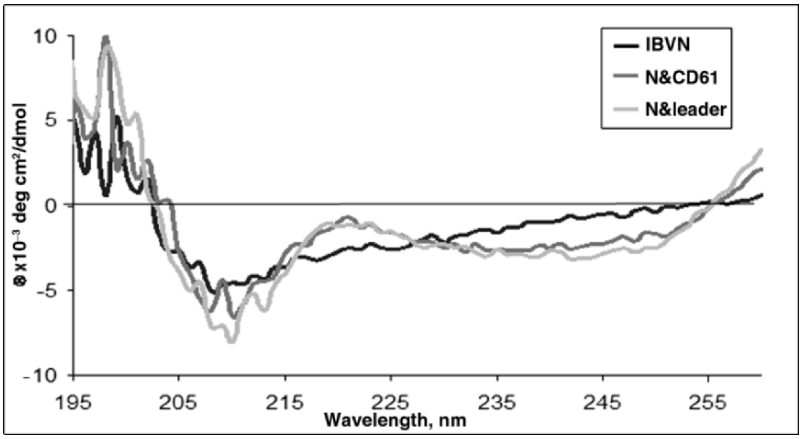

Figure 2. Purified recombinant IBV N protein was analyzed using far-UV circular dichroism spectroscopy. Measurements were taken of IBV N protein without the addition of RNA (black line), in the presence of CD61, an IBV genomic RNA analogue (gray line), and a synthetic RNA model of the 3' leader sequence (light gray line).

3.3. Structural Changes Can Be Mapped to Specific Regions of IBV N

IBV N has a molecular weight of approximately 45 kDa. It is therefore likely that it contains more than one type of secondary structure motif, such as a mixture of α-helices and ß-sheets, or ß-sheets involving ß-turns, together these motifs generate "noisy" spectra. In order to overcome this problem, IBV N was broken down into its three regions, NI, NII, and NIII, and analyzed by CD. The far UV CD-spectrum of NI, the N-terminal region of IBV N protein, can be seen in Fig. 3a. The curve is consistent with a random coil with the major trough at 210 nm. Upon the addition of leader RNA, no change can be seen in the overall shape of the spectrum, indicating that the secondary structure of IBV NI doesn't alter in the presence of viral RNA. NII, the central 15 kDa of IBV N, conversely undergoes a substantial conformational change in the presence of RNA. The far UV CD-spectrum of NII protein in the absence of RNA has a trough at around 215 nm which is indicative of random coil or α-helical secondary structures. Addition of leader RNA results in changes in the spectrum, namely the strong negative trough at 215 nm is lower in intensity and the peak at 235 nm shifts along nearer to 220 nm; these changes suggest that the protein-RNA complex may contain more ß-turns and sheets than that of protein alone (Fig. 3b). NIII, the C-terminal region of IBV N, was analyzed in the far UV. Due to the nature of this protein and problems with purity, strong CD signals proved difficult to obtain, and the resulting spectrum are therefore less definitive and assigning secondary structure problematic. As can be seen on the NIII spectrum (Fig 3c), a strong negative signal is obtained upon the addition of viral RNA,

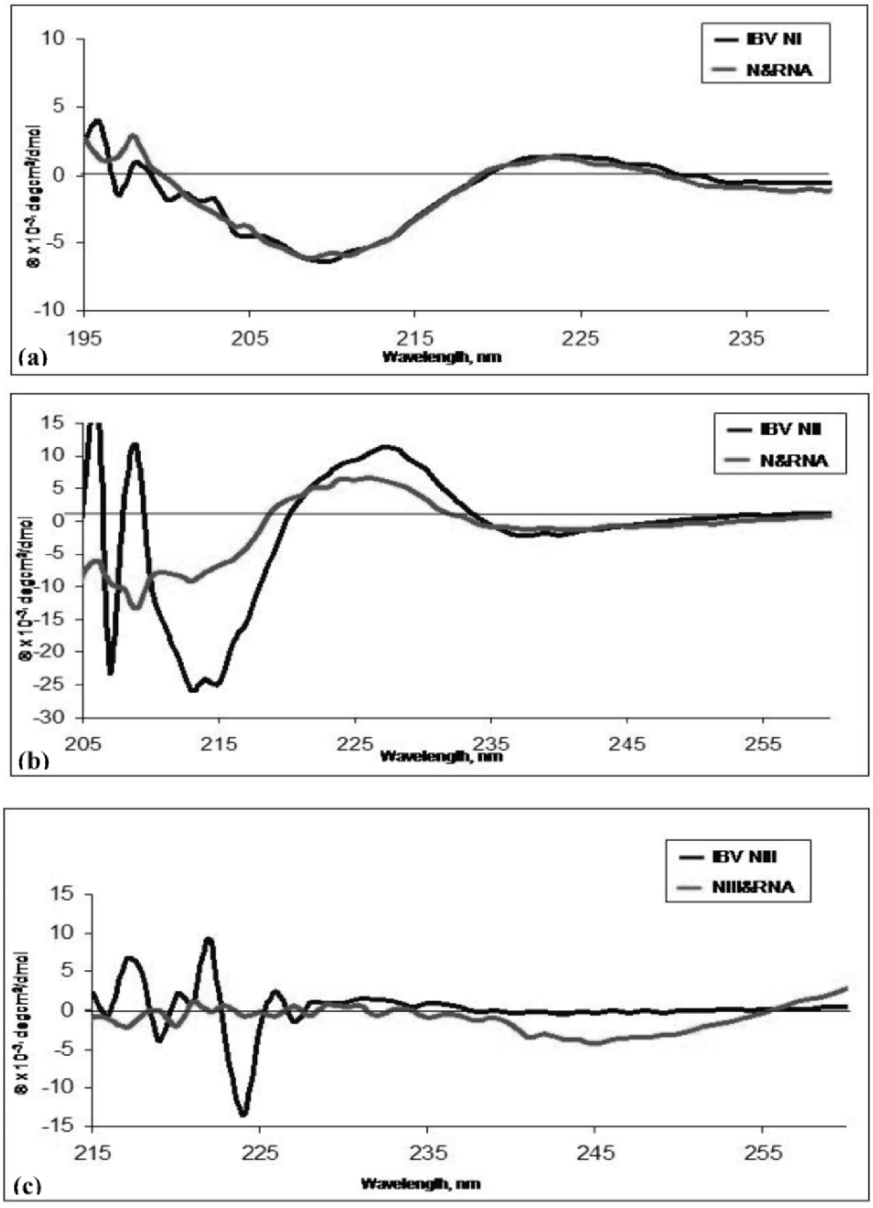

Figure 3. Far-UV spectra of IBV N regions (a) IBV NI, (b) IBV NII, (c) IBV NIII. Black line: protein without the addition of RNA, gray line: protein in a 1:1 ratio with viral RNA.

indicating that this interaction may stabilize the protein and prevent it from precipitating out of solution. The overall spectrum shows some features consistent with ß-sheet motifs, such as the negative signal between 230 and 250 nm, and the spectrum obtained is very weak in intensity, a feature often observed for ß-sheet proteins.[7]

In conclusion, our data would support the hypothesis that N protein undergoes conformational change upon binding viral RNA. These changes would appear to be confined to regions II and III, where RNA binding motifs may be located. Certainly our mass spectroscopic study indicated that regions II and III contained two phosphorylation sites each, whereas region I contained no such sites.

4. ACKNOWLEDGMENTS

This work was supported by the BBSRC and Guildhay Ltd. J.A.H. and K.A.S. would like to thank Prof. Sheena Radford and Dr. Sue Jones for their help and expertise with CD spectroscopy.

5. REFERENCES

1. E. Calvo, D. Escors, J. A. Lopez, J. M. Gonzalez, A. Alvarez, E. Arza, and L. Enjuanes, Phosphorylation and subcellular localization of transmissible gastroenteritis virus nucleocapsid protein in infected cells, *J. Gen. Virol.* **86**, 2255-2267 (2005).
2. H. Chen, A. Gill, B. K. Dove, S. R. Emmett, F. C. Kemp, M. A. Ritchie, M. Dee, and J. A. Hiscox, Mass spectroscopic characterisation of the coronavirus infectious bronchitis virus nucleoprotein and elucidation of the role of phosphorylation in RNA binding using surface plasmon resonance, *J. Virol.* **79**, 1164-1179 (2005).
3. P. S. Masters, Localization of an RNA-binding domain in the nucleocapsid protein of the coronavirus mouse hepatitis virus, *Arch. Virol.* **125**, 141-160 (1992).
4. M. L. Zhou and E. W. Collinson, The amino and carboxyl domains of the infectious bronchitis virus nucleocapsid protein interact with 3' genomic RNA, *Virus Res.* **67**, 31-39 (2000).
5. S. A. Stohlman, J. O. Fleming, C. D. Patton, and M. M. C. Lai, Synthesis and subcellular-localization of the murine coronavirus nucleocapsid protein, *Virology* **130**, 527-532 (1983).
6. N. J. Greenfield, Methods to estimate the conformation of proteins and polypeptides from circular dichroism data, *Anal. Biochem.* **235**, 1-10 (1996).
7. R. W. Woody, Circular dichroism, *Methods Enzymol.* **246**, 34-71 (1995).