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.

ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

VIROLOGY

Virology 316 (2003) 135–145

www.elsevier.com/locate/yviro

Peptide domains involved in the localization of the porcine reproductive and respiratory syndrome virus nucleocapsid protein to the nucleolus

Raymond R.R. Rowland,[a,*] Paula Schneider,[a] Ying Fang,[b] Sarah Wootton,[c] Dongwan Yoo,[c] and David A. Benfield[b]

[a] *Department of Diagnostic Medicine and Pathobiology, 1800 Denison Avenue, Kansas State University, Manhattan, KS 66506, USA*
[b] *Department of Veterinary Science, South Dakota State University, Brookings, SD 57007, USA*
[c] *Department of Pathobiology, Ontario Veterinary College, University of Guelph, Guelph, Ontario, Canada N1G 2W1*

Received 10 June 2002; returned to author for revision 30 May 2003; accepted 4 June 2003

Abstract

The nucleocapsid (N) protein of porcine reproductive and respiratory syndrome virus (PRRSV) is the principal component of the viral nucleocapsid and localizes to the nucleolus. Peptide sequence analysis of the N protein of several North American isolates identified two potential nuclear localization signal (NLS) sequences located at amino acids 10–13 and 41–42, which were labeled NLS-1 and NLS-2, respectively. Peptides containing NLS-1 or NLS-2 were sufficient to accumulate enhanced green fluorescent protein (EGFP) in the nucleus. The inactivation of NLS-1 by site-directed mutagenesis or the deletion of the first 14 amino acids did not affect N protein localization to the nucleolus. The substitution of key lysine residues with uncharged amino acids in NLS-2 blocked nuclear/nucleolar localization. Site-directed mutagenesis within NLS-2 identified the sequence, KKNKK, as forming the core localization domain within NLS-2. Using an in vitro pull-down assay, the N protein was able to bind importin-α, importin-β nuclear transport proteins. The localization pattern of N-EGFP fusion peptides represented by a series of deletions from the C- and N-terminal ends of the N protein identified a region covering amino acids 41–72, which contained a nucleolar localization signal (NoLS) sequence. The 41–72 N peptide when fused to EGFP mimicked the nucleolar–cytoplasmic distribution of native N. These results identify a single NLS involved in the transport of N from the cytoplasm and into nucleus. An additional peptide sequence, overlapping NLS-2, is involved in the further targeting of N to the nucleolus.

Keywords: Porcine reproductive and respiratory syndrome virus (PRRSV); Nucleocapsid (N) protein; Nuclear localization signal (NLS); Nucleolar localization signal (NoLS)

Introduction

Porcine reproductive and respiratory syndrome (PRRS) is caused by an enveloped, positive-stranded RNA virus, PRRSV, belonging to the family *Arteriviridae* (Benfield et al., 1992; Plagemann, 1996; Snijder and Meulenberg, 1998). PRRSV isolates are further divided into North American and European genotypes, represented by VR-2332 and the Lelystad virus, respectively (Nelsen et al., 1999). Other members of the arterivirus group include lactate dehydrogenase-elevating virus (LDV) of mice, equine arteritis virus (EAV), and simian hemorrhagic fever virus (SHFV). The arteriviruses, toroviruses, and coronaviruses are members of a single order, *Nidovirales* (Cavanagh, 1997). The arteriviruses structurally resemble togaviruses but similar to coronaviruses replicate via a nested 3′-coterminal set of subgenomic mRNAs, which possess a common leader and a poly-A tail (Lai, 1990; Snijder and Meulenberg, 1998).

The 15-kDa nonglycosylated N protein does not possess a signal peptide and is translated and maintained as mature 123 and 128 amino acid proteins, for VR-2332 and Lelystad viruses, respectively (Mardassi et al., 1996). Even though the North American and European PRRSV N proteins are

* Corresponding author. Fax: +1-785-532-4481.
E-mail address: browland@vet.ksu.edu (R.R.R. Rowland).

0042-6822/$ – see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S0042-6822(03)00482-3

only about 63% identical at the amino acid level, they share a common antigenic region between amino acids 52 and 69 (Rodriguez et al., 1997; Meulenberg et al., 1998; Wootton et al., 1998). During virus replication, the N protein localizes to the cytoplasm and nucleolus of cultured porcine macrophages and MARC-145 cells infected with either North American or European PRRSV genotypes (Rowland et al., 1999). An identical nucleolar–cytoplasmic distribution pattern is observed for the N protein of another arterivirus, EAV (Tijms et al., 2002), and for the N proteins of group I, II, and III coronaviruses (Hiscox et al., 2001; Wurm et al., 2001). The arterivirus and coronavirus N proteins, when expressed alone or fused to the red-shifted enhanced green fluorescent protein (EGFP), localize to the nucleolus, demonstrating that translocation across the nuclear pore complex (NPC) and accumulation in the nucleolus are independent of other viral proteins (Rowland et al., 1999; Hiscox et al., 2001; Wurm et al., 2001; Tijms et al., 2002).

Within a population of infected cells, the intracellular distribution of arterivirus and coronavirus N proteins is highly variable. For example, a small percentage of coronavirus-infected cells possess N in the nucleolus (Hiscox et al., 2001). In a similar study using PRRSV, we observed that the percentage of infected cells containing N in the nucleolus was dependent on when observations were made. In addition, cells were identified that contained N in the nucleoplasmic region surrounding the nucleoli. These observations suggest that the localization of N to the nucleolus is a complex and highly regulated process.

The translocation of a protein from the cytoplasm, across the NPC, and into the nucleus is initiated by the binding of a nuclear localization signal (NLS) on the cargo protein to the importin-α shuttle protein. The cargo/importin-α complex binds to a second shuttle protein, importin-β, which is then targeted to the NPC. In the presence of additional accessory proteins, including Ran-GDP, and ATP as a source of energy, the cargo/importin complex is transported across the NPC and released into the nucleoplasm (reviewed in Gorlich and Kutay, 1999). Classical NLS sequences incorporate regions enriched in basic amino acids and generally conform to one of three types (Nakai and Kanehisa, 1992; Hicks and Raikhel, 1995). The “pat4” NLS consists of a continuous stretch of four basic amino acids (lysine or arginine) or three basic amino acids associated with histidine or proline. The “pat7” NLS starts with a proline and is followed within three residues by a segment containing three basic residues out of four. The third type of NLS, known as a “bipartite” motif, consists of two basic amino acids, a 10-amino-acid spacer, and a 5-amino-acid segment containing at least three basic residues. There are exceptions to these established rules. Proteins that lack an NLS can localize to the nucleus as a result of cotransport with other nuclear proteins. Proteins which possess an NLS may remain cytoplasmic, especially if the NLS is not exposed on the protein’s surface (Roberts et al., 1987). And finally, small proteins less than 50–70 kDa can passively diffuse through the nuclear pore complex (Paine, 1975; Gorlich and Kutay, 1999). The addition of an NLS to a small protein that passively diffuses through the NPC confers to that protein all of the properties of NLS-dependent transport (Jans and Hubner, 1996).

Unlike the nucleus, the nucleolus is not a membrane-bound organelle. Therefore, the targeting of a protein to the nucleolus occurs through the diffusion of the protein through the nucleoplasm and accumulation in the nucleolus. Localization to the nucleolus is mediated by a nucleolar targeting sequence (NoTS), which functions by forming an interaction with a nucleolar component, such as a nucleolar protein, rRNA, or rDNA, or small nucleolar RNAs. By definition, the NoTS by itself cannot translocate a protein across the NPC, but once in the nucleoplasm is capable of targeting the protein to the nucleolus (Lixin et al., 2001). Another route to the nucleolus is the direct translocation of a protein from the cytoplasm to the nucleolus by cotransport with a nucleolar shuttle protein, such as nucleolin, fribrillarin, and B-23, which normally shuttle between the cytoplasm and the nucleolus. The purpose of this study was to gain insight into the mechanism of PRRSV N protein localization to the nucleolus by characterizing the individual oligopeptides domains, which participated in the translocation of the N protein across the NPC and localization to the nucleolus.

Results

Import of the N protein into the nucleus is NLS dependent

Previous work by us described the accumulation of the PRRSV N protein in MARC-145 cells during infection and after expression of recombinant N protein tagged with EGFP (Rowland et al., 1999). The purpose of this series of experiments was to determine if the translocation of N protein was dependent on a specific NLS domain. Peptide sequence analysis of the N protein of the North American PRRSV isolate SDSU-23983 using the PSORT computer program identified two NLS motifs, a “pat-4” NLS, 10-KRKK, and a “pat-7” NLS, 41-PGKKNKK (Fig. 1). Based on a hydrophilicity plot of the peptide sequence, both NLSs were predicted to be located in regions of the N protein that are highly hydrophilic and accessible for interaction with shuttle proteins. For convenience, the putative NLSs were designated NLS-1 and NLS-2, respectively. Peptide sequences of the N protein from 70 North American isolates deposited in GenBank showed that both NLSs are highly conserved among North American PRRSV isolates. The principal difference between isolates was the conserved substitution of lysines with arginines (data not shown). The European Lelystad virus N protein also contains NLS-1 and NLS-2 domains, which are located at residues 9 and 47, respectively.

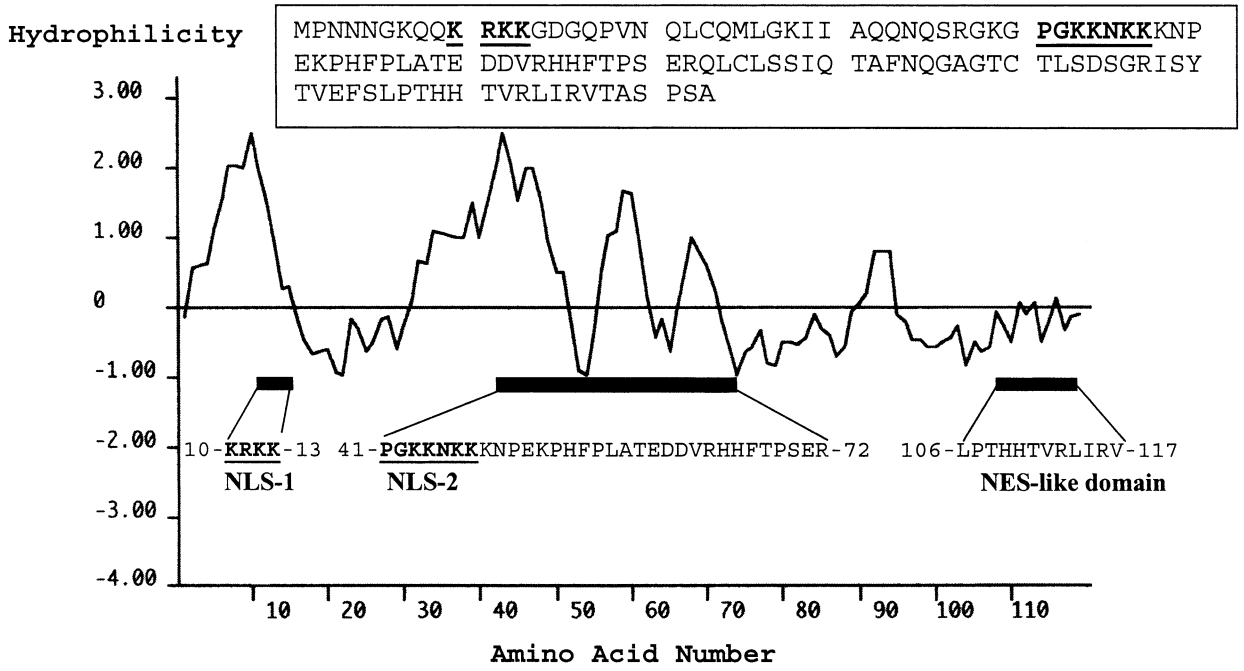

Fig. 1. Hydrophilicity plot of PRRSV N protein. The analysis of hydrophilicity was performed according to Hopp and Woods (1981). The inset contains the complete amino acid sequence of the N protein from SDSU-23983, the PRRSV isolate used in this study. Domains that formed NLS motifs were identified using the PSORT computer program (Nakai and Kanehisa, 1992) and labeled NLS-1 and NLS-2. A hydrophobic region, between amino acids 106 and 117, contained a potential nuclear export signal (NES) sequence.

The intracellular properties of the N protein were studied in MARC-145 cells transfected with N protein gene fragments cloned into a pEGFP-N1 expression vector. The nomenclature "N1" describes the insertion of the fragment upstream of the EGFP and within the context of the first open reading frame. MARC-145 cells, a monkey kidney cell line, were chosen because they support PRRSV replication to relatively high levels (Kim et al., 1993). As shown in Fig. 2A and B, the 25-kDa EGFP localized to the nucleoplasm and cytoplasm, which is consistent with passive diffusion. EGFP was excluded from the nucleolus and some cytoplasmic compartments. The fusion of N protein to EGFP altered the distribution of EGFP, resulting in the accumulation of fluorescent label in the nucleolus and decreased accumulation in the nucleoplasm surrounding the nucleoli (Fig. 2C and D). In previous work, colocalization of anti-N protein and nucleolar-specific antibodies were used to identify the nucleolus as the site of N protein accumulation (Rowland et al., 1999). Nucleoli, when viewed under phase-contrast microscopy, are apparent as highly refractive bodies within the nucleus.

To determine if NLS-1 and NLS-2 could function as legitimate localization signal sequences, N peptides 1–14 and 39–48 were tagged with EGFP and expressed in MARC-145 cells. The protein products of NP(1–14)-EGFP and NP(39–48)-EGFP accumulated in the nucleus and nucleolus, with a small amount of fluorescence remaining in the cytoplasm (see Fig. 2E and F). These data indicated that NLS-1 and NLS-2 could function as NLS domains sufficient to localize EGFP to the nucleus, but provided no information on which NLS participated in the translocation of the native N protein to the nucleus. Therefore, the localization properties of NP-EGFP were studied after inactivation of each NLS. Site-directed mutagenesis was used to replace key lysine and arginine residues, the essential residues involved NLS function, with uncharged alanines or glycines. The mutant protein, NP(1–123,G10G11)-EGFP, which lacked an intact NLS-1, localized to the nucleolus (see Fig. 2I). In a second experiment, NLS-1 was completely removed by deleting the first 14 amino acids from the N-terminus. The protein product, NP(14–123)-EGFP, also localized to the nucleolus, indicating that NLS-1 was not involved in nuclear translocation.

Inactivation of NLS-2 was accomplished by replacing lysine residues at positions 43 and 44 with glycines. Removal of either lysine reduced the accumulation of NP-EGFP in the nucleolus (data not shown), whereas NP(1–123,G43G44)-EGFP, which lacked both lysines, was retained in the cytoplasm and excluded from the nucleus/nucleolus, as shown in Fig. 2H. A phase-contrast image of the same cell (Fig. 2G) confirmed that nucleoli were present. These results suggested that lysines-43 and -44 formed part of the functional NLS peptide domain, which was required for the entry of N protein into the nucleus. Another possibility was that nuclear localization activity was lost in the mutant protein as a result of refolding caused by the lysine substitutions. We took advantage of a panel of six anti-N monoclonal antibodies (mAbs) to probe the con-

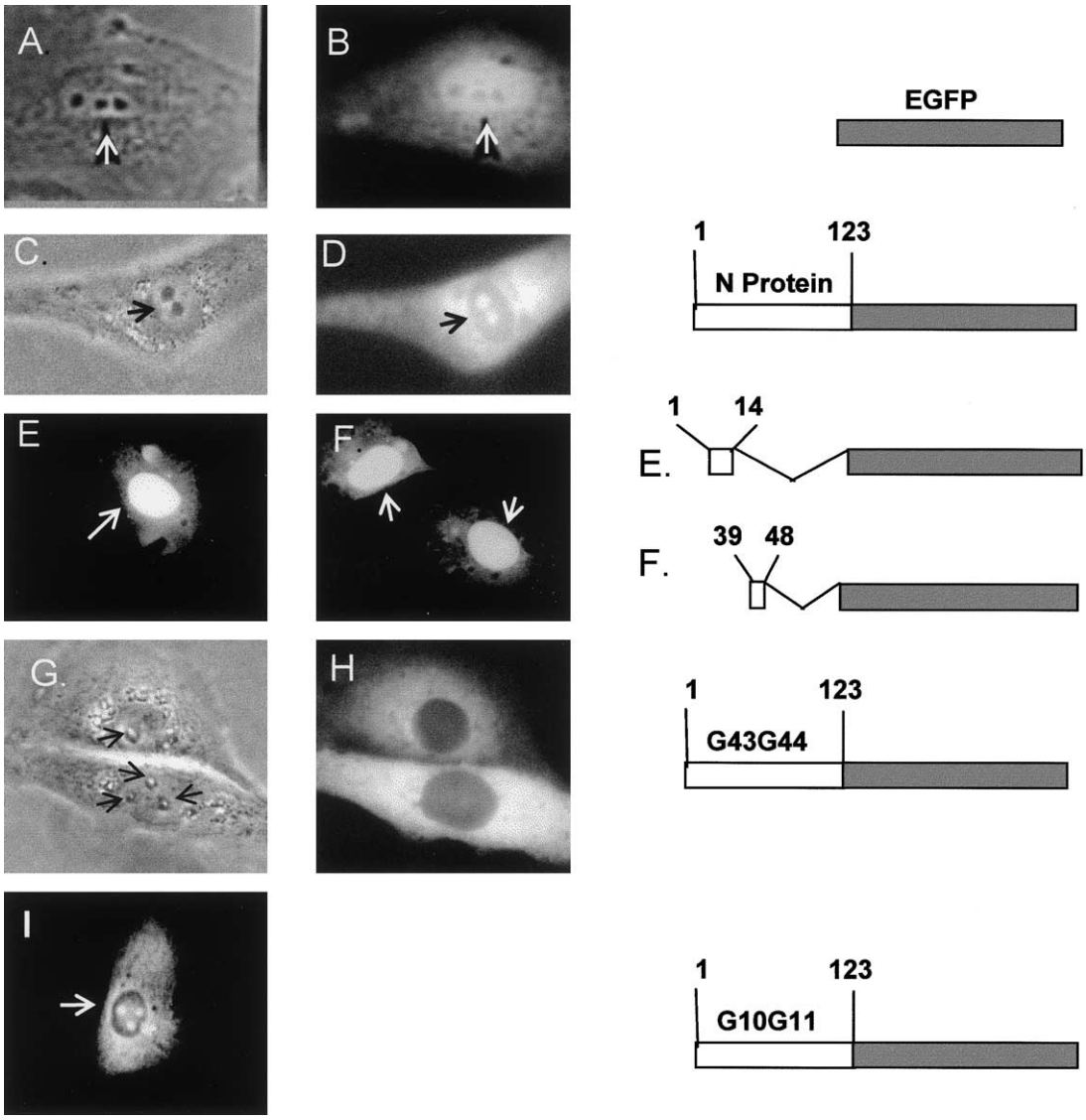

Fig. 2. Localization of EGFP-tagged wild-type and mutant N proteins. (A and B) Phase-contrast (left) and fluorescence (right) photomicrographs of the same cell at 15 h after transfection with the pEGFP vector. Arrows locate the nucleus. The nucleoli in both phase-contrast and fluorescence photomicrographs are identified as the dark spots in the nuclei. (C and D) Localization of wild-type N gene (construct NP (1–123)-EGFP) showing the accumulation of NP-EGFP in cytoplasmic and nucleolar compartments. (E and F) Localization of N protein fragments containing the NLS-1 sequence, 10-KRKK, and the NLS-2 sequence, 41-PGKKNKK. Arrows point to the nuclei. (G and H) Phase-contrast (left) and fluorescence (right) photomicrographs of the same cells after transfection with a gene construct that expressed the NLS-2-minus N protein gene (nucleoli are identified with arrows). (I) Intracellular localization of the NLS-1-minus mutant NP-EGFP protein. The arrow points to the nucleus of a cell containing three fluorescent nucleoli.

formation of wild-type and mutant N proteins. All mAbs are known to interact with conformation-dependent epitopes distributed throughout the last 90 amino acids of the PRRSV N protein (see Table 1; Nelson et al., 1993; Wootton et al., 1998, 2001). The results of the antibody binding studies are summarized in Table 1. All antibodies reacted with the full-length NP-EGFP fusion protein, indicating that EGFP did not affect mAb recognition and presumably did not alter N protein conformation (Table 1). There was no immunoreactivity with EGFP alone or EGFP fused to a small N peptide, such as NP(43–72), which contained a small linear segment of the conformational epitope recognized by mAb. All mAbs reacted with the NLS-1 and NLS-2 mutant proteins, NP(14–123)-EGFP and NP(1–123,G43G44)-EGFP, respectively. Even though four of the six mAbs (MR-40, EP-147, JP-24, VO-17) recognize the same N proteins regions, amino acids 52 to 69 and 112 to 123, these mAb are distinguished from each other based on differences in reactivity with field isolates and mutant recombinant N proteins (personal communication, Eric Nelson, South Dakota State University; Wootton et al., 2001). Together these data identify NLS-2 as the principal localization signal sequence involved in the translocation of the native N protein into the nucleus.

Additional site-directed mutagenesis studies were performed within NLS-2 to further delineate key residues in-

Table 1
Reactivity of N protein-EGFP constructs with anti-N monoclonal antibodies

mAb	Epitope[b]	Reactivity with EGFP constructs[a]						
		No peptide	43–72	1–123	1–123 G43,G44	1–123 G46,G47	14–123	1–115
SDOW-17	30–52, 112–123	−	−	+	+	+	+	+
SR-30	69–123	−	−	+	+	+	+	−
MR-40	52–69, 112–123	−	−	+	+	+	+	+
EP-147	52–69, 112–123	−	−	+	+	+	±	ND[c]
JP-24	52–69, 112–123	−	−	+	+	+	+	+
VO-17	52–69, 112–123	−	−	+	+	+	+	−

[a] Peptide region of N protein fused to EGFP.
[b] The epitope regions recognized by each monoclonal antibody (mAb) are described in Wootton et al. (1998) and are identified by beginning and ending amino acid.
[c] ND, not determined.

volved in nuclear translocation. These results are summarized in Table 2. NP(1–123,G46G47), which lacked lysines 46 and 47, remained in the cytoplasm and retained reactivity with the panel of mAbs. The removal of other amino acid pairs, such as asparagine-45/lysine-46, lysine-47/lysine-48, and lysine-48/asparagine-49, resulted in noticeable reduction in nucleolar localization. The substitution of proline-41 and asparagine-42 with alanines did not affect the amounts of nucleolar fluorescence, but did produce a noticeable increase in the accumulation of NP-EGFP in the nucleoplasm surrounding the nucleoli with what appeared to be a corresponding decrease of fluorescence in the cytoplasm.

Residues located outside of the predicted NLS also affected localization. For example, the substitution of lysine-48 and asparagine-49 with alanines produced a noticeable reduction in nucleolar fluorescence. Together these data identify 43-KKNKK as forming the core localization domain within the 41-PGKKNKK NLS-2 sequence. Furthermore, residues outside of this core sequence can affect N protein localization.

Recovery of NLS-1 activity in NLS-2-minus mutant N proteins

The 1–14 peptide, which contained the NLS-1 sequence, KKRK, was able to function as an NLS sufficient to accumulate the EGFP tag in the nucleus, but was inactive in the context of the native protein. The lack of localization activity from NLS-1 could reflect the inhibitory activity of adjacent residues, which could block the interaction between the NLS and the receptor domain on the shuttle protein. Another possibility was that the NLS was blocked as a result of the tertiary or quaternary conformation of the N protein. Since mutations in the C-terminal 11 amino acids can significantly alter the conformation of the N protein (Wootton et al., 1998), we investigated the possibility that mutations in the C-terminus, sufficient to affect recognition by mAb, would reveal nuclear localization activity from NLS-1. The experimental approach was to follow the intracellular localization of the NLS-2 mutant, NP(1–115,G43G44), which contained an eight-amino-acid dele-

Table 2
Mutational analysis of the NLS-2 region

Construct[a]	Sequence[b]	Localization[c]		
		Cy	Nu	No
NP (1–123)	41-P G K K N K K K N P-50	+++	−	+++
NP (1–123, A41A42)	A A	++	+	+++
NP (1–123, G43)	. . G	+++	−	++
NP (1–123, G44)	. . . G	+++	−	++
NP (1–123, G43G44)	. . G G	+++	−	−
NP (1–123, A45A46)	. . . . A A	+++	−	+[d]
NP (1–123, A46A47)	. . . . . G G . . .	+++	−	−
NP (1–123, A47A48)	. . . . . G G . .	+++	−	+[d]
NP (1–123, A48A49)	. . . . . A A .	+++	−	+[d]

[a] Constructs were expressed as full-length N proteins fused to EGFP. The indicated amino acids were substituted with glycines or alanines.
[b] Underlined amino acids show the composition of the pat7 NLS motif.
[c] The relative amount of fluorescence in each compartment of a representative cell. Cytoplasm (Cy), nucleoplasm (Nu), and nucleolus (No). +++, Strong fluorescence signal; ++, medium fluorescence; +, weak fluorescence; −, no visible fluorescence in any cell.
[d] Approximately 10% of fluorescent cells showed weak nucleolar fluorescence.

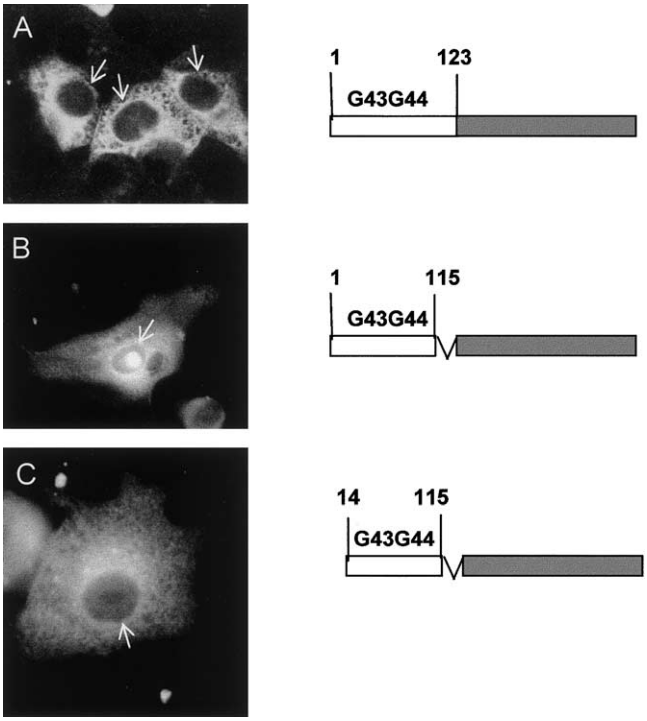

Fig. 3. Recovery of NLS-1 activity in N proteins that contain a truncation of eight amino acids from the C-terminus. Mutant N proteins were expressed as EGFP fusion proteins. The composition of each construct is illustrated to the right of the photomicrographs. Arrows point to the nucleus in each cell.

tion from the C-terminus. The result, presented in Fig. 3B, showed that NP(1–115,G43G44)-EGFP localized to the nucleolus, whereas the parent mutant protein NP(1–123,G43,G44)-EGFP remained cytoplasmic. To confirm that localization of NP(1–115,G43G44) was NLS-1 dependent, we followed the localization of a NLS double mutant, NP(14–115,G43G44)-EGFP. As shown in Fig. 3C, the protein product was retained in the cytoplasm, demonstrating the loss of NLS activity after removal of NLS-1. Removal of eight amino acids from the N protein resulted in the loss of immunoreactivity with SR-30 and VO-17 mAbs (Table 1), but did not alter recognition by SDOW-17, MR-30, or JP-24. The ability of a mutation in the C-terminus of the N protein to recover activity from NLS-1, which is located at the opposite end of the protein, demonstrates that the overall conformation of the native N protein is responsible for blocking NLS-1 activity. Furthermore, since the deletion of amino acids did not affect binding by three of the six mAbs, only a small conformational change is necessary to expose NLS-1.

The N protein binds importin shuttle proteins

To further support the notion that N protein nuclear transport is NLS-mediated, we investigated the ability of N to interact with importin-α and -β proteins in a GST-bead binding assay. Immunoprecipitation of radiolabled proteins from infected cell lysates or in vitro translated products using N-specific antibodies yielded a protein, which migrated at approximately 15 kDa (Fig. 4A, lane 1, Fig. 4B, lane 3). Immunoprecipitation of in vitro translated N protein, derived from the pCITE vector, yielded a slightly slower migrating product (Fig. 4B, lane 3). The slower migration was consistent with the addition of amino acid residues contributed by the pCITE vector. Since the PRRSV N protein typically forms noncovalent homodimers (Wootton and Yoo, 2003), GST-N was able to precipitate radiolabeled N from cell lysates from infected but not mock-infected cells (Fig, 4B, lanes 4 and 5). As a negative control, Sepharose GST alone failed to pull down proteins corresponding to the size of native (Fig. 4B lanes 1 and 2) or in vitro translated N proteins (data not shown). N proteins from PRRSV-infected cells or from in vitro translation were

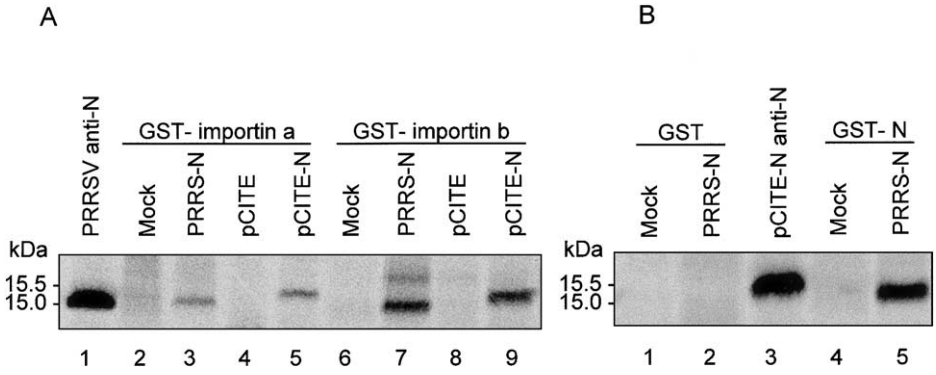

Fig. 4. Interaction of recombinant and virus-produced N proteins with importin-α, importin-β nuclear transport proteins. N proteins from PRRSV-infected cells or from in vitro translated products (pCITE) were incubated with Sepharose-GST immobilized murine importin-α or murine importin-β (A) or human CRM1 (B). Negative controls included incubation of GST-Sepharose beads with radiolabeled proteins from mock-infected and infected cells (B, lanes 1 and 2) or incubation of GST-fusion proteins with products from mock-infected and mock-translated cells (A, lanes 2, 4, 6, 8, for importin-α and -β, and B, lanes 6 and 8 for CRM1). Positive controls included PRRSV immunoprecipitated with anti-N rabbit serum (A, lane 1 and B, lane 3) and N coprecipitated with GST-N (B, lanes 4 and 5). N protein products derived from in vitro translation were slightly larger due to the incorporation of a His-Tag and additional amino acids incorporated from multiple cloning site of the pCITE vector.

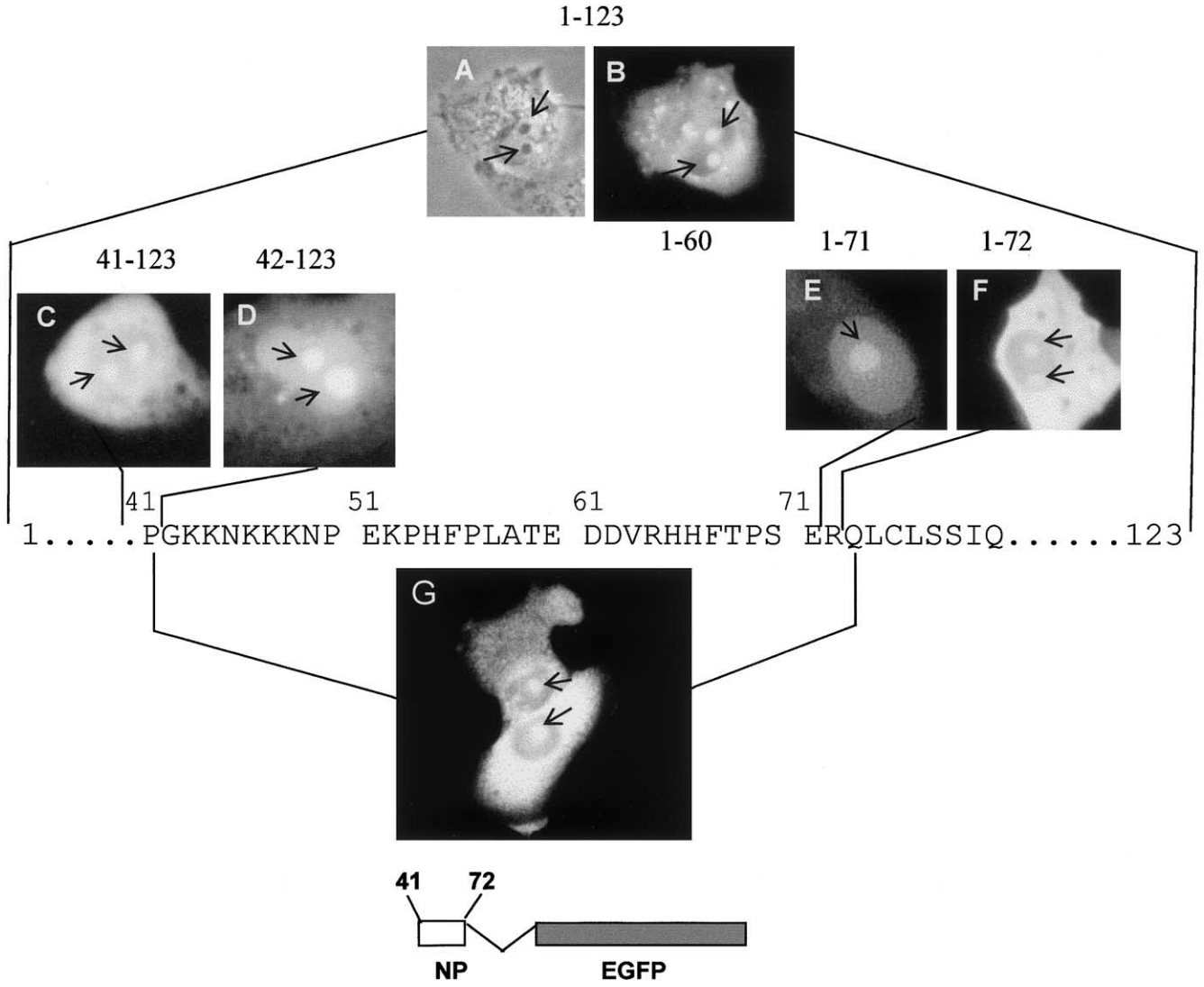

Fig. 5. Identification of a nucleolar localization signal (NoLS) sequence peptide. N protein gene fragments were truncated from the N-terminal end (C and D) or the C-terminal end (E-F) and expressed in MARC-145 cells as peptide-EGFP fusion proteins. (G) shows the distribution of the 41–72 peptide fragment fused to EGFP. Arrows in each photomicrograph point to the locations of the individual nucleoli.

able to form an interaction with GST-importin-α and GST-importin-β (see Fig. 4A). No bands of the correct molecular size were observed when radiolabeled proteins from mock-infected cells or translation products from pCITE vector RNA were incubated with GST-importin proteins. These data show that nuclear shuttle proteins have the capacity to form specific interactions with the PRRSV N protein. Furthermore, the capacity of importin-α and importin-β to bind in vitro translated N protein indicates that N-importin interactions are independent of other viral proteins.

The 41 and 72 peptide contains NoLS activity

Proteins that localize to the nucleolus typically possess a nucleolar localization signal (NoLS) motif. The NoLS associated with viral proteins is usually no longer than 20 amino acid residues and possesses at least nine basic residues overall, including one continuous stretch of four basic amino acids, or two stretches of three basic amino acids. Furthermore, an NLS is usually embedded within the NoLS peptide sequence (Kubota et al., 1999). Peptide sequence analysis of the regions surrounding NLS-1 and NLS-2 showed no identity with this type of NoLS motif (see Fig. 1). Therefore, to locate the smallest peptide sufficient to function as an NoLS, amino acids we followed the localization of NP-EGFP fusion proteins after sequentially removal of amino acid residues from the N- and C-terminal ends. A summary of the results, presented in Fig. 5, showed that the N protein construct lacking the first 40 amino acids (NP(41–123)) or last 51 amino acids (NP(1–72)) produced a nucleolar–cytoplasmic distribution pattern similar to the full-length N protein, with a small amount of fluorescence in the nucleoplasmic region surrounding the nucleoli. Removal of a single amino acid from either construct, NP(42–

Table 3
Comparison of possible PRRSV NoLS sequences with NoLS sequences from other viral proteins

Viral protein	Sequence[a]	Reference
HIV-1 Rev	RQARRN**RRRR**WRERQR	Kubota et al. (1989)
HIV-1 Tat	GR**KKRR**QRRRAHQ	Dang and Lee (1989)
HTLV-1 Rex	**PKTRRRPRR**SQ**RKR**PPTP	Siomi et al. (1988)
Semliki NoLS-1	KP**KKKK**TTKPKPKTQP	Favre et al. (1994)
Semliki NoLS-2	**RRRKR**NRDAA**RRRRK**Q	Favre et al. (1994)
MDV MEQ	**RRRKR**NRDAA**RRRRK**	Liu et al. (1997)
PRRSV NoLS (37)	RGKG**PGKKNKK** NPEK	Hiscox et al. (2001)
PRRSV NoLS (41)	**PGKKNKK**KNPEKPHFPLATEDDVRHHFTPSER	This study

[a] Bold amino acid residues identify NLS motifs. Underlined residues are basic amino acids.

123) or NP(1–71), produced a change in the distribution of EGFP fluorescence, which was most noticeable as an increase in the accumulation of the fusion protein in the nucleoplasm surrounding the nucleoli combined with a corresponding decrease of fluorescent protein in the cytoplasm (Fig. 5D and E). The 41–72 peptide, when fused to EGFP, localized to the cytoplasm and the nucleolus (Fig. 5G) demonstrating that the 41–72 oligopeptide alone can produce a localization pattern similar to the full-length wild-type N protein, supporting the possibility that this domain contains the NoLS sequence.

Discussion

In previous work, we characterized the nucleolar localization of the PRRSV N protein during infection of MARC-145 cells with North American and European PRRSV isolates and following transfection of cells with an N protein-EGFP fusion gene (Rowland et al., 1999). The results of this current study extend these observations by identifying peptide domains, which participate in N protein translocation across the NPC and accumulation in the nucleolus.

Peptide sequence analysis of the PRRSV N protein revealed two putative NLSs, which were predicted to be located in regions with high accessibility (Fig. 1). However, only one NLS, NLS-2, was found to be involved in nuclear transport of the native protein. Further analysis of NLS-2 revealed a sequence, 43-KKNKK, which formed the core of the NLS (Table 2). The removal of lysine residue pairs 43-KK or 46-KK abolished nuclear transport, but did not affect recognition by a panel of conformation-dependent anti-N mAbs. Therefore, in the context of the native N protein, NLS-2 is necessary and sufficient to transport the N protein into the nucleus.

The inactivation or complete removal of NLS-1 had no effect on N protein nuclear translocation, since mAbs prepared against the whole N protein molecule generally fail to recognize the region covered by the first 30 amino acids, even though this region possesses several charged residues (Wootton et al., 1998). Therefore, the absence of immunoreactivity suggests that the N-terminal region, including NLS-1, is normally hidden from the surface of the protein. NLS-1 nuclear localization activity was recovered following the removal of eight amino acids from the C-terminus. The loss of SR-30 and VO-17 reactivity indicated a change in the conformation of the N protein, sufficient to expose NLS-1. Another possibility is that NLS-1 is normally blocked as a result of oligomerization with other N protein molecules. Oligomerization with other N protein molecules may block NLS-1 without affecting the accessibility of NLS-2. In the context of the native protein, NLS-1 can be considered a cryptic NLS.

NLS-dependent nuclear transport is initiated by the interaction between the NLS on the cargo protein with importin-α (Gorlich and Mattaj, 1996). A separate importin-β-binding (IBB) domain on the N-terminus of the importin-α is responsible for forming the interaction between the importin-α/cargo complex and importin-β (Moroianu et al., 1996). Importin-β does not contain an NLS binding site; however, the NLS on viral proteins, such as HIV Rev, can bind directly to importin-β (Henderson and Percipalle, 1997). Using a pull-down assay, we demonstrated that the N protein of PRRSV is able to form interactions with murine GST-importin-α and -β proteins (Fig. 4). Furthermore, the formation of an N protein/importin complex using in vitro translated N demonstrates that N protein binding is independent of other viral proteins. Even though these results demonstrate the capacity of N to form stable interactions with importin, these data do not exclude the possibility for the participation of other shuttle proteins in the translocation of N into the nucleus/nucleolus. For example, major nucleolar proteins, such as nucleolin and B-23, can function in the transport of viral proteins from the cytoplasm to the nucleolus through the interaction between long stretches of acidic residues in B-23 and nucleolin with highly basic NoLS domains. B-23 can form stable interactions with Tat, Rev, Rex, and adenovirus V proteins (Frankhauser et al., 1991; Stauber and Pavlakis, 1998; Adachi et al., 1993; Matthews, 2001). Hiscox et al. (2001), after comparing the peptide sequences of infectious bronchitis virus and PRRSV N proteins, concluded that the PRRSV NoLS was located between amino acids 37 and 52 (Table 3). However, our results present a modified picture in which

an overlapping peptide covering amino acids 41–72 was capable of reproducing the nucleolar–cytoplasmic localization pattern of the wild-type N protein (Fig. 5). Even though the 41–72 domain contains an NLS, which is consistent with other viral NoLS sequences, there are some notable differences. The overall length is longer than the 20 amino acids for the typical viral NoLS sequences (Table 3) (Kubota et al., 1999). The second difference is a paucity of basic residues (eight lysines/arginines) and the incorporation of five negatively charged amino residues, making the 41–72 domain relatively more acidic when compared with other viral NoLS domains. At this time we propose that the 41–72 peptide contains an NLS, represented by the core peptide KKNKK, and a separate nucleolar localization domain, which targets N to the nucleolus.

The EAV N protein contains a potential nuclear export signal (NES) sequence located at residue 54, which may participate in the translocation of N from the nucleus and into the cytoplasm (Tijms et al., 2002). The classical NES motif, found in HIV Rev and cellular proteins, is X-$R_{(2-4)}$-X-R_2-X-R-X, where X represents a hydrophobic amino acid, such as leucine, isoleucine, or valine, and the letter R represents any amino acid (Fischer et al., 1995; Hope, 1999; Henderson and Eleftheriou, 2000). Almost all classical NES sequences contain the X-R-X tripeptide core sequence. The HIV Rev protein NES is LGLPPLERLTL (the hydrophobic residues are underlined). The PRRSV N protein contains two regions, sufficiently enriched in hydrophobic amino acids, to function as potential NES sequences. The first, located between amino acids 19 and 30, does not contain an X-R-X sequence. The second peptide, 106-LPTHHTVR-LIRV, contains at least one core tripeptide sequence (in bold and see Fig. 1). A nearly identical peptide sequence is found in the N protein of the Lelystad virus. It remains to be determined if the PRRSV protein is exported from the nucleus by way of a classical NES.

The function of N in the nucleolus is not known. Chen et al. (2002) showed that a coronavirus N protein was able to bind nucleolin, which was associated with the inhibition of cytokinesis, a possible viral strategy to divert biosynthetic resources from the dividing nucleus and into the cytoplasm. Nucleolar N could also have a neutral effect on cell function, i.e., a simple case of molecular mimicry. For instance, peptide regions enriched in basic amino acids, required to bind viral RNA, can mimic NLS and/or NoLS sequences (LaCasse and Lefebvre, 1995). One intriguing possibility is that localization of N protein to the nucleus/nucleolus represents a novel host cell protection strategy. The sequestration of large quantities of nucleocapsid precursor in the nucleolus would deprive a cytoplasmic virus of an essential structural protein. The incorporation of an export signal in the N protein would represent a counterstrategy, which would ensure that a sufficient quantity of N protein was returned to the cytoplasm where it could participate in the assembly of virions.

Materials and methods

Plasmid construction and mutagenesis

Restriction sites were added to the 5′ end of the respective sense and antisense PCR primers used to amplify the desired regions of the N gene. PCR products were double-digested with the appropriate restriction enzymes and ligated into pEGFP-N1 (Clontech Inc.). Site-directed mutagenesis was performed by overlapping extension PCR (Ho et al., 1989; Jespersen et al., 1997). Briefly, two fragments of N gene were amplified separately using overlapping primers that contained the desired mutation. Equal amounts of each PCR product were added to the reaction and the overlap extension was carried out for five cycles at 95°C for 30 s, 58°C for 30 s, and 72°C for 1 min.

Transfection of DNA into MARC-145 cells, a PRRSV-permissive cell line (Kim et al., 1993), was performed using Lipofectamine (Invitrogen) according to the manufacturer's directions. Approximately 24 h prior to transfection, 2×10^5 cells/well were seeded into a 6-well plate. Twenty-four hours later, the cells were viewed under a fluorescence microscope (Olympus) equipped with a standard FITC excitation filter. In some experiments, cells were trypsinized, placed on 10-well slides, and incubated in a humidified chamber. After further incubation to allow cells to reattach, slides were fixed in acetone for 10 min.

Immunofluorescent staining with anti-N protein mAbs

Cells on glass slides were fixed in acetone, air dried, and then stained with anti-N protein monoclonal antibodies (mAbs) SDOW-17, SR-30, JP-24, EP-147, VO-17, or MR-39 (a generous gift of Eric Nelson, South Dakota State University, Brookings, SD). All antibodies were diluted 1:100 in PBS with 5% FBS and incubated on slides for 2 h at room temperature. Mouse antibodies were detected with biotinylated anti-mouse IgG (Sigma) diluted 1:200 and incubated on slides for 1 h at room temperature, followed by a 20-min incubation with a 1:1,000 dilution of streptavidin-Texas Red (Molecular Probes). EGFP and Texas Red were visualized in the same cell under a fluorescence microscope using 488-nm (FITC) and 540-nm (rhodamine) excitation filters, respectively.

GST-bead pull-down assay

Plasmids mImp-α (PTAC58) and mImp-β (PTAC97), used to express the mouse importin-α and importin-β glutathione-*S*-transferase (GST) fusion proteins, respectively, were a generous gift from Dr. David Jans (John Curtin School of Medical Research, Canberra City, Australia). GST fusion proteins were expressed in *Escherichia coli* BL21 (Novagen) as described in Wootton et al. (1998). One hundred milliliters of log phase bacterial culture was induced with 1 mM IPTG, and after 3 h bacteria were pelleted

by centrifugation, resuspended in 5 ml of PBS, and sonicated on ice three times for 30 s. Triton X-100 was added to a final concentration of 1%, and proteins were disassociated for 30 min at 4°C with constant agitation. The insoluble fraction was removed by centrifugation, and the supernatant, containing the GST fusion proteins, was incubated with 100 μl of a 50% slurry of glutathione Sepharose 4B beads (Pharmacia) overnight at 4°C. Beads complexed to the GST fusion proteins were collected at 500*g* and washed three times in PBS and resuspended in a final volume of 250 μl of PBS containing 30% glycerol and 0.1% Triton X-100. To produce radiolabeled recombinant N protein, the N gene was cloned into pCITE and transcribed in vitro using T7 RNA polymerase MEGAscript (Ambion) according to the manufacturer's instructions. The resulting RNA was translated using the rabbit reticulocyte lysate translation kit (Promega) in the presence of 50 μCi/ml of [^{35}S]methionine (New England Nuclear, specific activity 3000 Ci/mmol). To produce the radiolabeled authentic viral N protein, PRRSV-infected MARC-145 cells were labeled at 30 h post-infection with 50 μCi/ml of [^{35}S]methionine for 8 h. Cells were washed and lysed using RIPA buffer (TNE with 0.5% Nonidet P-40, 0.5% sodium deoxycholate, 1% SDS).

Approximately equal amounts GST fusion proteins complexed to glutathione-Sepharose beads in a 20% slurry were incubated with radiolabeled N protein in incubation buffer (final concentration: 40 mM Hepes [pH 7.5], 100 mM KCl, 0.1% Nonidet P-40, and 20 mM β-mercaptoethanol) for 2 h or overnight at 4°C with constant agitation. The beads were rinsed four times with incubation buffer, boiled for 5 min in reducing Laemmli sample buffer, and analyzed by electrophoresis on 15% bis-polyacrylamide gel containing SDS. Gels were dried and exposed to Phosphorimager (Molecular Dynamics).

Acknowledgments

This work is dedicated to the memory of Harvey Tolleson. Special thanks go to Melissa Bien in South Dakota for technical contributions and to Susan Carpenter, Iowa State University, for contributing the EIAV Rev clones. This work was supported by PHS NIH Grant R15GM61319-01 and USDA NRI Grant 97-35204-5071.

References

Adachi, Y., Copeland, T.D., Hatanaka, M., Oroszlan, S., 1993. Nucleolar targeting signal of Rex protein of human T-cell leukemia virus type I specifically binds to nucleolar shuttle protein B-23. J. Biol. Chem. 268, 13930–13934.

Benfield, D.A., Nelson, E., Collins, J.E., Harris, L., Goyal, S.M., Robison, D., Christianson, W.T., Morrison, R.B., Gorcyca, D., Chladek, D., 1992. Characterization of swine infertility and respiratory syndrome (SIRS) virus (Isolate ATCC VR-2332). J. Vet. Diag. Invest. 4, 127–133.

Cavanagh, D., 1997. Nidovirales: a new order comprising Arterividae and Coronaviridae. Arch. Virol. 142, 629–633.

Chen, H., Wurm, T., Britton, P., Brooks, G., Hiscox, J.A., 2002. Interaction of the coronavirus nucleoprotein with nucleolar antigens and the host cell. J. Virol. 76, 5233–5250.

Dang, C.V., Lee, W.M., 1989. Nuclear and nucleolar targeting sequences of c-erb-A, c-myb, N-myc, p53, HSP70, and HIV tat proteins. J. Biol. Chem. 264, 18019–18023.

Favre, D., Studer, E., Michel, M.R., 1994. Two nucleolar targeting signals present in the N-terminal part of Semliki Forest virus capsid protein. Arch. Virol. 137, 149–155.

Fischer, U., Huber, J., Boelens, W.C., Mattaj, I.W., Luhrmann, R., 1995. The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. Cell 82, 475–483.

Frankhauser, C., Izaurralde, E., Adachi, Y., Wingfield, P., Laemmli, U.K., 1991. Specific complex of human immunodeficiency virus type 1 rev and nucleolar B23 proteins: dissociation by the Rev response element. Mol. Cell. Biol. 11, 2567–2575.

Gorlich, D., Mattaj, I.W., 1996. Nucleocytoplasmic transport. Science 271, 1513–1518.

Gorlich, D., Kutay, U., 1999. Transport between the cell nucleus and the cytoplasm. Annu. Rev. Cell Dev. Biol. 15, 607–660.

Henderson, B.R., Percipalle, P., 1997. Interactions between HIV rev and nuclear import and export factors: the rev nuclear localization signal mediates specific binding to human importin-β. J. Mol. Biol. 274, 693–707.

Henderson, B.R., Eleftheriou, A., 2000. A comparison of the activity, sequence specificity, and CRM1-dependence of different nuclear export signals. Exp. Cell Res. 256, 213–224.

Hicks, G.R., Raikhel, N.V., 1995. Protein import into the nucleus: an integrated view. Annu. Rev. Cell Dev. Biol. 11, 155–188.

Hiscox, J.A., Wurm, T., Wilson, L., Britton, P., Cavanagh, D., Brooks, G., 2001. The coronavirus infectious bronchitis virus nucleoprotein localizes to the nucleolus. J. Virol. 75, 506–512.

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., Pease, L.R., 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene. 77, 51–59.

Hope, T.J., 1999. The ins and outs of HIV Rev. Arch. Biochem. Biophys. 365, 186–191.

Hopp, T.P., Woods, K.R., 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA 78, 3824–3828.

Jans, D.A., Hubner, S., 1996. Regulation of protein transport to the nucleus: central role of phosphorylation. Physiol. Rev. 76, 651–685.

Jesperson, T., Mogens, D., Pedersen, F.S., 1997. Efficient non-PCR-mediated overlap extension of PCR fragments by exonuclease "end polishing." Biotechniques 23, 48–52.

Kim, H.S., Kwang, J., Yoon, L.J., Joo, H.S., Frey, M.L., 1993. Enhanced replication of porcine reproductive and respiratory syndrome (PRRS) virus in a homogeneous subpopulation of MA-104 cell line. Arch. Virol. 133, 477–483.

Kubota, S., Siomi, H., Satoh, T., Endo, S., Maki, M., Hatanaka, M., 1989. Functional similarity of HIV-I rev and HTLV-I rex proteins: identification of a new nucleolar targeting signal in rev protein. Biochem. Biophys. Res. Commun. 162, 963–970.

Kubota, S., Copeland, T.D., Pomerantz, R.J., 1999. Nuclear and nucleolar targeting of human ribosomal protein S25: common features shared with HIV-1 regulatory proteins. Oncogene 18, 1503–1514.

LaCasse, E.C., Lefebvre, Y.A., 1995. Nuclear localization signals overlap DNA- or RNA-binding domains in nucleic acid-binding proteins. Nucleic Acids Res. 23, 1647–1656.

Lai, M.M., 1990. Coronavirus: organization, replication, expression of genome. Annu. Rev. Micro. 44, 303–333.

Liu, J.L., Lee, L.F., Ye, Y., Qian, Z., Kung, H.J., 1997. Nucleolar and nuclear localization properties of a herpesvirus bZIP oncoprotein, MEQ. J. Virol. 71, 3188–3196.

Lixin, R., Efthymiadis, A., Henderson, B., Jans, D.A., 2001. Novel properties of the nucleolar targeting signal of human angiogenin. Biochem Biophys Res. Commun. 284, 185–193.

Mardassi, H., Massie, B., Dea, S., 1996. Intracellular synthesis, processing, and transport of proteins encoded by ORFs 5 to 7 of porcine reproductive and respiratory syndrome virus. Virology 221, 98–112.

Matthews, D.A., 2001. Adenovirus protein V induces redistribution of nucleolin and B23 from nucleolus to cytoplasm. J. Virol. 75, 1031–1038.

Meulenberg, J.J., van Nieuwstadt, A.P., van Essen Zandbergenl, A., Bos de Ruijter, J.N., Langeveld, J.P., Meloen, R.H., 1998. Localization and fine mapping of antigenic sites on the nucleocapsid protein N of porcine reproductive and respiratory syndrome virus with monoclonal antibodies. Virology 252, 106–114.

Moroianu, J., Blobel, G., Radu, A., 1996. The binding site of karyopherin alpha for karyopherin beta overlaps with a nuclear localization sequence. Proc. Natl. Acad. Sci. USA 93, 6572–6576.

Nakai, K., Kanehisa, M., 1992. A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics 14, 897–911.

Nelsen, C.J., Murtaugh, M.P., Faaberg, K.S., 1999. Porcine reproductive and respiratory syndrome virus comparison: divergent evolution on two continents. J. Virol. 73, 270–280.

Nelson, E.A., Christopher-Hennings, J., Drew, T., Wensvoort, G., Collins, J.E., Benfield, D.A., 1993. Differentiation of United States and European isolates of porcine reproductive and respiratory syndrome (PRRS) virus using monoclonal antibodies. J. Clin. Micro. 31, 3184–3189.

Paine, P.L., Moore, L.C., Horowitz, S.B., 1975. Nuclear envelope permeability. Nature 245, 109–114.

Plagemann, P.G.W., 1996. Lactate dehydrogenase-elevating virus and related viruses, in: Fields, B. (Ed.), Fields Virology, 3rd Edition. Lippincott–Raven, Philadelphia, pp. 1105–1120.

Roberts, B.L., Richardson, W.D., Smith, A.E., 1987. The effect of protein context on nuclear location signal function. Cell 50, 465–475.

Rodriguez, M.J., Sarraseca, J., Garcia, J., Sanz, A., Plana-Duran, J., Ignacio-Casal, J., 1997. Epitope mapping of the nucleocapsid protein of European and North American isolates of porcine reproductive and respiratory syndrome virus. J. Gen. Virol. 78, 2269–2278.

Rowland, R.R.R., Kervin, R., Kuckleburg, C., Sperlich, A., Benfield, D.A., 1999. The localization of porcine reproductive and respiratory syndrome virus nucleocapsid protein to the nucleolus of infected cells and identification of potential nucleolar localization signal sequence. Virus Res. 64, 1–12.

Siomi, H., Shida, H., Nam, S.H., Nosaka, T., Maki, M., Hantanaka, M., 1988. Sequence requirements for nucleolar localization of human T-cell leukemia virus type I pX proteins, which regulates viral RNA processing. Cell 55, 197–209.

Snijder, E.J., Meulenberg, J.J., 1998. The molecular biology of arteriviruses. J. Gen. Virol. 79, 961–979.

Stauber, R.H., Pavlakis, G.N., 1998. Intracellular trafficking and interactions of the HIV-1 Tat protein. Virology 252, 126–136.

Tijms, M.A., van der Meer, Y., Snijder, E.J., 2002. Nuclear localization of non-structural protein 1 and nucleocapsid protein of equine arteritis virus. J. Gen. Virol. 83, 795–800.

Wootton, S.K., Nelson, E.A., Yoo, D., 1998. Antigenic structure of the nucleocapsid protein of porcine reproductive and respiratory syndrome virus. Clin. Diagn. Lab. Immunol. 5, 773–779.

Wootton, S., Koljesar, G., Yang, L., Yoon, K.J., Yoo, D., 2001. Antigenic importance of the carboxy-terminal beta-strand of the porcine reproductive and respiratory syndrome virus nucleocapsid protein. Clin. Diagn. Lab. Immunol. 8, 598–603.

Wootton, S.K., Yoo, D., 2003. Homo-oligomerization of the porcine reproductive and respiratory syndrome virus nucleocapsid protein and the role of disulfide linkages. J. Virol. 77, 4546–4557.

Wurm, T., Chen, H., Hodgson, T., Britton, P., Brooks, G., Hiscox, J.A., 2001. Localization to the nucleolus is a common feature of coronavirus nucleoproteins, and the protein may disrupt host cell division. J. Virol. 75, 9345–9356.