Delineation and Modelling of a Nucleolar Retention Signal in the Coronavirus Nucleocapsid Protein

Mark L. Reed1, Brian K. Dove1, Richard M. Jackson1,2, Rebecca Collins1, Gavin Brooks3 and Julian A. Hiscox1,2*

1Institute of Molecular and Cellular Biology, University of Leeds, Leeds, UK
2Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, UK
3School of Pharmacy, University of Reading, Reading, UK
*Corresponding author: Julian A. Hiscox, j.a.hiscox@leeds.ac.uk

Unlike nuclear localization signals, there is no obvious consensus sequence for the targeting of proteins to the nucleolus. The nucleolus is a dynamic subnuclear structure which is crucial to the normal operation of the eukaryotic cell. Studying nucleolar trafficking signals is problematic as many nucleolar retention signals (NoRSs) are part of classical nuclear localization signals (NLSs). In addition, there is no known consensus signal with which to inform a study. The avian infectious bronchitis virus (IBV), coronavirus nucleocapsid (N) protein, localizes to the cytoplasm and the nucleolus. Mutagenesis was used to delineate a novel eight amino acid motif that was necessary and sufficient for nucleolar retention of N protein and colocalize with nucleolin and fibrillarin. Additionally, a classical nuclear export signal (NES) functioned to direct N protein to the cytoplasm. Comparison of the coronavirus NoRSs with known cellular and other viral NoRSs revealed that these motifs have conserved arginine residues. Molecular modelling, using the solution structure of severe acute respiratory (SARS) coronavirus N-protein, revealed that this motif is available for interaction with cellular factors which may mediate nucleolar localization. We hypothesise that the N-protein uses these signals to traffic to and from the nucleolus and the cytoplasm.

Key words: nuclear export signal, nucleocapsid protein, nucleolar retention signal, nucleolar targeting, nucleolus, virus

Received 8 November 2005, revised and accepted for publication 10 March 2006, published online 25 May 2006

The nucleolus is a dynamic subnuclear structure involved in ribosome subunit biogenesis, in RNA processing, in cell cycle control and as a sensor for cell stress (1–4). Morphologically, the nucleolus can be divided into fibrillar centre(s) (FC), a dense fibrillar component (DFC) and an outer granular component (GC). A directed proteomic analysis, followed by subsequent bioinformatic analysis revealed that the nucleolus is composed of at least 700 proteins (5–7). Whilst the rules and signals governing the nuclear localization and nuclear export of proteins are well defined, those concerning nucleolar localization/retention are not. In the case of nucleolar localization signals (NLSs), these can be classified into several categories. The majority of motifs identified thus far such as ’pat4’, ’pat7’ and bipartite signals are composed of basic amino acids within a given sequence length (8–10). Protein nuclear export signals (NESs) again vary, but one of the most common and well characterized is an approximately 11 amino acid leucine-rich signal, typified by LxxxLxlLxxL, where a number of hydrophobic amino acids can substitute for L, and the spacer regions (x) can vary in number (9,11).

In contrast to the NLSs and NESs, nucleolar localization/retention signals (NoRSs) and pathways are not well characterized (12), and the signals can vary, but are usually rich in arginine and lysine, although there is no obvious consensus. For example, the MQRKPTIRRNLRLRRK motif identified in survivin-deltaEx3 protein (13) and the RSRKYGWAVALKR motif of the 18-kDa fibroblast growth factor-2 (14). Nucleolar localization can also be regulated by binding accessory proteins, such as nucleostemin binding to GTP (15). Compared with NESs where the leucine-rich export signal (for CRM-1 type nuclear export receptors) is accessible for interaction with carrier proteins (16), the structural context of a NoRS is not well characterized. In many cases, proteins localizing to both the cytoplasm/nucleus/nucleolus contain multiple signals to determine their subcellular localization (17–20). This highlights the difficulty in identifying NoRSs, in that, many proteins which localize to the nucleolus also localize to the nucleus and contain both classical nuclear and nucleolar signals which can also overlap (14,21,22).

We use the avian infectious bronchitis virus (IBV), coronavirus nucleocapsid (N) protein, as a model to study nucleolar retention. We have previously shown that N protein is present in the cytoplasm but can also actively localize to the nucleolus (23–25) and interact with the nucleolar proteins nucleolin and fibrillarin (26). Coronaviruses can also interact with host cell processes, such as the cell cycle (27,28) and signal transduction pathways (29,30). Coronaviruses together with the closely related arteriviruses are nidoviruses, a group of positive strand RNA viruses, which cause a variety of different diseases (31,32). For example, IBV and severe acute respiratory syndrome coronavirus (SARS-CoV) both cause respiratory disease (33,34), whereas murine coronaviruses can cause hepatitis and demyelination (35) and porcine coronaviruses, respiratory and gastroenteric diseases (36).
IBV (Beaudette strain) N protein, composed of 409 amino acids with a predicted molecular weight of 45 kDa, is a phosphoprotein which can bind viral RNA with high affinity (37) and also modulate cellular processes (24,26). On the basis of amino acid sequence comparison, three conserved regions (1, 2 and 3) have been identified in the coronavirus N protein, which in the case of IBV N protein can map to the N-terminal (amino acids 1–133), central (amino acids 134–265) and C-terminal parts of the protein (amino acids 266–409), respectively (38). Also, mass spectroscopy revealed that conserved phosphorylation sites are present in regions 2 and 3 of both avian and porcine coronavirus N proteins (37,39). Given the subcellular localization of IBV N protein to the nucleolus, but not nucleus, we hypothesized that the protein would contain a unique NoRS and possibly a signal for export of the protein into the cytoplasm. Therefore, using a combination of deletion and substitution mutagenesis, coupled with live cell imaging and confocal microscopy, we have tested this prediction.

Results

Bioinformatic analysis and preliminary molecular investigation of nuclear import and export signals in IBV N protein

To identify whether there were potential NLSs (which could form part of a NoRS) and/or NESs in IBV N protein, we first conducted a bioinformatic analysis of the protein using existing motif prediction algorithms. PredictNLS (8) and PSORTII (40) were used to identify potential NLSs, and the NES predictor (NetNES) was used (16) to identify potential NESs. PredictNLS found no NLSs, whereas PSORTII indicated that IBV N protein contained two potential overlapping NLSs in region 3 between residues 358–366, a pat4 motif (RPKK) and a pat7 motif (PKKEKKKL) (Figure 1A). NetNES predicted a potential CRM-1-dependent NES between residues 291 and 298 (LQLDGLHL) (also in region 3) (Figure 1A). To investigate whether these and other unknown signals operated to determine the subcellular trafficking of N protein, we divided the protein into three regions [based on conservation between IBV strains and other coronavirus N proteins (38)] and combinations of regions 1 plus 2 and 2 plus 3 and cloned these downstream of enhanced cyan fluorescent fusion protein (ECFP) generating plasmids pECFP-IBVNR1+2, pECFP-IBVNR2+3, pECFP-IBVNR1, pECFP-IBVNR2 and pECFP-IBVNR3. When expressed in cells, these would lead to the synthesis of ECFP fused to regions 1 plus 2, 2 plus 3 and expressed individually, regions 1, 2 and 3, respectively (Figure 1B). Vero cells, a model cell line to study IBV-cell interactions (24,26,41,42), were transfected with plasmids pECFP-IBVNR1+2, pECFP-IBVNR2+3, pECFP-IBVNR1, pECFP-IBVNR2 and pECFP-IBVNR3. Recombinant fusion proteins were imaged at 24 h post-transfection using live cell imaging (direct fluorescence). As a control, cells were also transfected with pECFP-C1, which leads to the expression of ECFP only and also wild-type N protein cloned downstream of ECFP, plasmid pEGFP-IBVN, as described (25) (Figure 2). In a parallel series of experiments, cells were cotransfected with pECFP-IBVNR1+2, pECFP-IBVNR2+3, pECFP-IBVNR1, pECFP-IBVNR2 and pECFP-IBVNR3 and pEGFP nucleolin (Figure 2). This latter construct allowing the expression of a nucleolar marker protein, nucleolin, tagged to EGFP (25). Dual-transfected cells were fixed at 24 h post-transfection for confocal analysis by direct fluorescence.

Live cell imaging indicated that as previously shown, ECFP localized to both the cytoplasm and nucleus, but not the nucleolus, whereas EGFP-IBVN protein localized to both the cytoplasm and nucleolus but not the nucleus, as described previously (23,25). ECFP-IBVNR1+2 protein localized to the nucleus and nucleolus, whereas ECFP-IBVNR2+3 protein was predominantly cytoplasmic in localization. Confocal analysis of dual-transfected cells was used to confirm the presence of the nucleolus using the marker protein, EGFP nucleolin, and reflected the live cell imaging results (Figure 2). The data also indicated that ECFP-IBVNR1+2 colocalized with nucleolin but not ECFP-IBVNR2+3. Live cell analysis of the single region constructs (pECFP-IBVNR1, pECFP-IBVNR2 and pECFP-IBVNR3) indicated that ECFP-IBVNR1 localized predominantly to the nucleolus and also localized to the nucleus, ECFP-IBVNR2 localized predominantly to the nucleus and appeared also to accumulate in the nucleolus to the same level as the nucleus, whereas ECFP-IBVNR3 localized predominantly to the cytoplasm. Confocal analysis confirmed these findings and indicated that ECFP-IBVNR1 colocalized with nucleolin, whereas ECFP-IBVNR2 was indeterminate and ECFP-IBVNR3 did not (Figure 2).

This data suggested that a potential NoRS could be located in region 1, that region 2 contained a potential NLS(s) not identified by the bioinformatic analysis and also a possible NoRS. The data also suggested that because region 3 was directed to the cytoplasm, the putative NES was dominant to the predicted NLSs. None of these fusion proteins had a distribution similar to ECFP only. Therefore the potential of regions 1 and 2 to direct nucleolar localization and region 3 to promote nuclear export was investigated.

Delineation of a NoRS in region 1 of N protein

Having identified that region 1 of IBV N protein localized to the nucleolus, the hypothesis was tested that it contained a NoRS to target the protein to the nucleolus. To investigate this prediction, a series of expression constructs containing fragments of region 1 were constructed, based on the rationale of keeping sequences of basic and non-basic amino acids discrete, pECFP-IBVNR1–50, pECFP-IBVNR151–100 and pECFP-IBVNR1101–132 (Figure 3A). Vero cells were transfected with these constructs and analyzed by live cell imaging at 24 h post-transfection. The data indicated that ECFP-IBVNR11–50 and ECFP-
IBVNR1\textsubscript{101–132} had a similar localization pattern to ECFP only, whereas ECFP-IBVNR\textsubscript{151–100} localized predominantly to the nucleus and nucleolus (Figure 3B). These data were confirmed by cotransfected cells with either pECFP-IBVNR\textsubscript{1–60}, pECFP-IBVNR\textsubscript{151–100} or pECFP-IBVNR\textsubscript{101–132}, and pEGFP-nucleolin, fixed at 24 h post-transfection and analyzed using confocal microscopy (Figure 3B). The data also indicated that ECFP-IBVNR\textsubscript{151–100} colocalized with nucleolin, whereas the other two fusion proteins did not. Therefore, the fragment containing amino acids 51–100 of IBV N protein could direct an exogenous protein to the nucleolus.

To further refine the amino acids involved in nucleolar retention, 20 amino acid overlapping motifs encompassing amino acids 61–100 were cloned downstream of ECFP, creating plasmids pECFP-IBVNR\textsubscript{161–80}, pECFP-IBVNR\textsubscript{171–90} and pECFP-IBVNR\textsubscript{181–100} for the expression of recombinant fusion proteins (Figure 4A). Amino acids 51–60 were not included in this analysis as there were no arginine and lysine residues present in this sequence and basic amino acids form part of known NoRSs. Vero cells were transfected with pECFP-IBVNR\textsubscript{161–80}, pECFP-IBVNR\textsubscript{171–90} and pECFP-IBVNR\textsubscript{181–100} and analyzed 24 h post-transfection using live cell imaging. Also, cells were cotransfected with these constructs and pEGFP-nucleolin, fixed at 24 h post-transfection and analyzed using confocal microscopy (Figure 4B). The data indicated that ECFP-IBVNR\textsubscript{171–90} and ECFP-IBVNR\textsubscript{181–100} localized to the nucleus and nucleolus (and colocalized with nucleolin). However, ECFP-IBVNR\textsubscript{181–100} localized to the cytoplasm and nucleus but not the nucleolus.
(and did not colocalize with nucleolin). Therefore, the amino acids at positions 61–90 in IBV N protein were able to direct an exogenous protein to the nucleus/nucleolus.

To further define the amino acids involved in nucleolar trafficking, we conducted a tetra-alanine substitution mutagenesis of amino acids 71–90. These were placed downstream of ECFP, creating expression plasmids, pECFP-IBVNR1_71WRRQAAAA, pECFP-IBVNR1_75ARFKAAAA, pECFP-IBVNR1_79PGKGAADA, pECFP-IBVNR1_83GKPAAAA and pECFP-IBVNR1_87VPDAAAAA. Therefore, in some cases, the wild-type alanine was not substituted. Amino acids 61–70 were excluded from the substitution analysis as no basic residues were present. These expression plasmids were transfected into Vero cells and the distribution of the respective fluorescent fusion proteins analyzed by live cell imaging at 24 h post-transfection (Figure 5A). The data indicated that substituting 71WRRQ with AAAA (pECFP-IBVNR1_71WRRQAAAA) abolished nucleolar retention of the recombinant fusion protein and that substitution of 75ARFK with AAAA (pECFP-IBVNR1_75ARFKAAAA) resulted in reduced nucleolar retention. The remaining tetra-alanine substitutions had no effect on nucleolar retention, indicating that amino acids 71WRRQARFK78 were involved in nucleolar retention.
To test whether this amino acid sequence was involved in directing the nucleolar localization of N protein, this motif was deleted in the context of full length N protein tagged to EGFP (plasmid pEGFP-IBV\textsubscript{N} \textsubscript{NoRS}). This plasmid was transfected into Vero cells and the subcellular localization of the resulting fusion protein EGFP-IBV\textsubscript{N} \textsubscript{NoRS} investigated using live cell imaging. There was no nucleolar localization at 24 h post-transfection (Figure 5B, several examples are shown) compared with approximately 50% in cells expressing EGFP-IBV\textsubscript{N} protein (data not shown). This data also indicated that the NoRS identified in region 1 was necessary for nucleolar retention in IBV N protein. As described, the preliminary investigation using the single and double region constructs indicated a potential NoRS in region 2. If the latter signal was functional, then we would have expected a proportion of cells expressing EGFP-IBV\textsubscript{N} \textsubscript{NoRS} in the nucleolus. However, no nucleolar localization was observed, indicating that the WRRQARFK motif was necessary for nucleolar localization.

To determine whether the eight amino acid sequence identified in region 1 was sufficient to direct nucleolar retention, this motif was placed downstream of ECFP and DsRed, creating vectors pECFP-WRRQARFK and pDsRed-WRRQARFK, respectively. As controls, an amino acid sequence C-terminal of the region 1 NoRS, GRKPVPDA, identified in the tetra-alanine substitution mutagenesis as not being involved in nucleolar retention, was placed downstream of ECFP, creating vector pECFP-GRKPVPDA. Transfection of Vero cells with pECFP-WRRQARFK and pECFP-GRKPVPDA indicated that WRRQARFK directed ECFP to the nucleolus and nucleus, whereas GRKPVPDA did not (Figure 6A). In the case of the former construct, relative fluorescence revealed that there was approximately fourfold more ECFP in the nucleolus than in the nucleus. To investigate the role of both WRRQ and ARFK in nucleolar targeting, each motif was substituted for alanine and placed downstream of ECFP, creating expression vectors pECFP-WRRQA\textsubscript{AAA}ARFK and pECFP-AAAAARFK. Transfection and expression of
these constructs in Vero cells (Figure 6A) indicated that whilst WRRQ could direct ECFP to the nucleus and nucleolus, ARFK was required for efficient nucleolar retention. In contrast, ECFP-AAAAARFK by itself localized predominantly to the cytoplasm, weakly to the nucleus and was absent from the nucleolus (Figure 6A).

To investigate whether the WRRQARFK motif targeted a specific part of the nucleolus, cells were cotransfected with pDsRed-WRRQARFK and either pEGFP-nucleolin or pEGFP-fibrillarin (Figure 6B, several examples are shown). These fusion proteins provided distinct markers for the nucleolus (25). The data indicated that WRRQARFK could also direct DsRed to the nucleolus (as well ECFP, Figure 6A) and that WRRQARFK tagged to the appropriate fluorescent protein colocalized predominately with EGFP-fibrillarin and EGFP-nucleolin and formed a punctate appearance in the nucleolus, which we tentatively defined as the DFC. Note that DsRed-WRRQARFK was also observed in the nucleus and cytoplasm, but the images are resolved in the linear range for the nucleolar signal, which is the predominant component of this localization.

Investigation of potential nucleolar targeting signals in IBV N protein region 2
Similar to the approach used to identify the NoRS in region 1, region 2 was subdivided into two distinct components. Amino acids 133–200 and 201–265 were placed downstream of ECFP, creating expression vectors pECFP-IBVN133–200 and pECFP-IBVN123–200 (Figure 7A). Expression of these plasmids in Vero cells indicated that amino acids 133–200 directed ECFP to the cytoplasm and nucleus and had a subcellular localization similar to ECFP. In contrast, amino acids 201–265 directed ECFP to the nucleus with no evidence of nucleolar exclusion. Further investigation revealed that amino acids 201–220 and 211–230 when fused to ECFP directed this protein to the nucleus and nucleolus, whereas amino acids 221–240 directed ECFP to the nucleus but not to the nucleolus (Figure 7B). Relative fluorescence indicated that the ratio of ECFP between the nucleus and nucleolus with amino acids 201–220 and 211–220 was approximately 1:1, which is in contrast to that observed with the NoRS identified in region 1, in which there was four times more signal in the nucleolus than the nucleus. Taken together with the lack of nucleolar retention observed in cells expressing the region 1 NoRS deletion mutant, in the context of full length N protein, we propose that WRRQARFK is necessary and sufficient to direct N protein to the nucleolus, and no other signals are involved.

Comparison of the IBV N protein NoRS with known cellular and viral NoRS
To investigate whether our identified NoRS was unique or had a mammalian cellular or viral equivalent, we conducted a search using the basic local alignment search
tool (BLAST) (http://www.ncbi.nlm.nih.gov/BLAST/). No known mammalian cellular or non-avian coronavirus viral protein sequences were highlighted. Using AlignX, a consensus NoRS was derived by comparing the IBV N protein NoRS with known cellular and viral NoRS which have been shown to target exogenous proteins to the nucleolus (Figure 8). The consensus sequence was composed of two basic groups, both of which were arginine rich. Compared with this consensus sequence, analysis revealed that the arg72 and arg73 of the IBV N protein NoRS were conserved, arg76 and lys78 were similar and Gln74 were weakly similar.

**Figure 5:** (A) Amino acid sequence of tetra-alanine substitution mutagenesis of pECFP-NR1\textsubscript{71-90}, the position of the substituted amino acids is underlined. In some cases, the substitution maintained the wild-type alanine and analysis of the subcellular localization of these peptides (ECFP-IBVNR1\textsubscript{71WRRQ→AAAA}, ECFP-IBVNR1\textsubscript{75ARFK→AAAA}, ECFP-IBVNR1\textsubscript{79PGKG→AAAA}, ECFP-IBVNR1\textsubscript{83GRKP→AAAA} and ECFP-IBVNR1\textsubscript{87VPDA→AAAA}) using live cell imaging of Vero cells at 24 h post-transfection. (B) Live cell imaging of cells expressing EGFP-IBVNoRS and corresponding bright field images. Several examples are shown. Scale bar is 10 μm, and the nucleolus (No) is arrowed where appropriate.
Delineation of a leucine-rich NES in IBV N protein

Bioinformatic analysis of N protein indicated that a predicted NES was located between amino acids 291–298, LQLDGLHL. To test whether this motif functioned as an NES, these amino acids were deleted in the context of wild-type N protein-fused C-terminal of ECFP, creating plasmid pECFP-IBVN<sup>D</sup><sub>291</sub>/C0<sub>298</sub>. As a control, amino acids 268–275 (VTAMLNLV), encompassing a hydrophobic region, were deleted in the context of wild-type N protein-fused C-terminal of ECFP, creating plasmid pECFP-IBVN<sup>D</sup><sub>268</sub>/C0<sub>275</sub>. Vero cells were transfected with these plasmids and analyzed by live cell imaging at 24 h post-transfection. The data indicated that the predicted NES deletion mutant (ECFP-IBVN<sup>D</sup><sub>291</sub>/C0<sub>298</sub>) localized predominately to the nucleus and nucleolus, whereas the control deletion (ECFP-IBVN<sup>D</sup><sub>268</sub>/C0<sub>275</sub>) had no apparent effect on the subcellular localization of the protein (Figure 9A) when compared to wild-type N protein-fused C-terminal of EGFP (Figures 2 and 9A) (23,25,43).

Given that the bioinformatic analysis indicated that region 3 contained potential NLSs, we investigated...
whether the NES was dominant to these signals. The above NES deletion was made in the context of the region 3 fusion protein, pECFP-IBVNR3, creating plasmid pECFP-IBVNR3 for the expression of recombinant fusion protein. Live cell imaging of this protein in Vero cells at 24 h post-transfection indicated that when the NES was deleted (Figure 9B) region 3 had a similar localization pattern to ECFP only (Figure 2), in that the fragment localized to the nucleus and cytoplasm but not nucleolus. This data indicated that despite containing two predicted NLSs, region 3 did not accumulate in the nucleus or nucleolus, as would be predicted if it contained such active signals.

To investigate the relative importance of leucine residues in nuclear export, appropriate alanine substitutions were made either in the context of the NES in region 3 only or in the context of wild-type N protein. To determine whether residues 293L and 298L were involved in nuclear export, these positions were substituted for alanine both individually and together in the context of ECFP fused to region 3 (pECFP-IBVNR3), creating plasmids, pECFP-IBVNR3293L→A, pECFP-IBVNR3298L→A and pECFP-IBVNR3293/298L→A (respectively). These plasmids were transfected into Vero cells and the subcellular localization of the resulting fusion proteins analyzed by direct fluorescence using live cell imaging (Figure 9B). The data
indicated that none of these changes affected the distribution of the fusion protein, suggesting that these amino acids were not involved in nuclear export. Amino acids 291L was substituted for alanine in the context of EGFP-IBVN, creating plasmid pEGFP-IBVN

Expression of this fusion protein in Vero cells and analysis using relative fluorescence indicated an increased level of N protein in the nucleus (Figure 9C) when compared with expression of the wild-type N protein (Figure 2), thus suggesting that position 291L is involved in nuclear export. This is in contrast with EGFP-IBVN

where there is no apparent difference in the localization to wild-type N protein (Figure 2).

Discussion

NoRSs are not well characterized, and we have made use of the avian IBV coronavirus N protein to study these. To investigate whether IBV N protein contained a NoRS, initially the protein was expressed as a series of single and overlapping regions. This preliminary analysis indicated that IBV N protein contained a NoRS in region 1. Deletion mutagenesis delineated a 20 amino acid motif that modulated nucleolar retention. Subsequent tetra-alanine substitution mutagenesis highlighted that four residues were crucial for the targeting function, 71WRRQ, with residues 75ARFK also promoting retention. The role of this novel octa-peptide in nucleolar localization was confirmed by deletion mutagenesis in the context of full length N protein and by placing the motif C-terminal of ECFP andDsRed. These latter constructs also colocalized with nucleolin and fibrillarin, suggesting that the WRRQARFK motif directs a protein to the DFC. Certainly, IBV N protein has been shown to localize to the DFC (43).

This is the first description of a defined NoRS in a coronavirus N protein which localizes to the nucleolus. Although region 2 could localize ECFP to the nucleolus, the NoRS identified in region 1 resulted in nucleolar accumulation of fluorescent fusion proteins when comparing the ratio of protein in the nucleus versus the nucleolus. Region 2, although containing no predicted NLSs (and subcellular localization motifs), localized predominately to the nucleus. When region 2 was fused to region 3, ECFP localized predominately to the cytoplasm, suggesting that if a NLS is present, it is submissive to the NES in region 3.

Studies have suggested that phosphorylation can control nucleolar retention of certain proteins (44). However, mass spectroscopic analysis revealed that no phosphorylated amino acids are present in region 1 (37), and therefore, we hypothesize that phosphorylation plays no role in the specific nucleolar targeting

![Figure 8: AlignX analysis of the IBV N protein NoRS with known cellular and viral NoRSs which can target an exogenous protein to the nucleolus.](image-url)
activity. However, in full length, protein conformational changes may be induced by phosphorylation and/or cleavage to expose relevant motifs to direct the protein to appropriate subcellular localizations. Bioinformatic analysis found that some of the most abundant motifs within the nucleolar proteome were the RNA-recognition motif and the DEAD/H box helicase domain (6). Although the non-phosphorylated form of IBV N protein binds cellular RNA with high affinity (37), this protein contains no known cellular RNA-binding motifs. In addition, phosphorylated N protein (which is present inside the cell) has low affinity for non-viral RNA (37). Sequence analysis reveals that region 3 of IBV N protein contains the amino acid sequence 371DEAD. However, the helicase activity of this protein is unknown. In addition, region 3 does not localize to the nucleolus, even when the NES was deleted. Also, other tri-peptide or tetra-peptide motifs containing either GR or RG motifs are over-represented in the nucleolar proteome (6) but not represented in the IBV N protein sequence. Therefore, we propose that WRRQARFK motif acts exclusively as the NoRS for IBV N protein.
The three dimensional structure of IBV N protein region 1 was modelled based on the solution structure of the equivalent region of SARS-CoV N protein (45). Amino acid sequence alignment between the two regions (Figure 10A) and superimposition of the model on the structure of SARS-CoV N protein (Figure 10B) indicated that a credible molecular model of IBV N protein region 1 could be generated by comparative modelling methods (Figure 10C). Coupled to the tetra-alanine amino acid substitution mutagenesis analysis and subcellular localization of ECFP-WRRQARFK and DsRed-WRRQARFK to the nucleolus and comparison with cellular and viral NoRSs, the data indicated that Arg72 and Arg73 were present at the bottom of a pocket and were thus accessible for interaction with cellular factors which may be involved in nucleolar targeting/retention.

Although the focus of this study was to elucidate whether IBV N protein contained a NoRS, a functional NES was also identified in region 3. Interestingly, region 3 also contained two predicted NLSs, which could either have been submissive to the NES or not functional. Such basic regions are found in the C-terminal regions of other coronavirus N proteins and in the case of SARS-CoV proposed to act as an NLS (46). However, subsequent experimental data showed that this sequence was non-functional (25,47). Deletion mutagenesis of the IBV NES in the context of a region 3 peptide again suggested that the region 3 NLS was non-functional in IBV. Our current and previous studies (23,24) indicated that by fusing EGFP/ECFP with IBV N protein increased the molecular weight of this protein above the size exclusion limit of the nuclear pore complex. The fusion peptides generated in this study are close to the size exclusion limit (50–60 kDa) (9). The region 3 fusion protein (~43 kDa) is below the size exclusion limit and could theoretically diffuse into and out of the nucleus. However, as discussed above, the peptide localized to the cytoplasm, again indicating the presence of a functional NES.

Similar to other RNA viruses, IBV replication is error prone and the genome also undergoes recombination (48,49). Therefore, due to selection pressure, there is sequence variability causing the formation of different IBV strains (50), which has led to different clinical outcomes, such as predominately respiratory or nephropathogenic disease (34). This variability is reflected in the amino acid sequence of the N protein, with presumably essential amino acids remaining unchanged or being conserved. Comparison of the Beaudette strain N protein amino acid sequence (used in this study) with nine other strains of IBV including a nephropathogenic strain (accession numbers are shown in square brackets) Ark99 [M85244], DE072 [AF203001], M41 [M28566], N2/75 [U52598], N1/62 [U52596], N9/74 [U52597], QXIBV [AF199412], KB8523 [M21515] and LX4 (AAC21592) revealed that the essential NoRS 71 WRRQ motif was identical in all strains, as was amino acid 78K. With regard to the NES located between amino acids 291–298, amino acids 291L and 296L were identical, whereas 293L was not, and 298L was identical in all but one strain, where it was substituted for a valine (and thus was conserved). Substitution of appropriate leucine residues in the NES would support the role of the conserved leucines being the functional in nuclear export.

---

**Figure 10:** (A) Sequence alignment of IBV N protein region 1 to SARS-CoV N protein region 1. Identical amino acids are indicated by an asterisk and the position of the tetra-alanine substitutions on IBV N protein are indicated by coloured boxes which correspond to their position on the three-dimensional model. (B) Mapping the modelled structure of the IBV N protein region 1 (red) on the solution structure of SARS-CoV protein region 1 (white). (C) Three-dimensional model of region 1 of IBV N protein, the position of the tetra-alanine substitutions, is indicated by the appropriate colour as shown in (A).
Analysis of the SARS-CoV N protein revealed that although it contains a potential NoRS(s) which has been mapped to region 2 (25), the protein localizes to the cytoplasm only in infected cells (25,47). However, it can localize to the nucleus/nucleolus when expressed in the absence of other viral proteins (51,52), albeit with low frequency when compared with IBV N protein (25). Unlike IBV N protein, SARS-CoV N protein does not possess a recognizable CRM-1-dependent NES but instead contains an uncharacterized cytoplasmic retention/NES region in region 3 (25).

Coronaviruses are closely related to arteriviruses, and although the arterivirus N protein has no discernable homology to the coronavirus N protein, and has a molecular weight of 15 kDa, several arterivirus N proteins have been shown to localize to the nucleolus (53,54). In the case of the arterivirus porcine reproductive and respiratory syndrome virus (PRRSV) N protein, two NLSs were identified, one which directed the protein to the nucleus and one which directed the protein to the nucleus and nucleolus (53). In comparison to the IBV N protein NoRS, the PRRSV NoRS was lysine rather than arginine rich. Bioinformatic analysis using NetNES predictor revealed no obvious NES in this protein, although the equine arteritis virus N protein is sensitive to leptomycin B treatment (54).

An emerging paradigm is that both plant and animal positive-strand RNA viruses can interact with nucleus, the nucleolus and nucleolar proteins to recruit factors to aid in virus replication and/or subvert host cell function (65–63). If viral proteins target subnuclear structures such as the nucleolus, then they must contain appropriate signalling motifs, not only for localization but crucially for export back to the cytoplasm. If such proteins were to be retained in the nucleolus, then we would predict that virus replication would be down-regulated, as the principal site of replication for positive-strand RNA viruses is the cytoplasm. Disrupting the efficiency of nucleolar localization/nuclear export of viral proteins may therefore be a way of attenuating the replication of HIV-1. Certainly, the replication of HIV-1 can be inhibited by interfering with the interaction of HIV-1 with the nucleolus (64,65). In summary, IBV N protein contains an eight amino acid motif which is necessary and sufficient for nucleolar retention and a functional NES to traffic the protein to the cytoplasm.

Materials and Methods

Cell culture

Vero (monkey-derived kidney epithelial) cells were grown at 37 °C with 5% CO2 in minimum Eagles media (MEM) supplemented with 10% foetal calf serum and penicillin/streptomycin as described previously (24). The N gene from the Beaudette strain (accession number: AAA46214 of IBV served as a template for PCR of region and subregion constructs. Primers used incorporated T7Xhol site and 3’-SacII site for cloning into pECFP-C1 (enhanced cyan fluorescent protein) (Clontech, Palo Alto, CA, USA). At all times, numbers used in primer or construct names denoting amino acid numbers refer to their position on the full-length N protein. Primers used for the constructions were as follows: pECFP-NR1 (amino acids 1–133), forward primer GCCGGTCTCTGCAAGCGGACTTTGGCTTACATGGTGTCTTGTAATCCCTGTGATGTC; pECFP-NR2 (amino acids 133–265), forward primer GCCGGTCTCTGCAAGCGGACTTTGGCTTACATGGTGTCTTGTAATCCCTGTGATGTC and reverse primer GCCGGTCTCTGCAAGCGGACTTTGGCTTACATGGTGTCTTGTAATCCCTGTGATGTC and ECFP-NR3 (amino acids 265–409), forward primer GCCGGTCTCTGCAAGCGGACTTTGGCTTACATGGTGTCTTGTAATCCCTGTGATGTC and reverse primer GCCGGTCTCTGCAAGCGGACTTTGGCTTACATGGTGTCTTGTAATCCCTGTGATGTC.

Construction of plasmids

To produce the full-length inserts used for delineation of the NoRS in N1, the following oligonucleotides comprising restriction overhangs: (5´, XhoI; 3´, SacII) pECFP-C1 (enhanced cyan fluorescent protein) (Clontech, Palo Alto, CA, USA). Primers used incorporated 5´-XhoI site and 3´-SacII site for cloning into pECFP-C1 (enhanced cyan fluorescent protein) (Clontech, Palo Alto, CA, USA). At all times, numbers used in primer or construct names denoting amino acid numbers refer to their position on the full-length N protein. Primers used for the constructions were as follows: pECFP-NR1 (amino acids 1–133), forward primer GCCGGTCTCTGCAAGCGGACTTTGGCTTACATGGTGTCTTGTAATCCCTGTGATGTC; pECFP-NR2 (amino acids 133–265), forward primer GCCGGTCTCTGCAAGCGGACTTTGGCTTACATGGTGTCTTGTAATCCCTGTGATGTC and reverse primer GCCGGTCTCTGCAAGCGGACTTTGGCTTACATGGTGTCTTGTAATCCCTGTGATGTC and ECFP-NR3 (amino acids 265–409), forward primer GCCGGTCTCTGCAAGCGGACTTTGGCTTACATGGTGTCTTGTAATCCCTGTGATGTC and reverse primer GCCGGTCTCTGCAAGCGGACTTTGGCTTACATGGTGTCTTGTAATCCCTGTGATGTC.

Materials and Methods

Cell culture

Vero (monkey-derived kidney epithelial) cells were grown at 37 °C with 5% CO2 in minimum Eagles media (MEM) supplemented with 10% foetal calf serum and penicillin/streptomycin as described previously (24).

Traffic 2006; 7: 833–848
To investigate whether the octa-peptide was sufficient to direct nuclear localization, the signal and appropriate control peptides were placed C-terminal of ECFP and/or DsRed by generating overlapping oligonucleotides with 5´-XhoI and 3´-SacII restriction overhangs for direct ligation into either pECFP-C1 or pDsRed-C1 which had been digested with the appropriate restriction enzymes. WRRQARFK was placed C-terminal of ECFP and DsRed, creating pECFP-WRRQARFK and pDsRed-WRRQARFK, forward primer TCGAGCCATGTTACAAGCGCAACGGCCATGTTGCGTACCTCGCC and reverse primer GGCTACTTAAACCTGGCGGCGGCGGCCCATGGCC. The same cloning strategy was used to generate pEEFP-GRKPVFD (forward primer TCGAGCCATGTTACAAGCGCAACGGCCATGTTGCGTACCTCGCC and reverse primer GGCTACTTAAACCTGGCGGCGGCGGCCCATGGCC), pEEFP-WRRQAAA (forward primer TCGAGCCATGTTACAAGCGCAACGGCCATGTTGCGTACCTCGCC and reverse primer GGCTACTTAAACCTGGCGGCGGCGGCCCATGGCC), pEEFP-AAGTGAAGCCCATCTAGTTGGGCTTTGGGTGTCACTCTACTTC (as described above for region 1): pEEFP-IBVNR2 291L for A (pEGFP-IBVNR2 291L) and 293L for A (pEEFP-IBVNR2 293L), in the context of EGF-IBV. Unique primer combinations were 291L for A (forward primer GAAGTAGAGGACCAAAAGCGCCTAGGGCTCCTACCT and reverse primer AAGAGCCACCATCTTTGCTGGTGTTGTCAGC) and 293L for A (forward GAAGTAGAGGACCAAAAGCGCCTAGGGCTCCTACCT and reverse primer TCTAAAGGAGCCCATCGGTGCTGTTGTCAGC). The sequences of all constructs used in this study were confirmed by sequencing and where appropriate expression of the resulting fusion proteins by Western blot (data not shown).

**Live cell imaging**

Vero cells were transfected with 1 μg DNA to 5 μg polyethyleneimine (PEI), DNA and PEI were mixed in a total volume of 200 μl serum-free Dulbecco’s modified Eagles media (DMEM) and incubated at 20 °C for 30 min. Thirty-millimetre cell cultures dishes were seeded with 2 × 105 cells in MEM as per cell culture methods. Transfection mix was added in drop-wise to cells and incubated at 37 °C with 5% CO2 for 24 h. Live cell imaging was performed using a Nikon Eclipse TS100 microscope utilizing the appropriate filter for each tag (e.g., Filter B-2A, excitation 450–490 nm for ECFP/EGFP). Fluorescence and bright-field images were captured using a Nikon Digital Sight DS-L1.

**Confocal microscopy**

Confocal sections of fixed samples were captured on an LSM510 META microscope (Carl Zeiss Ltd., Oberkochen, Germany) equipped with a x 40 and x 63, NA 1.4, oil immersion lens. Pinholes were set to allow optical sections of 1 mm to be acquired. In singly transfected cells, ECFP was excited with the 488 nm argon laser running at 10%, and emission was collected through a BP435-485 emission filter. EGFP was excited with the 488 nm argon laser line running at 2%, and emission was collected through a LP505 filter. DsRed was excited with the helium : neon 543 nm laser line in all cases, and emission was collected through a LP560 filter. Due to excitation of the EGFP molecule by the 458 nm argon laser line, EGFp and ECFP cotransfected samples were linearly unmixed using the META detector. Lambda plots of EGFp and ECFP were generated from singly transfected reference samples excited with the 458 nm argon laser line and collected with the META detector between 460 and 520 nm, showing IBV N protein or derived peptides and nucleolar marker proteins for the cytoplasm, nucleus and nucleolus. In this study, all confocal images showing IBV N protein or derived peptides and nucleolar marker proteins are presented in green and red, respectively (false coloured where appropriate using the Zebrafish IMAGE Browser). Colocalization is shown in yellow.

**Protein alignment and comparative modelling**

AlignX (VectorNTI, version 9.0) was used to align amino acid sequences for generation of a consensus NoRS. The application uses a modified Clustal W algorithm. The sequences of the IBV N protein region 1 was searched against the Protein Databank and the N-terminal RNA-binding domain of SARS-CoV nucleocapsid protein (i45); PDB entry 1ask chain A) identified as a suitable template structure for comparative modelling (with a BLAST E-value of 1 × 10–18 for residues of the protein region 1). The model was based on the coordinates of the NMR, minimized average structure. They share 38% sequence identity over 137 residues. The sequence alignment (Figure 8B) was used to construct a model for IBV N protein region 1 using the comparative protein structure modelling program MODELLER (67).
Acknowledgments

This work was funded by the BBSRC, project grant number BB/B03416 to GB and JAH and studentship BB/B030104/34 to JAH. The confocal microscopy facility in the Astbury Centre for Structural Molecular Biology was funded by the Wellcome Trust and SRIF, and we thank Gareth Howell for his help in using this facility and Jae-Hwan You in our laboratory for the provision of the pEGFP-IBV clone. The authors also thank Mark Harris, Eric Hewitt and John Findlay and members of the coronavirus group for critically reviewing this manuscript and suggesting appropriate experiments.

References

1. Lam YW, Trinkle-Mulcahy L, Lamond AI. The nucleolus. J Cell Sci 2005;118:1335–1337.
2. Pederson T. The plurifunctional nucleolus. Nucleic Acids Res 1998;26:3871–3876.
3. Pederson T, Politz JC. The nucleolus and the four ribonucleoproteins of translation. J Cell Biol 2000;148:1091–1095.
4. Rubbi CP, Milner J. Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. Embo J 2003;22:6068–6077.
5. Andersen JS, Lyon CE, Leung AK, Mann M, Lamond AI. Directed proteomic analysis of the human nucleolus. Curr Biol 2002;12:1–11.
6. Leung AK, Andersen JS, Mann M, Lamond AI. Bioinformatic analysis of the nucleolus. Biochem J 2003;376:553–569.
7. Andersen JS, Lam YW, Leung AK, Ong SE, Lyon CE, Lamond AI, Mann M. Nucleolar proteome dynamics. Nature 2005;433:77–83.
8. Cokol M, Nair R, Rost B. Finding nuclear localization signals. EMBO Rep 2000;1:411–415.
9. Macara IG. Transport into and out of the nucleus. Microbiol Mol Biol Rev 2001;65:570–584.
10. Garcia-Bustos J, Heitman J, Hall MN. Nuclear protein localization. Biochim Biophys Acta 1991;1071:83–101.
11. Ossareh-Nazari B, Gwizdek C, Dargemont C. Protein export from the nucleolus and protein degradation. Oncogene 2005;24:2723–2734.
12. Carmo-Fonseca M, Mendes-Soares L, Campos I. To be or not to be in the nucleus. Traffic 2001;2:684–689.
13. Carson-Fonseca M, Mendes-Soares L, Campos I. To be or not to be in the nucleolus. Nat Cell Biol 2000;2:E107–E112.
14. Song Z, Wu M. Identification of a novel nucleolar localization signal and a degradation signal in Survivin-deltaEx3: a potential link between nucleolus and protein degradation. Oncogene 2005;24:2723–2734.
15. Tsai RYL, McKay RDG. A multistep, GTP-driven mechanism controlling the dynamic cycling of nucleostemin. J Cell Biol 2005;168:179–184.
16. la Cour T, Kiemer L, Molgaard A, Gupta R, Skriver K, Brunak S. Analysis and prediction of leucine-rich nuclear export signals. Protein Eng Des Sel 2004;17:527–536.
17. Croc CF, Garcia-Sastre A, Palese P. An unconventional NLS is critical for the nuclear import of the influenza A virus nucleoprotein and ribonucleoprotein. Traffic 2005;6:205–213.
18. Tsukahara F, Maru Y. Identification of novel nuclear export and nuclear localization-related signals in human heat shock cognate protein 70. J Biol Chem 2004;279:8867–8872.
19. Ryabov EV, Kim SH, Taliansky M. Identification of a nuclear localization signal and nuclear export signal of the umbraviral long-distance RNA movement protein. J Gen Virol 2004;85:1329–1333.
20. Ladd AN, Cooper TA. Multiple domains control the subcellular localization and activity of ETR-3, a regulator of nuclear and cytoplasmic RNA processing events. J Cell Sci 2004;117:3519–3529.
21. Cheng G, Brett ME, He B. Signals that dictate nuclear, nucleolar, and cytoplasmic shuttling of the gamma(1) protein of herpes simplex virus type 1. J Virol 2002;76:9434–9445.
22. Valdez BC, Perlaky J, Henning D, Saio Y, Chan PK, Busch H. Identification of the nuclear and nucleolar localization signals of the protein p120. Interaction with translation protein B23. J Biol Chem 1994;269:23776–23783.
23. Hiscox JA, Wurm T, Wilson L, Cavanagh D, Britton P, Brooks G. The coronavirus infectious bronchitis virus nucleoprotein localizes to the nucleolus. J Virol 2001;75:506–512.
24. Wurm T, Chen H, Britton P, Brooks G, Hiscox JA. Localization to the nucleolus is a common feature of coronavirus nucleoproteins and the protein may disrupt host cell division. J Virol 2001;75:9345–9356.
25. You J-H, Dove BK, Enjuanes L, DeDiego ML, Alvarez E, Howell G, Heinen P, Zambon M, Hiscox JA. Sub-cellular localisation of the severe acute respiratory syndrome coronavirus nucleocapsid protein. J Gen Virol 2005;86:3303–3310.
26. Chen H, Wurm T, Britton P, Brooks G, Hiscox JA. Interaction of the coronavirus nucleoprotein with nuclear antigens and the host cell. J Virol 2002;76:5233–5250.
27. Dove BK, Brooks G, Bicknell KA, Wurm T, Hiscox JA. Cell cycle perturbations induced by infection with the coronavirus infectious bronchitis virus and their effect on virus replication. J Virol 2006;80:4147–4156.
28. Chen CJ, Makino S. Murine coronavirus replication induces cell cycle arrest in G0/G1 phase. J Virol 2004;78:5658–5668.
29. He R, Leeson A, Andonov A, Li Y, Bastien N, Cao J, Osiowy C, Dobbie F, Cutts T, Ballantine M, Li X. Activation of AP-1 signal transduction pathway by SARS coronavirus nucleocapsid protein. Biochem Biophys Res Commun 2003;311:870–876.
30. Leong WF, Tan HC, Ooi EE, Koh DR, Chow VT. Microarray and real-time RT-PCR analyses of differential human gene expression patterns induced by severe acute respiratory syndrome (SARS) coronavirus infection of Vero cells. Microbes Infect 2005;7:248–259.
31. Lai MM, Cavanagh D. The molecular biology of coronaviruses. Adv Virus Res 1997;48:1–100.
32. Weiss SR, Navas-Martin S. Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. Microbiol Mol Biol Rev 2005;69:635–664.
33. Peiris JS, Yuen KY, Osterhaus AD, Stohr K. The severe acute respiratory syndrome. N Engl J Med 2003;349:2431–2441.
34. Cavanagh D. Coronaviruses in poultry and other birds. Avian Pathol 2005;34:439–448.
35. Bergmann CC, Lane TE, Stohlmann SA. Coronavirus infection of the central nervous system: host-virus stand-off. Nat Rev Microbiol 2006;4:121–132.
36. Enjuanes L, Sanchez C, Gebauer F, Mendez A, Dopazo J, Ballesteros ML. Evolution and tropism of transmissible gastroenteritis coronavirus. Adv Exp Med Biol 1993;342:35–42.
37. Chen H, Gill A, Dove BK, Emmett SR, Kemp FC, Ritchie MA, Dee M, Hiscox JA. 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 2005;79:1164–1179.
38. Parker MM, Masters PS. Sequence comparison of the N genes of 5 strains of the coronavirus mouse hepatitis-virus suggests a 3 domain-structure for the nucleocapsid protein. Virology 1996;179:463–468.
39. Calvo E, Escors D, Lopez JA, Gonzalez JM, Alvarez A, Arza E, Enjuanes L. Phosphorylation and subcellular localization of transmissible gastroenteritis virus nucleocapsid protein in infected cells. J Gen Virol 2005;86:2255–2267.

Traffic 2006; 7: 833–848

Coronavirus Nucleocapsid Protein Trafficking
40. Nakai K, Horton P. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem Sci 1999;24:34–35.

41. Pendleton AR, Machamer CE. Infectious bronchitis virus 3a protein localizes to a novel domain of the endoplasmic reticulum. J Virol 2005;79:6142–6151.

42. Lontok E, Corse E, Machamer CE. Intracellular targeting signals contribute to localization of coronavirus spike proteins near the virus assembly site. J Virol 2004;78:5913–5922.

43. Dove BK, You JH, Reed ML, Emmett SR, Brooks G, Hiscox JA. Changes in nucleolar architecture and protein profile during coronavirus infection. Cell Microbiol 2006 Doi: 10.1111.j.1462-5822.2006.00698.x.

44. Catez F, Erard M, Schaerer-Uthurralt N, Kindbeiter K, Madjar JJ, Diaz JJ. Unique motif for nucleolar retention and nuclear export regulated by phosphorylation. Mol Cell Biol 2002;22:1126–1139.

45. Huang Q, Yu L, Petros AM, Gunasekera A, Liu Z, Xu N, Hajduk P, Mack J, Fesik SW, Olejniczak ET. Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. Biochemistry 2004;43:6059–6063.

46. Marra MA, Jones SJ, Astell CR, Holt RA, Brooks-Wilson A, Butterfield YS, Reed et al. SARS-associated coronavirus. Science 2003;300:1399–1404.

47. Johnston A, Cavanagh D, Britton P. Experimental evidence of recombination in coronavirus infectious bronchitis virus. Virology 1999;263:569–580.

48. Storrs K, Shaw K, Evans S, Dalton K, Cavanagh D, Britton P. Leader switching occurs during the rescue of defective RNAs by heterologous strains of the coronavirus infectious bronchitis virus. J Gen Virol 2000;81:791–801.

49. Williams AK, Wang L, Sneed LW, Collier RW. Comparative analyses of the nucleocapsid genes of several strains of infectious bronchitis virus and other viruses. Virus Res 1992;25:213–222.

50. Surjit M, Kumar R, Mishra RN, Reddy MK, Chow VT, Lal SK. The severe acute respiratory syndrome coronavirus nucleocapsid protein is phosphorylated and localizes in the cytoplasm by 14-3-3-mediated translocation. J Virol 2005;79:11476–11486.

51. Li FQ, Xiao H, Tam JP, Liu DX. Sumoylation of the nucleocapsid protein of severe acute respiratory syndrome coronavirus. FEBS Lett 2005;579:2387–2396.

52. Rowland RR, Schneider P, Fang Y, Wootten S, Yoo D, Benfield DA. Peptide domains involved in the localization of the porcine reproductive and respiratory syndrome virus nucleocapsid protein to the nucleolus. Virology 2003;316:135–145.

53. Tijms MA, van der Meer Y, Snijder EJ. Nuclear localization of non-structural protein 1 and nucleocapsid protein of equine arteritis virus. J Gen Virol 2002;83:795–800.

54. Hiscox JA. Brief review: the nucleolus – a gateway to viral infection? Arch Virol 2002;147:1077–1089.

55. Hiscox JA. The interaction of animal cytoplasmic RNA viruses with the nucleus to facilitate replication. Virus Res 2002;86:13–22.

56. Gustin KE. Inhibition of nucleo-cytoplasmic trafficking by RNA viruses: targeting the nucleolar pore complex. Virus Res 2003;95:35–44.

57. Weidman MK, Sharma R, Raychaudhuri S, Kundu P, Tsai W, Dasgupta A. The interaction of cytoplasmic RNA viruses with the nucleus. Virus Res 2003;95:75–85.