The Nucleocytoplasmic Transport of Viral Proteins[*]

Qiong DING[†], Lei ZHAO[†], Hong GUO and Alan C. ZHENG[**]

(State Key Laboratory, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, China)

Abstract: Molecules can enter the nucleus by passive diffusion or active transport mechanisms, depending on their size. Small molecules up to size of 50-60 kDa or less than 10 nm in diameter can diffuse passively through the nuclear pore complex (NPC), while most proteins are transported by energy driven transport mechanisms. Active transport of viral proteins is mediated by nuclear localization signals (NLS), which were first identified in Simian Virus 40 large T antigen and had subsequently been identified in a large number of viral proteins. Usually they contain short stretches of lysine or arginine residues. These signals are recognized by the importin super-family (importin α and β) proteins that mediate the transport across the nuclear envelope through Ran-GTP. In contrast, only one class of the leucine-rich nuclear export signal (NES) on viral proteins is known at present. Chromosome region maintenance 1 (CRM1) protein mediates nuclear export of hundreds of viral proteins through the recognition of the leucine-rich NES.

Key words: Nuclear localization signal (NLS); Nuclear export signal (NES); Nuclear pore complex (NPC); Viral proteins

In eukaryotic cells, the translational machinery and metabolic systems of the cytoplasm are separated from the genetic material and transcriptional machinery of the nucleus by the nuclear envelope. Transport of macromolecules larger than 60 kDa between the nucleus and the cytoplasm is an active, energy-dependent and essential cellular process that is mediated by sequence-specific motifs, nuclear localization signals (NLS), and nuclear export signals (NES)[3]. This review summarizes the mechanisms having been identified so far for the nucleocytoplasmic transport of macromolecules and viral proteins.

NUCLEOCYTOPLASMIC TRANSPORT

The bidirectional transport of macromolecules between the nucleus and cytoplasm is a selective process that occurs exclusively through nuclear pore complexes (NPC)[1,10]. The NPC, is a 125-MDa macromolecular assembly of 50 to 100 polypeptides[4], consists of a cylindrical structure with an 8-fold symmetry and a central diameter of about 10 nm[3].

The NPC allows the passive diffusion of small molecules, including ions, metabolites, and globular proteins of up to ca. 60 kDa[9], without energy consumption. However, transport of larger proteins between the cytoplasmic and nuclear compartments is an active process and facilitated by specific soluble carrier proteins that are collectively referred to as "karyopherins", with "importins" and "exportins"[20]. The karyopherin family comprises about 14 known members in yeast and 19 in humans[35]. Many transport receptors are collectively called "β-karyopherins" because the majority of them are members of the importin β superfamily[20]. Although cargo proteins can bind directly to β-karyopherins, the interaction between the β-karyopherin and the cargo is mediated by the importin α during classical nuclear import [20].

The energy for nuclear transport is provided by Ran (a small Ras family of GTPases)[30], which cycles between a GTP- and a GDP-bound state. An asymmetric distribution of Ran-GTP and Ran-GDP between the nucleus and the cytoplasm controls the cargo import and export, and this gradient is maintained by various Ran associated regulatory factors[15]. Ran-GTP is concentrated in the nucleus, while Ran-GDP is concentrated in the cytoplasm, and Ran can impart directionality to nuclear transport by acting as a molecular switch that controls the binding and release of cargo for the compartmentalization[15,34].

In import pathways, Ran-GTP and substrates bind Kapßs competitively, allowing substrate binding in the cytoplasm and Ran-GTP-mediated release in the nucleus[21]. In the cytoplasm, a cargo molecule containing an NLS binds to an importing karyopherin (importin) in the absence of Ran-GTP. The import complex docks at the cytoplasmic side of an NPC and is then translocated into the nucleus [29]. Once inside

the nucleus, imported Kapßs bind the small GTPase Ran-GTP and release their substrates[35]. By interaction with the active, guanosine 5′-triphosphate (GTP)-loaded form of the small guanosine nucleotide-binding protein Ran, i.e. with Ran-GTP, the import complex is dissociated, and the cargo is set free to search the nuclear contents for a specific binding partner[29]. In contrast, in export pathways, Ran-GTP, substrates, and Kapßs bind cooperatively, resulting in substrate binding in the nucleus and release in the cytoplasm as the Ran bound nucleotide is hydrolyzed [21]. Whereas a cargo containing an NES forms a trimeric export complex with exporting karyopherin (exportin) and Ran-GTP in the nucleus. The export complex docks at the nuclear side of an NPC and is translocated into the cytoplasm[29]. In the cytoplasm, Ran-bound GTP is hydrolyzed, inducing the dissociation of the complex. Importing and exporting karyopherins, although quite large (~100 kDa), can pass through the NPC freely [29]. The diameter of the NPC can increase to some 30 nm in order to allow larger complexes come through. Once transport has been completed, the cargo is released and the transporters are recycled [3], while the import of Ran into the nucleus is facilitated by NTF2, a nuclear transport receptor belonging to the nuclear export factor (NXF) family of nuclear transport receptors[3,29].

Full understanding of nucleocytoplasmic transport is important for both medical science and biotechnology. The role of nucleocytoplasmic transport plays in the biogenesis of pathological conditions and the possibilities for new therapeutic strategies have been intensively studied. In addition, the NPC provides an intriguing paradigm of a molecular machine to screen native protein molecules with high efficiency [29].

CLASSICAL NUCLEAR LOCALIZATION SIGNAL AND IMPORTINS

Many proteins must enter the nucleus to fulfill their physiological roles[4].The best understood system for the transport of macromolecules is the classical nuclear import pathway, which was first proposed when a nuclear targeting signal in the large T antigen of the simian virus 40 (SV40) was characterized[16]. Since then, several pathways for nucleocytoplasmic transport have been described[20].

NLS motifs consist of grouped positively charged amino acids and can be mono- or bipartite. 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 an amino acid sequence containing three basic residues out of four, such as PKKKRKV[32]. The 'pat4' NLS and the 'pat7' NLS belongs to monopartite signals. The third type of NLS, known as a "bipartite" motif, consists of two groups of 2-3 basic amino acids separated by about 10 residues[14], such as KRPAATKKAGQAKKKKLD[7]. This information has been incorporated into computer programs to identify NLS motifs, but cannot always successfully predict the localization of a protein to the nucleus[32]; proteins without an NLS can be co-transported with other nuclear proteins and proteins that contain NLS sequences may remain cytoplasmic, especially if the NLS sequence is blocked or buried within the protein[31]. It is also possible that the sequence is not linear, in which case the basic residues may form a conformational epitope on the surface of the protein[3]. Furthermore, amino acid replacement analyses revealed that the consensus basic patterns of the classical NLS are not essential for activity, thereby generating more unconventional patterns, including redox-sensitive NLS. These results explain the causes of the NLS diversity [17]. Although the consensus sequences of the classical NLS have been defined, there are still many NLS that do not match the consensus rule and many nonfunctional sequences that match the consensus. It is reported that M9 is a novel type of nuclear localization domain as it does not contain sequences similar to the classical basic-type NLS, which is necessary and sufficient for nuclear localization[33]. Rather, it is a new NLS, PY-NLS, a C-terminal consensus sequence $R/K/H-X_{(2-5)}-P-Y$ (where X is any residue), that recognized by Kapβ2[21]. Many studies have shown that the M9-NLS peptide is both necessary and sufficient for Kapβ2 mediated nuclear import [33].

In our previous studies, the subcellular localization of bovine herpesvirus type 1 (BHV-1) VP8 and VP22 have been characterized and their nuclear localization signals had also been identified. The results indicated that VP8 contained a pat4 NLS and VP22 contained a functional nonclassical NLS[36,37]. Recently, we have identified and confirmed a functional NLS and nucleolar localization signal (NoLS) of BHV-1 infected cell protein 27 (BICP27)[13].

NLS confers interaction with members of the importins. There are two types of importins, α and β, both of which are importin β super-family numbers. Importin α serves as an adaptor that links cargos and importin β1 and recognizes NLS within the cargos[3]. Importin α recognizes two classes of classical NLS: monopartite NLS and bipartite NLS[20]. Once the cargo is bound by importin α, the complex can be recognized and bound by importin β that subsequently binds to the fibrils of the NPC and is responsible for the actual translocation [2]. Nuclear import of proteins

is generally initiated by the formation of a ternary complex with importin α, importin β1, and a cargo, where importin β1 docks the complex to the NPC to release the cargo into the nucleus through the binding of Ran-GTP to importin β1 [17]. This is not the only way molecules can be transported into the nucleus [25] but it is the most common way. Importin α has several isoforms that recognize different variants of NLS but they are all recognised by a single importin β1. Therefore, these isoforms could be considered as adaptors for different signal transmission to the importin β1 molecule[6]. In the "classical" or "canonical" import pathway, an NLS-containing cargo is bound first by importin α. In the second step, the cargo-importin-α complex is bound by importin β1 and only the trimeric complex is translocated into the nucleus finally (Fig. 1)[29].

Each import Kapβ appears to bind distinct sets of substrates, suggesting that each Kapβ recognizes different NLS. However, large sequence diversity among various substrates has hindered identification of NLS for most Kapβs, and it remains extremely difficult to predict NLS in candidate import substrates[21]. It is reported that six different NLS classes specifically bind to distinct binding pockets of importin α[17]. The well-defined consensus patterns and properties of importin α-dependent NLS will provide useful information for NLS identification.

The nuclear import pathways benefit viral pathogens to achieve efficient nuclear import of both viral particles during the initial stages of infection and individual viral proteins following their synthesis in the cytoplasm [4].

NUCLEAR EXPORT SIGNAL AND EXPORTINS

NES have an important regulatory function in the subcellular localization of viral proteins, and also have an impact on transcription and other nuclear processes[19]. Although several classes of NLS have been characterized, only one class of NES, the leucine-
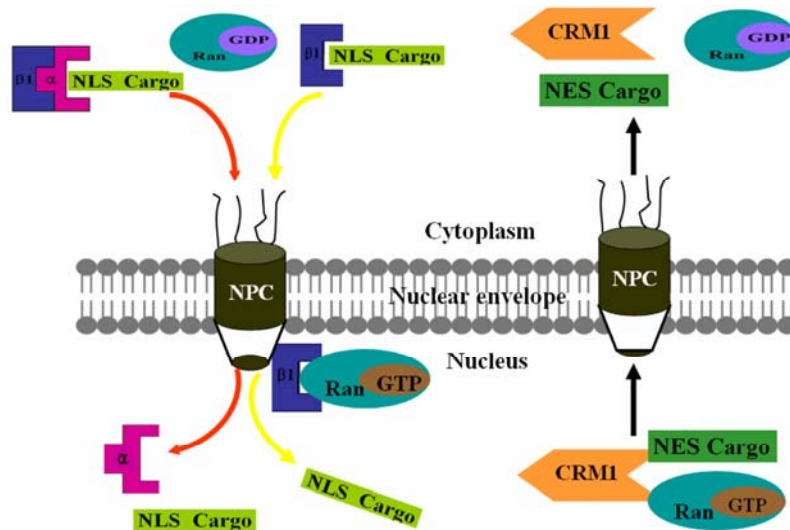

Fig. 1. Simplified model of importin/exportin-mediated nucleocytoplasmic transport of proteins. In the cytoplasm, importin-α (α) binds to the NLS-containing cargo protein and forms a ternary complex with importin-β1 (β1) to enter into the nucleus through the NPC. Some NLS cargo can directly bind to importin-β1. Once within the nucleus, binding of Ran-GTP to importin-β1 elicits a conformational change resulting in cargo release into the nucleus. For nuclear export, Ran-GTP stimulates binding of exportin CRM1 to the NES cargo in the nucleus and the complex is exported to the cytoplasm, where hydrolysis of Ran-GTP to Ran-GDP.

rich NES, is known at present[8]. The 10-15-residue leucine-rich NES were first characterized in the human immunodeficiency virus type 1 (HIV-1) Rev protein and cyclic-AMP-dependent protein kinase inhibitor[8]. Several mutagenesis and computational studies have identified a consensus sequence Φ-X$_{(2-3)}$-Φ-X$_{(2-3)}$-Φ-X-Φ, in which Φ refers to Lys, Val, Ile, Phe or Met, X represents any amino acid, and the numbers in parentheses denote the number of repeats. Although this sequence matches most known leucine-rich NES, it is so broadly distributed that it also exists in most helix-containing proteins[8]. Prolines situated between the hydrophobic residues can disrupt the function of nuclear export [8]. Moreover, NES that do not belong to the family of leucine-rich NES have also been described for various proteins, as demonstrated by the atypical NES in the hepatitis D antigen HDAg-L[22].

Direct interaction of proteins bearing leucine-rich NES with the nuclear export factor chromosome region maintenance 1 (CRM1), also known as exportin-1, is very well documented and essential for the active nuclear export of these proteins [11, 12, 27]. CRM1 mediates nuclear export of hundreds of proteins through the recognition of the leucine-rich NES. The NES recognized by CRM1 is the consensus sequence Φ-X$_{(2-3)}$-Φ-X$_{(2-3)}$-Φ-X-Φ [3]. The large hydrophobic residues are the most conserved in NES which suggests that the interaction with CRM1 is mediated through these hydrophobic residues [19]. CRM1, contains armadillo repeats, belongs to the importin β super family, and is the regular exportin for RNAs, proteins and ribonucleoproteins complexes. Like importins, there exist exportins that directly recognize various NES and then bind to the second exportin molecule that interacts with the nuclear pore[3].

CRM1 cannot leave the nucleus with its cargo without the assistance of Ran [3]. It's believed that Ran-GTP and leucine-rich NES bind cooperatively to CRM1 upon formation of a ternary CRM1-Ran-GTP-NES complex (Fig. 1)[11, 28]. Moreover, the NES substrate RanBP3/Yrb2p functioning as an essential export cofactor, enhances the affinity of CRM1 for Ran-GTP [19].

Recently, a study showed that CRM1 binds an export substrate snurportin 1 (SNUPN) and Ran-GTP cooperatively in the nucleus[8]. Multipartite recognition in SNUPN may be used more generally to increase CRM1 binding beyond the usually low affinity leucine-rich NES[8]. The crystal structure of the CRM1–SNUPN complex has been presented [8]. This trimeric complex is translocated from the nucleus to the cytoplasm through the NPC [9]. Soon after the complex has left the nucleus, GTP is hydrolysed by Ran through interaction with Ran-GAP (Ran GTPase Activating Protein) which is associated with cofactors RanBP1 and RanBP2 (Ran Binding Proteins) that are localized on the fibrils on the cytoplasmic side of the NPC. After hydrolysis of GTP, the conformation of Ran changes and the complex falls apart and the exportins and Ran-GDP go back to the nucleus leaving the cargo remain in the cytoplasm[3].

Several ways of regulating NES-dependent export have been reported, such as masking/unmasking of NES phosphorylation and even disulfide bond formation resulted from oxidation [19].

Leptomycin B (LMB), a metabolite that disrupts the interaction of NE with the CRM1 receptor by binding to a cysteine residue, is localized in the central domain of the receptor[18]. Antifungal antibiotic LMB inhibits nuclear export by alkylating Cys 528 of human CRM1[18]. It is unclear whether LMB blocks substrate binding sterically or by a conformational change [8].

Besides, LMB is used to identify CRM1 involved nuclear export pathways. Although the CRM1-dependent transport is the best-characterized pathway, CRM1-independent nuclear export mechanisms have been proposed for various proteins [5, 22, 23].

In conclusion, it is hard to define NES by its hydrophobic pattern only. For correct classification of leucine-rich regions with and without NES activity, detection of less apparent sequence features, secondary structure, flexibility and/or surface exposure is still necessary[19]. In addition, there are more complex signal sequences which exist, some of which confer both import and export functions [24]. In all cases, NLS and NES are constant constituents of the cargo and not removed during transport. They can be located at any site within a polypeptide and frequently occur in more than one copy per molecule[26].

Both BHV-1 VP8 and VP22 were identified to contain an NES and located in leucine-rich regions, consisting of amino acids ^{485}LSAYLTLFVAL^{495} and ^{204}LDRMLKSAAIRIL^{216} respectively. Moreover, the nuclear export activity of VP22 NES was completely blocked by LMB [36, 37]. Therefore, VP22 is a LMB-sensitive protein, which suggests that VP22 mediates nuclear export through a leucine-rich NES by a CRM-1-dependent pathway [37].

CONCLUSION

Many viral proteins shuttle between the cytoplasm and the nucleus to exert their specific biological functions. The transport of molecules across the nuclear membrane relies on the NPC, the transport carriers (including importins and exportins) and the Ran. Karyopherins bind to their cargoes by recognition of specific NLS or NES. Although transport mechanisms of molecules have been elucidated to a

certain extent, many details remain unknown. With further investigation of the molecular basis of nuclear transport, it will significantly promote the design of anti-virus strategies and gene vector therapy.

References

1. **Allen T D, Cronshaw J M, Bagley S***, et al.* 2000. The nuclear pore complex: mediator of translocation between nucleus and cytoplasm. **J Cell Sci,** 113 (Pt 10): 1651-1659.

2. **Bayliss R, Littlewood T, Stewart M.** 2000. Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. **Cell,** 102: 99-108.

3. **Boulo S, Akarsu H, Ruigrok R W***, et al.* 2007. Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes. **Virus Res,** 124: 12-21.

4. **Bremner K H, Seymour L W, Pouton C W.** 2001. Harnessing nuclear localization pathways for transgene delivery. **Curr Opin Mol Ther,** 3: 170-177.

5. **Chen T, Brownawell A M, Macara I G.** 2004. Nucleocytoplasmic shuttling of JAZ, a new cargo protein for exportin-5. **Mol Cell Biol,** 24: 6608-6619.

6. **Conti E, Muller C W, Stewart M.** 2006. Karyopherin flexibility in nucleocytoplasmic transport. **Curr Opin Struct Biol,** 16: 237-244.

7. **Dingwall C, Robbins J, Dilworth S M***, et al.* 1988. The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. **J Cell Biol,** 107: 841-849.

8. **Dong X, Biswas A, Suel K E***, et al.* 2009. Structural basis for leucine-rich nuclear export signal recognition by CRM1. **Nature,** 458: 1136-1141.

9. **Eulalio A, Nunes-Correia I, Carvalho A L***, et al.* 2004. Two African swine fever virus proteins derived from a common precursor exhibit different nucleocytoplasmic transport activities. **J Virol,** 78: 9731-9739.

10. **Fahrenkrog B, Aebi U.** 2002. The vertebrate nuclear pore complex: from structure to function. **Results Probl Cell Differ,** 35: 25-48.

11. **Fornerod M, Ohno M, Yoshida M***, et al.* 1997. CRM1 is an export receptor for leucine-rich nuclear export signals. **Cell,** 90: 1051-1060.

12. **Fukuda M, Asano S, Nakamura T***, et al.* 1997. CRM1 is responsible for intracellular transport mediated by the

nuclear export signal. **Nature,** 390: 308-311.

13. **Guo H, Ding Q, Lin F,** *et al.* 2009. Characterization of the Nuclear and Nucleolar Localization Signals of Bovine Herpesvirus-1 Infected Cell Protein 27. **Virus Res**, 145: 312-320.

14. **Hodel M R, Corbett A H, Hodel A E.** 2001. Dissection of a nuclear localization signal. **J Biol Chem,** 276: 1317-1325.

15. **Kalab P, Weis K, Heald R.** 2002. Visualization of a Ran-GTP gradient in interphase and mitotic Xenopus egg extracts. **Science,** 295: 2452-2456.

16. **Kalderon D, Richardson W D, Markham A F,** *et al.* 1984. Sequence requirements for nuclear location of simian virus 40 large-T antigen. **Nature,** 311: 33-38.

17. **Kosugi S, Hasebe M, Matsumura N,** *et al.* 2009. Six classes of nuclear localization signals specific to different binding grooves of importin alpha. **J Biol Chem,** 284: 478-485.

18. **Kudo N, Matsumori N, Taoka H,** *et al.* 1999. Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region. **Proc Natl Acad Sci USA,** 96: 9112-9117.

19. **la Cour T, Kiemer L, Molgaard A,** *et al.* 2004. Analysis and prediction of leucine-rich nuclear export signals. **Protein Eng Des Sel,** 17: 527-536.

20. **Lange A, Mills R E, Lange C J,** *et al.* 2007. Classical nuclear localization signals: definition, function, and interaction with importin alpha. **J Biol Chem,** 282: 5101-5105.

21. **Lee B J, Cansizoglu A E, Suel K E,** *et al.* 2006. Rules for nuclear localization sequence recognition by karyopherin beta 2. **Cell,** 126: 543-558.

22. **Lee C H, Chang S C, Wu C H,** *et al.* 2001. A novel chromosome region maintenance 1-independent nuclear export signal of the large form of hepatitis delta antigen that is required for the viral assembly. **J Biol Chem,** 276: 8142-8148.

23. **Lischka P, Rosorius O, Trommer E,** *et al.* 2001. A novel transferable nuclear export signal mediates CRM1-independent nucleocytoplasmic shuttling of the human cytomegalovirus transactivator protein pUL69. **EMBO J,** 20: 7271-7283.

24. **Michael W M, Choi M, Dreyfuss G.** 1995. A nuclear export signal in hnRNP A1: a signal-mediated,

temperature-dependent nuclear protein export pathway. **Cell,** 83: 415-422.

25. **Miyamoto Y, Imamoto N, Sekimoto T,** *et al.* 1997. Differential modes of nuclear localization signal (NLS) recognition by three distinct classes of NLS receptors. **J Biol Chem,** 272: 26375-26381.

26. **Nair R, Carter P, Rost B.** 2003. NLSdb: database of nuclear localization signals. **Nucleic Acids Res,** 31: 397-399.

27. **Ossareh-Nazari B, Bachelerie F, Dargemont C.** 1997. Evidence for a role of CRM1 in signal-mediated nuclear protein export. **Science,** 278: 141-144.

28. **Ossareh-Nazari B, Dargemont C.** 1999. Domains of Crm1 involved in the formation of the Crm1, RanGTP, and leucine-rich nuclear export sequences trimeric complex. **Exp Cell Res,** 252: 236-241.

29. **Peters R.** 2006. Introduction to nucleocytoplasmic transport: molecules and mechanisms. **Methods Mol Biol,** 322: 235-258.

30. **Quimby B B, Dasso M.** 2003. The small GTPase Ran: interpreting the signs. **Curr Opin Cell Biol,** 15: 338-344.

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

32. **Rowland R R, Yoo D.** 2003. Nucleolar-cytoplasmic shuttling of PRRSV nucleocapsid protein: a simple case of molecular mimicry or the complex regulation by nuclear import, nucleolar localization and nuclear export signal sequences. **Virus Res,** 95: 23-33.

33. **Siomi H, Dreyfuss G.** 1995. A nuclear localization domain in the hnRNP A1 protein. **J Cell Biol,** 129: 551-560.

34. **Smith A E, Slepchenko B M, Schaff J C,** *et al.* 2002. Systems analysis of Ran transport. **Science,** 295: 488-491.

35. **Suel K E, Chook Y M.** 2009. Kap104p imports the PY-NLS-containing transcription factor Tfg2p into the nucleus. **J Biol Chem**, 284：15416－15424.

36. **Zheng C, Brownlie R, Babiuk L A,** *et al.* 2004. Characterization of nuclear localization and export signals of the major tegument protein VP8 of bovine herpesvirus-1. **Virology,** 324: 327-339.

37. **Zheng C, Brownlie R, Babiuk L A,** *et al.* 2005. Characterization of the nuclear localization and nuclear export signals of bovine herpesvirus 1 VP22. **J Virol,** 79: 11864-11872.