I. Introduction

Chapter 1. Basics on general nucleocytoplasmic transport

Chapter 2. Biogenesis and nuclear export of cellular mRNA transcripts

Chapter 3. Biogenesis and nuclear export of retroviral mRNA transcripts

Chapter 4. The questions addressed in this PhD thesis

General Overview

Nucleocytoplasmic transport in eukaryotes

The nucleus is an organelle enclosed by a double membrane that is continuous with the rough endoplasmic reticulum. It exists exclusively in eukaryotic cells, where it is generally disrupted during mitosis. At interphase, it segregates the genetic material from other cellular components. Consequently, there is a spatial segregation of DNA replication and RNA transcription, which take place in the nucleus, from protein synthesis, occurring in the cytoplasm. This likely confers advantages to eukaryotic cells, which are able to handle large amounts of DNA, control gene expression patterns according to spatial and temporal cellular requirements, and produce post-transcriptionally modified RNA transcripts. It however, also poses a problem in that eukaryotic cells have to accommodate a continuous flow of molecules through the nuclear membrane: RNA transcripts must reach the cytoplasm where polysomes reside, whereas nuclear proteins must localize to their sites of action in the nucleus. To fulfill this need, eukaryotic cells perform recurrent rounds of regulated nucleocytoplasmic transport events.

Chapter 1. Basics on general nucleocytoplasmic transport

1.1) Nucleocytoplasmic transport

Nucleocytoplasmic transport is a facilitated diffusion process, which is mediated by specific soluble receptors that recognize their cargo and traverse the nuclear pore complexes (NPCs) in an energy-independent manner (reviewed by Fried and Kutay 2002).

1.2) The NPC structure

NPCs are aqueous channels embedded in the nuclear envelope. They are multiprotein assemblies of approximately 125 MDa in vertebrates (Reichelt *et al*., 1990) and 65 MDa in yeast (Rout and Blobel 1993). A proteomic analysis of the NPC has led to an approximate description of its stoichiometrical composition, both in yeast (Allen *et al.,* 2001) and in vertebrates (Cronshaw *et al*., 2002). A low-resolution three-dimensional view of the NPC has been obtained by electron microscopy and single particle reconstruction methods (Akey and Radermacher 1993; Goldberg and Allen 1993; Yang *et al*., 1998).

The membrane-embedded central core structure is an aqueous channel 120 nm in diameter and 70 nm thick. The nuclear pore channel is 60 nm long and has a diameter of 9 nm at rest, but can expand up to 25 nm during active transport. The NPC is asymmetric regarding its cytoplasmic and nuclear sides. The reason being that it is decorated by filament extensions, protruding outwards, about 30-50 nm from the central core. The filaments facing the cytoplasm are flexible, whereas those at the nucleoplasmic side referred to as the nuclear basket, tend to converge at their apices, giving the structure a cage-like appearance (figure 1) (reviewed by Adam 2001; Bayliss *et al*., 2000).

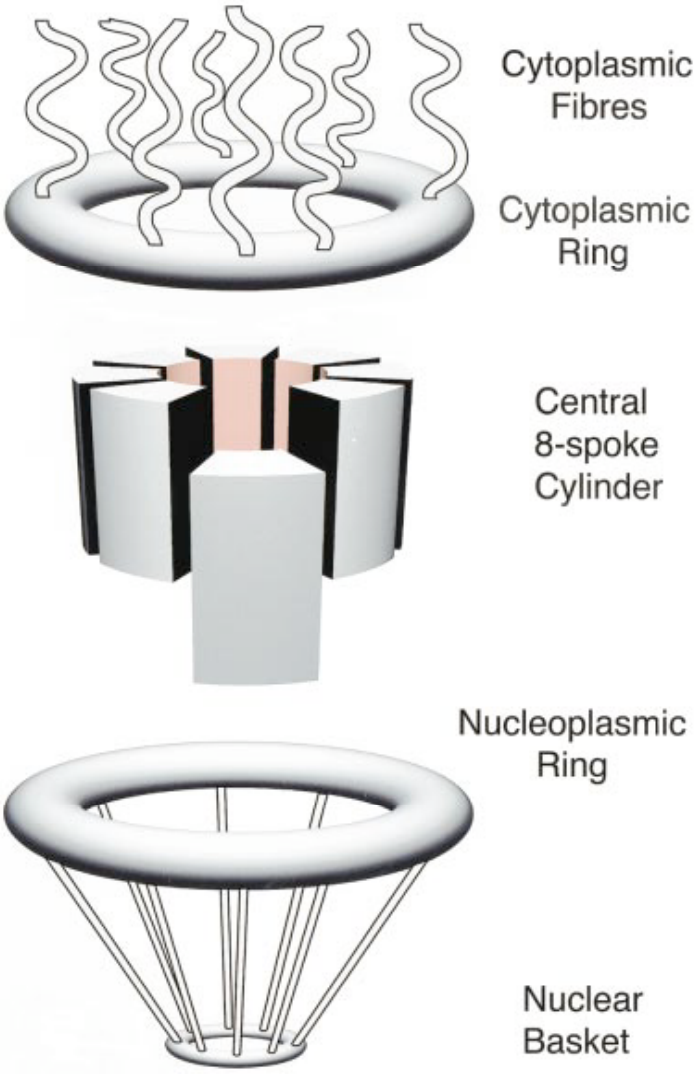


Figure 1. Schematic representation of the nuclear pore complex. Adapted from Bayliss *et al*., 2000.

1.3) Translocation across the NPC

The NPC is a 8-fold symmetrical structure, built up from repetitions of 30-50 different proteins, collectively called nucleoporins (Nups). Most Nups are stationary, but some have been described as mobile proteins. The role of Nups is dual, as they provide the structural framework of the NPC and represent transient docking sites for the translocation of nuclear transport receptors across the nuclear pore channel. The majority of the Nups feature repeated phenylalanine-glycine (FG)-rich sequences, in the form of FG, FxFG or GLFG peptide motifs. The FG-repeats are thought to represent a hydrophobic foot-path, for the active translocation of cargo/transport receptor complexes, across the nuclear pore channel (reviewed by Nakienly and Dreyfuss 1999; Görlich and Kutay 1999; Fried and Kutay 2002).

Different models have been proposed to date, from which the Brownian model and the Molecular Sieve model stand as the more plausible (figure 2). The aim of these models is to provide a rationale for the regulation, selectivity and energy-independence of transport receptor translocation across the NPC (reviewed by Fried and Kutay 2002).

The Brownian model is based on the interaction between transport receptors and the Nups that build-up the external filaments, on both sides of the NPC. It proposes that these interactions, which are transient, increase the residence time of transport complexes in the vicinity of the nuclear pore channel. This then leads to the translocation event by Brownian motion, where the nuclear receptor would continue its way through the nuclear pore channel by means of weak interactions with the Nups that face the nuclear pore lumen (figure 2). This model implies that only proteins capable of interacting with Nups, would access the NPC-tunnel and gives a rationale for the specificity of the translocation process. It fails however, to explain the transport of large ribonucleoprotein complexes (Rout *et al*., 2000).

The Molecular Sieve model, on the other hand, assumes the existence of a hydrophobic meshwork at the centre of and along the nuclear pore channel. This meshwork would be based on the interaction of Nups, with each other, through their

FG-repeat motifs, and would act as a permeability barrier. Only receptors compatible with these Nup-Nup interactions would disrupt such solubility-based barrier and gain access into the translocation tunnel (figure 2). This partition-based access into the nuclear pore channel would then be inaccessible to macromolecules unable to interact with Nups, but will still provide a solubility-based means for specific transport receptors to travel along the nuclear pore channel. Even if this model accounts for the transport of large complexes, the existence of such a meshwork has yet to be proven (Ribbeck and Görlich 2001).

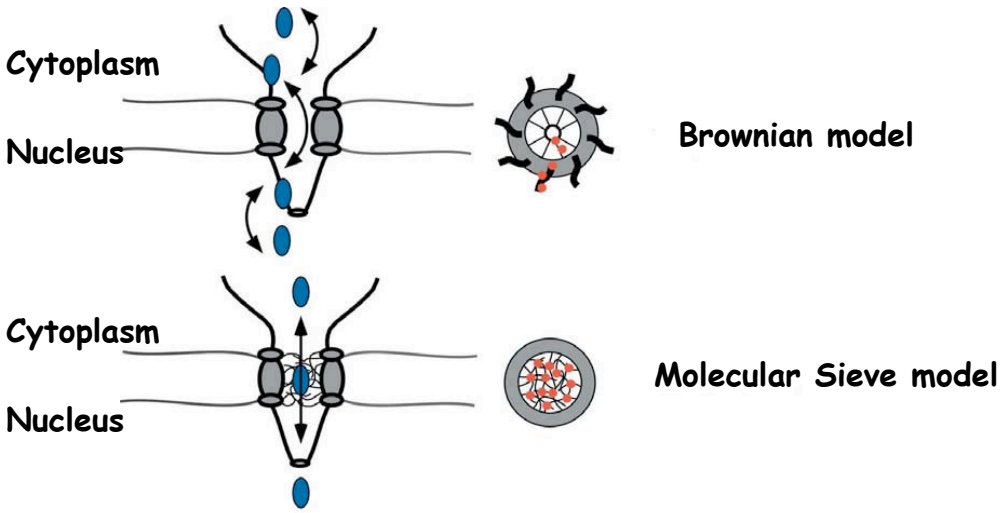


Figure 2. Schematic representation of the Brownian and Molecular Sieve models. Putative transport receptors are shown in blue and surface-exposed Nups are shown in orange. Adapted from Fried and Kutay 2003.

1.4) The nuclear transport receptors

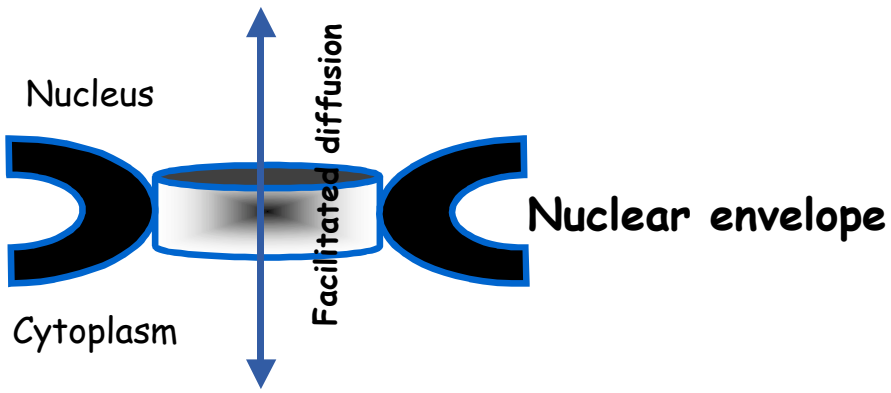
Three types of nuclear transport receptors, which participate in distinct nucleocytoplasmic transport pathways, have been described to date. They include the importin β-like family of proteins, the nuclear transport factor-2 (NTF2) and

TAP/nuclear export factor 1 (NXF1) (figure 3). Members of the importin β-like family of proteins are responsible for the export and import of the majority of nuclear-shuttling proteins, non-coding cellular RNA transcripts (e.g. rRNA, snRNA, tRNA and microRNA) and certain types of viral mRNA transcripts (mRNA transcripts of complex retroviruses, such as HIV-1), to their site of function. The transport receptor NTF2 imports RanGDP to the nucleus, where RanGTP is regenerated. Finally, the protein TAP/NXF1 participates in the nuclear export of cellular mRNA transcripts and certain types of viral mRNA transcripts (mRNA transcripts of type-D simple retroviruses). These three receptor types share in common their ability to interact directly with Nups, and to shuttle between the nucleus and the cytoplasm. Their functional role is to facilitate the passage of protein or RNA cargo through the NPC. Although they share a common NPC-translocation mechanism, different nuclear transport receptors are characterized by different cargo-uptake modalities and transport directionality mechanisms (reviewed by Fried and Kutay 2002).

1.5) Karyopherins: the importin β-like family of transport receptors

Most of the nuclear transport receptors described to date are members of the importin-β-like family of proteins, also known as karyopherins. These proteins are referred to as exportins or importins, depending on whether they direct cargoes from or to the nucleus (reviewed by Strom and Weiss 2001). Karyopherins bind their cargo either directly by recognizing specific peptide signals (e.g., nuclear export/NES or nuclear import/ NLS signals) (Dingwall and Laskey 1991) or through protein adaptors, such as importin α, that in turn recognize a given localization signal (a NLS, in the case of importin α) (Weis *et al*., 1995). All karyopherins contact the FG-rich motifs of the Nups through a surface-exposed hydrophobic patch (Kutay *et al*., 1997; Seedorf *et al*., 1999; Damelin and Silver 2000). In addition, they all interact with the small GTPase Ran (Görlich *et al*., 1997).

Export receptors	Cargo
Karyopherin family	
Exportin 1/Crm1	proteins, snRNA, rRNA, complex retroviral mRNA (HIV)
Exportin T	tRNA
Exportin 5	pre-micro-RNA
NXF family	
TAP/NXF1:Mex67p	cellular mRNA, type D simple retroviral mRNA



Import receptors	Cargo
Karyopherin family	
Importin β-like receptors	proteins, snRNA
NTF2	RanGDP

Figure 3. Nuclear transport receptors. Highlighted in blue are TAP and NTF2, the transport receptors that do not belong to the karyopherin family of proteins.

1.5.1) Structure of karyopherins

To date, structural information has been obtained for several importins, such as importin β1 (Cingolani *et al*., 1999) and importin β2/transportin (Chook and Blobel 1999). Importins fold into a spiral-like right-handed superhelical structure composed of two arches representing the N-terminal and C-terminal halves of the receptors. They are built up from tandem HEAT repeats, which are 40 amino acid long motifs that form two α helices connected by a short turn (reviewed by Chook *et al*., 1999; Conti and Izaurralde 2001). In general, karyopherins have a similar shape and are flexible molecules that may stretch and compress their domains by twisting, as shown by the available atomic structures of importins (reviewed by Chook *et al*., 1999; Conti and Izaurralde 2001), and the X-ray scattering solution studies for various importins and exportins (Fukuhara *et al*., 2004). Figure 4 shows a ribbon representation of importin β, in the context of various ligands.

1.5.2) Ran: the regulator of karyopherin-mediated nuclear transport

The GTPase Ran was first described in 1993 as a Ras-related signalling protein with GTP/GDP binding properties (Coutavas *et al*., 1993). Ran features a conserved G (guanidine nucleotide-binding or GTPase) domain with two flexible switch loops (switch I and II) and a C-terminal extension that assumes different conformations in the GTP- and GDP-bound states (reviewed by Fried and Kutay 2002). Successful rounds of karyopherin-mediated import and export events are regulated by an asymmetric distribution of RanGTP between the nucleus and the cytoplasm (Izaurralde *et al*., 1997; Nachury and Weis 1999).

During karyopherin-mediated nuclear transport, importin/cargo complexes form in the cytoplasm at low concentrations of RanGTP. The complex translocates into the nucleus where it is disrupted by the high levels of nuclear RanGTP. Exportins on the other hand, load their cargo cooperatively in the presence of nuclear RanGTP and translocate to the cytoplasm, where RanGTP is hydrolyzed to RanGDP and the cargo is released. The differential gradient of RanGTP nuclear/RanGDP cytoplasmic is exerted by the interplay of two enzymes: a nuclear guanidine nucleotide exchange

factor, RanGEF (RCC1) (Bischoff and Postingl 1991) and a cytoplasmic RanGTPase activating protein, RanGAP (Bischoff *et al*., 1994), which is co-activated by Ran Binding Protein (RBP1) (Bischoff *et al*., 1995), and probably in some cases by RBP2 (Yokoyama *et al*., 1995; Wu *et al*., 1995) (figure 5). RanGEF/RCC1 localizes to the chromatin where it promotes the interaction of Ran with GTP. Cytoplasmic RanBP1 on the other hand, binds to RanGTP and promotes the dissociation of RanGTP/cargo complexes. This triggers a conformation change on Ran, with a concomitant increase of its GTPase activity. During recurrent karyopherin-based nuclear transport events, RanGTP accumulates in the cytoplasm, where it is converted to RanGDP and is recycled to the nucleus as such, by the transport receptor NTF2. Inside the nucleus, RanGDP is rapidly converted to RanGTP. This cycle of events sets the basis for constitutive rounds of karyopherin-based nuclear transport events within the context of a dividing cell (reviewed by Nakienly and Dreyfuss 1999; Görlich and Kutay 1999 and Fried and Kutay 2002).

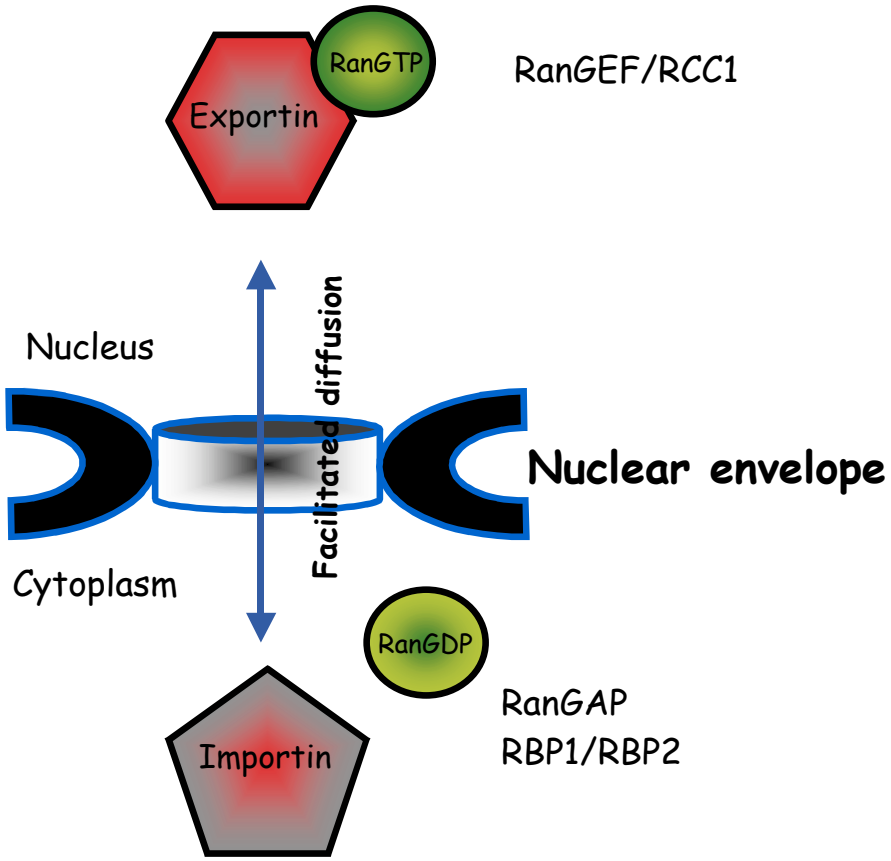
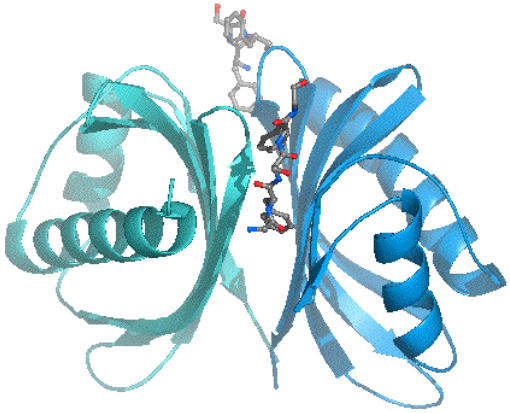


Figure 5. RanGTP/GDP gradient. The asymmetric distribution of RanGTP and GDP is given by a set of enzymes that localize specifically to the nucleus and to the cytoplasm respectively.

1.6) NTF2: the transport receptor of RanGDP

NTF2 binds to and recycles RanGDP into the nucleus, where RanGTP is regenerated (Ribbeck *et al*., 1998). Structural data is available for free NTF2 (Bullock *et al*., 1996) and NTF2 in complex with both RanGDP (Stewart *et al*., 1998) and an FG-rich peptide (figure 6) (Bayliss *et al*., 2002b). NTF2 is a homodimer that recognizes the conformation of Ran's switch II, when Ran is bound to GDP (Stewart *et al*., 1998). A complete description of the surface of NTF2 that interacts with nucleoporin FG-rich repeats, has been mapped by NMR chemical shift perturbation assays (Morrison *et al*., 2003) and complemented the available crystallographic data. Basically, NTF2 homodimers bind to the NPC through an extended surface-exposed hydrophobic patch centred on tryptophan 7 of each subunit (Morrison *et al*., 2003).

NTF2 homodimer- FG-rich repeat



Bayliss *et al*., 2002b

Figure 6. Ribbon representation of NTF2. Each domain is represented in different colour, marine and cyan. The FG-rich repeats are shown in grey.