CHAPTER 4

Nuclear Import and Export Signals

Toshihiro Sekimoto, Jun Katahira and Yoshihiro Yoneda

Eukaryotic cells are separated into two large compartments, namely the nucleus and the cytoplasm, by the nuclear envelope. As a result, macromolecules including RNAs, which are transcribed in the nucleus and nuclear proteins, which are translated in the cytoplasm must cross the double lipid bilayer to reach the intracellular sites where they function. In addition, cumulative evidence suggests that trafficking between the nucleus and the cytoplasm is rather dynamic and some proteins and RNAs cross the nuclear envelope again after being transported to one compartment.

The nuclear pore complex (NPC), a huge proteinaceous channel composed of 50 to 100 different proteins that are collectively termed nucleoporins, is the only known functional path through which soluble molecules are transported.[1,2] The functional diameter of the aqueous channel of the nuclear pore, which is estimated to be 9 to 10 nm, allows the non-selective passive diffusion of small molecules such as ions and low molecular weight metabolites. Small molecules can freely traverse the NPC in both directions in a concentration gradient-dependent manner. On the contrary, molecules with molecular weights in excess of 40 to 60 kDa cannot pass through the NPC by simple diffusion. To accomplish "active transport", which is probably synonymous with transport against a concentration gradient via the consumption of energy, it is generally thought such molecules are directed to the appropriate compartments by specific mechanisms.

Indeed, numerous efforts have revealed that even small proteins, which are smaller than the diffusion limit, are subjected to both the active nuclear import and export. As expected the transport process has been found to be mediated by transferable signals, harbored within the transported substrates themselves. Such signal sequences for nuclear import and export are called nuclear localization signals (NLS) and nuclear export signals (NES), respectively. The aim of this Chapter is to introduce the definition and the mechanism of how such signals function.

Definition of Nuclear Import and Export Signals

During the nuclear import or export process, transport substrates must pass through the NPC by interacting with nucleoporins. Conceptually this may be accomplished by direct interactions of NLS or NES with nucleoporins. The recent identification of various receptors for NLS and NES, however, suggest that this is rather exceptional. Instead, the import and export receptors for various NLS or NES that exhibit intrinsic activities for interacting with the repeat sequences found in some nucleoporins mediate this task.[3-7] To date, at least 21 and 14 members of the importin β (or also called karyopherin β) family proteins have been identified in human and yeast *S. cerevisiae*, respectively, of which at least 9 of the yeast proteins and 5 of the human proteins, their functions in nuclear import have been assigned, whereas 4 of the yeast proteins and 3 of the human proteins have been identified as being involved in the nuclear

Nuclear Import and Export in Plants and Animals, edited by Tzvi Tzfira and Vitaly Citovsky.

export of various substrates.[8] According to the directions of transport, importin β family proteins are collectively termed importins (for nuclear import) and exportins (for nuclear export). Unidirectional transport by these shuttling receptors is basically maintained by a mechanism whereby substrate-importin complexes are destabilized, whereas substrate-exportin complexes are stabilized in the nucleus. The GTP-bound form of Ran, which can flip between the GDP and the GTP bound state as other small GTPases, plays a pivotal role in the regulation of complex formation. Ran-GTP is enriched in the nucleus due to the biased localization of a specific GTPase activating protein (GAP) and a guanine-nucleotide exchange factor (GEF). For the nuclear import of NLS-bearing substrates, NLS is recognized in the cytoplasm by importins and the complexes are then translocated through the NPC. Once transported into the nucleus, the binding of Ran-GTP to importins destabilizes the import complexes. In contrast, complex formation between NES-bearing substrates and exportins occurs only in the presence of Ran-GTP. In the cytoplasm Ran is rapidly converted from the GTP-bound to the GDP-bound state by the concerted action of RanGAP and RanBP1/BP2, and the cargo molecules are released (for a review see refs. 8 and 9). Thus, NLS or NES recognized by the importin β family proteins basically act only as signals for import or export. In addition to this Ran-dependent regulation, recent publications have established that various mechanisms for regulating the activities of NLS or NES by posttranslational modifications exist. Individual examples of such regulation mechanisms will be described later in this Chapter.

As described above, now we can define the criteria for NLS and NES. One of their functional characteristics is that they could be experimentally identified based on their abilities to stimulate the migration of heterologous reporter proteins, which are otherwise restricted to the cytoplasm or the nucleus, to the opposite compartment. Another important prognosis for NLS or NES, which may also be determined experimentally, is their abilities to be directly recognized by the well-characterized transport receptors in the absence (for NLS) or presence (for NES) of Ran-GTP. In this case, since it has been recently shown that a certain member of the importin β family (Kap142p/Msn5p) has the ability to transport different cargos in an opposite compartment through the NPC,[10] it should be noted that the binding ability of the sequence to importin β family molecules does not necessarily define the direction of transport. The third functional property is that the directionality of transport is exclusively one way. However, various signal sequences that do not fulfill these functional characteristics have emerged. For example a couple of signal sequences enable reporter proteins to be translocated both back and forth through NPC. We will also describe such exceptional sequences below.

Basic Type NLSs

Dingwall et al elegantly showed, using partially digested nucleoplasmin, that a signal required for the nuclear localization of nucleoplasmin exists in a "tail" portion of the molecule.[11] Thereafter, the first NLS was identified in Simian Virus 40 (SV40) large T antigen. This NLS, which is rich in lysine and arginine, consists of only seven amino acids, and has proved to be sufficient for the nuclear localization of the T antigen.[12] Moreover, synthetic peptides containing this sequence act as an NLS when chemically conjugated with a non-nuclear protein such as bovine serum albumin.[13,14] A mutational analysis revealed that lysine and arginine residues in the NLS are essential for nuclear targeting. Although a number of other NLSs had been identified, most have been found to contain one or two basic amino acid clusters composed of several lysine and arginine residues, therefore, these NLSs are referred to as basic type NLS. Thus, it was assumed that most nuclear localizing proteins have a basic type NLS in their primary amino acid sequence. Interestingly, there is no obvious consensus sequence between basic type NLSs. The key word is "basic amino acid cluster". The basic type NLS can be largely classified into two groups: (1) the single basic amino acid cluster type, composed of four to six lysine/arginine resides such as SV40 large T antigen and (2) the bipartite basic amino acid