The transport of integral membrane proteins across the nuclear pore complex

Anne C. Meinema,1 Bert Poolman1 and Liesbeth M. Veenhoff2,*
1Department of Biochemistry; Groningen Biomolecular Sciences and Biotechnology Institute; Netherlands Proteomics Centre; Zernike Institute for Advanced Materials; University of Groningen; Groningen, The Netherlands; 2Department of Neuroscience; University Medical Centre Groningen; University of Groningen; European Research Institute for the Biology of Ageing; Netherlands Proteomics Centre; The Netherlands

The nuclear envelope protects and organizes the genome. The nuclear pore complexes embedded in the nuclear envelope allow selective transport of macromolecules between the cytosol and nucleoplasm, and as such help to control the flow of information from DNA to RNA to proteins. A growing list of integral membrane proteins of the nuclear envelope are described to function in the organization of the genome, as well as the assembly of the NPCs. Here, we discuss how the nuclear pore complex may sort these proteins to obtain a specific protein composition of the inner membrane.

The Nuclear Envelope

The nuclear envelope (NE) consists of two phospholipid bilayers, called the inner and the outer nuclear membrane (INM, ONM), with the lumenal space in between (Fig. 1A). The INM faces the nucleoplasm and contains proteins that interact with the chromatin, and in metazoan nuclei also with the nuclear lamina that aligns with the INM.1,2 The ONM is continuous with the endoplasmic reticulum (ER) of the cell and has partially overlapping functions in the transport, synthesis and folding of proteins, and the synthesis of lipids.3 While the NE is a cisternal membrane structure, the ER forms cisternae (flattened membrane disks) as well as tubules (tube structures).4 In mammalian cells, much of the ER network lies around the nucleus, while in baker’s yeast and in higher eukaryotic cells during the interphase. Recent reviews have summarized all potential routes to the INM5,10 and here we will summarize only those observations and hypothesis that relate to a transport route of membrane proteins between the INM and ONM via the pore membrane across the NPC. The data are derived from different organisms and also using different GFP-fusions and therefore generalizations should be taken with caution.
Using rodent heterokaryons (cells with two nuclei) Powel and Burke initially showed that the INM-located integral membrane protein p55 was able to diffuse from one nucleus to the other via the ER. This experiment indicated that the ONM and INM are connected via the NPC and suggested that nuclear transport of membrane proteins is a diffusive process. Later, Soullam and Worman found that the N-terminal domain of LBR was essential for its targeting to the INM in simian fibroblast cells. Soluble cargo was known to be imported across the NPC by soluble transport factors that recognize the cargo molecules by binding to a nuclear localization signal (NLS), but the N-terminal domain of LBR lacked a known NLS. They reasoned that the proteins accumulated in the INM "by diffusion and ligand binding". Subsequently, it was shown that the lateral diffusion of GFP-labeled LBR was reduced in the INM compared with the diffusion in the ER, probably due to specific and functional interactions with nuclear components. This led to the diffusion-retention model: a membrane protein is able to diffuse across the NPC and is retained in the nucleus by intranuclear interactions. The transmembrane domain(s) remain in the pore membrane during the transport across the NPC, while the soluble domains pass through a space in between the NPC scaffold and the pore membrane, the lateral channels. These lateral channels were observed in electron microscopy (EM) studies. Indeed, the intra nuclear interactions are important for the localization of several membrane proteins in different organisms.

An upper size limit for the soluble domains of membrane proteins of 60–75 kDa was found, above which membrane proteins would not cross the NPC to reach the INM. As discussed by Hinshaw et al. and later by Soullam et al. the narrow lateral channel in the NPC is about 10 nm wide and this is large enough to provide access for cytosolic domains up to 60–75 kDa. Specifically in yeast, Deng and Hochstrasser determined that Doa10 (Ssm4), a yeast transmembrane ubiquitin ligase, accumulated in the INM by diffusion and retention. Doa10 could accumulate in the INM by diffusion across the NPC while having the soluble domain of an ER-localized protein Hrd1 (Der3), with a mass of 38 kDa plus one copy of the 28 kDa GFP. Replacing this Hrd1-domain for a larger globular protein Pglk (~45 kDa), the cytoplasmic domain increased from ~66 to ~73 kDa, and nuclear accumulation was attenuated. A systematic analysis of the size limitations in yeast would be valuable, but at least the data are consistent with the proposed upper size limit based on earlier studies in higher eukaryotes.
Mechanisms in Addition to Diffusion-retention

The first deviation from the simple diffusion-retention model was observed for the nuclear transport of a membrane reporter based on the transmembrane segment of Lap2β, which was dependent on metabolic energy; after ATP-depletion, the reporter did not accumulate at the INM but remained in the ER of HeLa cells. The reporter did not have a known NLS and so there was no reason to think that the nuclear transport was dependent on transport factors. A model was proposed, where “the NPC undergoes continuous, energy dependent restructuring,” which, “would create transient channels through the NPC at the nuclear pore membrane, thereby permitting lateral diffusional movement of integral proteins in the lipid bilayer between the INM and ONM.”

In the same year, another targeting mechanism for INM proteins was presented by Braunagel et al. They reported the finding of a sorting motif in the occlusion-derived virus (ODV) protein E66, which targets the membrane protein toward the NPC immediately after its synthesis. A cluster of positively charged residues at the cytoplasmic site, immediately next to the predicted transmembrane segment, was marked as the INM sorting motif. This cluster would be recognized by an isoform of the transport factor p10 in Spodoptera frugiperda, p10 is a homolog of the vertebrate transport factor importin α. They hypothesized that this isoform could target the protein through the ER to the NPC and even across the NPC.

In 2006, two proteins, Src1/Heh1 and Heh2, caught our attention in a proteomics study of yeast nuclear envelopes (unpublished). They had predicted LEM-domains indicating localization at the INM and they featured putative NLSs for the transport factors Kap60 (importin α in vertebrates). This raised the exciting possibility that these proteins could be imported by a mechanism involving transport factors. For soluble cargo, an import reaction mediated by the transport factors Kap60 and Kap95 (importin α and importin β in vertebrates) involves binding of Kap60 to the NLS on a cargo. The complex of cargo/Kap60/Kap95 is shuttled across the NPC by interactions of Kap95 with a specific subset of NPC proteins, named FG-Nups, which encode repeats that are rich in phenylalanine and glycine, and serve as Kap binding sites. GTP-bound Ran in the nucleus and GDP-bound Ran in the cytoplasm coordinate the direction of transport; in the case of Kap60/95 mediated import, RanGTP stimulates the release of the cargo from the transport factors in the nucleus. Solid evidence, that NPC components, the transport factors Kap60 and Kap95 and RanGTP were needed for the INM-localization of these yeast proteins was published that same year by King, Lusk and Blobel. This was the first publication clearly showing the parallels between mechanisms responsible for targeting of membrane and soluble proteins.

Three other membrane proteins featuring an NLS were later found: the human Pom121, which is a component of the NPC and essential for NPC-assembly, and two members of the SUN family, the human SUN2 and the C. elegans Unc-84. However, it has been rightfully pointed out that the prevalence of these sequences within INM proteins, which are rich of positively charged residues, could be a contribution to their chromatin binding function. These putative chromatin-binding motifs could support the capture of INM proteins and their retention at the INM. Further complications arrived as it was actually shown that location of Unc-84 in the INM of C. elegans cells requires multiple targeting signals, namely two NLSs plus an INM sorting motif and a SUN-nuclear envelope localization signal. SUN2 targeting is also complicated, as in addition to a classical NLS, two other elements are needed for its proper INM localization. One element is an arginine cluster that conveys INM accumulation to the linker sequence is not important as randomized versions also work. However, the linker length is important, and a clear relationship between linker length and INM accumulation was observed. In the constructs with the randomized linker and the synthetic transmembrane segment, the only original Heh2 sequence that was still present, was the h2NLS, and we thus concluded that no other signals are required. It is unclear what the importance of the reported INM-sorting signal in Heh2 may be for INM sorting.

Kap60/95 Dependent Transport of GFP-reporters

We have studied the INM targeting of Heh2 and Heh1 in more detail and showed that they feature a high affinity NLS for Kap60 binding, named h2NLS and h1NLS respectively. The NLSs of Heh1 and Heh2 are spaced from the transmembrane domain (TM) by a 180 or 235 amino acid long intrinsically disordered linker (L). The h2NLS together with the long linker is a transplantable signal that conveys INM accumulation to a synthetic transmembrane segment and also to the normally ER-localized Sec61. The linker sequence is not important as randomized versions also work. However, the linker length is important, and a clear relationship between linker length and INM accumulation was observed. In the constructs with the randomized linker and the synthetic transmembrane segment, the only original Heh2 sequence that was still present, was the h2NLS, and we thus concluded that no other signals are required.
Figure 2. Proposed mechanism for Kap60/95 dependent import of Heh1 and Heh2. The structure of the scaffold of nuclear pore complex is based on reference 79 and the FG-Nups are an artist’s impression. The GFP-h2NLS-L-TM reporter is indicated in yellow. Kap60 and Kap95 (red and pink) bind the high affinity h2NLS of the membrane reporter. The intrinsically-disordered linker facilitates the interactions between Kap60/95 and the FG-Nups in the central channel of the NPC by dodging in the NPC scaffold while the transmembrane domains diffuse through the POM. Figure created by Graham Johnson of grahamj.com.

We then showed that INM accumulation results from Kap60/95 mediated import and Kap-independent leak. The reporters are mobile within the NE/ER network including the INM and there is no selective retention. From experiments where we trapped a reporter protein at the anchor domain of an FG-Nup, Nsp1, we concluded that the soluble domains with the h2NLS traverse the NPC central channel. Indeed also large (174 kDa) soluble domains linked terminal to the h2NLS were tolerated. We proposed a model for the transport mechanism of these membrane proteins where the Kap60/95 bound h2NLS would pass through the central channel interacting with the FG-Nups while the transmembrane segments diffuse through the pore membrane (Fig. 2). A sufficiently long linker domain would allow these interactions with the FG-Nups by dodging into the NPC scaffold. It has been shown that intrinsically-disordered domains can easily adopt a wide range of lengths and it will not cost much energy to stretch such a domain. Stretching the linker may be energetically favored by the enthalpy gain from binding of NLS-bound Kap60/95 to the multiple FG-repeat binding sites in the central channel of the NPC. This proposed transport mechanism of the Heh1 and Heh2-derived proteins resembles that of soluble proteins: it requires RanGTP, Kaps and FG-Nups, and the NLS-encoding domain passes through the central channel. A number of observations lead us to think that it is very unlikely that the reporters are transported though the NPC as soluble proteins. While polytopic membrane
proteins are generally co-translationally integrated in the ER via the signal receptor particle (SRP) and Sec61 translocation pathway (reviewed in ref. 50–52), tail-anchored membrane proteins use post-translational membrane insertion mechanism (reviewed in ref. 51, 53 and 54). To rule out the possibility that the membrane reporters are translocated across the NPC as soluble proteins, we studied the nuclear transport of a number of different reporter proteins. Most significant are our experiments with the Sec61-fusions that are certainly co-translationally inserted in the ER membrane: these reporters also accumulate at the INM in NLS- and linker-dependent fashion. Interesting would be to extend the studies with the Sec61 reporters to versions with large extraluminal domains, also to make the comparison with size restrictions found for Sun2.65 In addition, there are marked differences with soluble transport. First, trapping the reporter at the NPC protein, Nsp1, within the NPC reduces nuclear transport of membrane proteins but not that of soluble proteins. Also, the translocation of h2NLS-L-TM through the NPC is disrupted in a strain that lacks the NPC protein Nup170,66 similar as Heh2,35 while the transport of soluble cargo h2NLS-GFP is not affected (Fig. 3; ref. 55). Third, the Kap60/95-facilitated nuclear transport of the h2NLS-L-TM reporters and soluble GFP-cNLS reporters is affected differentially in the different FG-Nup mutants. Lastly, the accumulation at the INM of the reporters is reversible and molecules leak out when the RanGTP gradient is disrupted of when Kap60/95 is no longer available. If transport across the NPC would be as a soluble cargo, then reversible association of the reporters with the membrane would be required. Such a process is unlikely, particularly as it would have to be independent of the amino acid sequence of the transmembrane or linker domains. Altogether, we concluded that the most straightforward interpretation of the data are that the transmembrane segments of the reporter remain in the pore membrane while the soluble domains pass through the FG-repeat network of the NPC.

**Figure 3.** Transport of h2NLS-L-TM depends on Nup170 while transport of h2NLS-GFP does not. Confocal fluorescence images in cells where Nup170 is knocked out (nup170Δ) show mislocalization of GFP-h2NLS-L-TM but the soluble and mobile GFP-h2NLS-L is accumulated in the nucleus. The scale bar is 5 μm.

**Implications for the Structure of the NPC**

Can we reconcile our transport model with current data describing the NPC structure? The structure and architecture of the NPC have been investigated intensively with EM.20,21,56–60 In these EM-studies, the nuclear pore complex is shown as a spoke-ring complex with an 8-fold rotational symmetry along the nucleocytoplasmic axis through the central channel and a 2-fold symmetry in the plane of the membrane. In both the vertebrate and yeast NPC, the protein density is lower between the spokes, but lateral gates that connect the space close to the pore membrane and the central channel, and could host the linker region, have not been described in EM-studies. However, even the most detailed images of the NPC obtained by 3D EM-tomography may have too limited resolution to observe such features.20,57 EM-studies do reveal a high structural plasticity within the NPC. The NPCs can differ in their radial size and occasionally NPCs with more than eight spokes have been observed.61–63 Differences in the sizes of single spokes within one NPC were observed and could be explained by spoke compaction and extension or changing distances between spokes.20,57 Dilation of the complex might be needed to facilitate the transport of large particles across the NPC.64,65 Altogether, these studies show that the NPC is not static or rigid, but has radial flexibility, resulting from dynamics between the spokes.

From many biochemical and structural studies we know which proteins make up the NPC scaffold and with variable detail and certainty we know where they are within the NPC (reviewed in ref. 65–67). In baker’s yeast, only the transmembrane NPC protein Pom152, forms a continuous ring by homotypic interactions of the luminal domains,68 but these will not affect the movement of the linker. Other biochemically stable sub-complexes can be isolated or assembled from purified proteins, but none of them have a tendency to form stable oligomers, which could have been indicative for stable spoke-to-spoke interactions. This hold true for the Nup84 subcomplex from yeast,69–71 complexes of Nup170/157 and Nup192/188 from the eukaryotic thermophile Chaetomium thermophilum72 and the yeast membrane anchoring complex of Ndc1, Pom152 and Pom34.73 The stability of these complexes in solution may indicate that within the NPC they are also stable as independent units. In contrast, within a spoke there is evidence for direct protein-protein contacts that range from the pore membrane to the center of the NPC.74

Collectively, we judge that the available data on the structure of the NPC does not contradict the proposed lateral gates that connect the space close to the pore membrane and the central channel. It is mostly some cartoons and descriptions of the NPC that have suggested a scaffold structure built of continuous rings.
expression and genome stability have been described for Heh1. 74,76,77 Also relevant to note is that Heh1 and Heh2 contribute at discrete and different steps to the initiation of the assembly of new NPCs.78 The strong accumulation at the INM might be essential to determine the orientation of the NPC in the NE, e.g., Heh1 and Heh2 could mark the INM and assure initiation of NPC formation in the right orientation. The Heh1 and Heh2 membrane proteins may need to be mobile to define these new sites for pore formation. The reversible accumulation of mobile proteins may be a third advantage over a retention mechanism. Lastly, the presence of the linker may alleviate a restriction on the size of the extralumenal domains to be transported, at least if terminal to the linker. Future studies are needed so answers outstanding questions such as: where does the linker pass the NPC scaffold and do the transmembrane segments indeed stay in the pore membrane? What properties of the linker, next to its length, are important: flexibility, charge? What defines the high affinity interaction with Kap60?

**Outlook**

Although there are no other INM-localized proteins described that are actively targeted to the INM, there are a number of proteins, known to reside at the INM, that have confirmed Kap60 binding sites important for the INM-localization. The yeast protein Nur1 was identified at the INM and has a confirmed NLS.35,74 Human Pom121 and Sun2 have a bipartite NLS, interact with importinβ via importin α and their INM-localization was dependent on RanGTP 37 and the localization of the *Caenorhabditis elegans* Sun protein Unc-84 was partly dependent on the presence of two stretches of cNLS segments.41 The three INM-localized proteins Pom121, Sun2 and Unc-84 in metazoans share no identity at the amino acid level with the yeast Heh1, Heh2 or Nur1, but may share the ‘NLS-L-TM’-feature (Fig. 4). The region between the NLS and the first transmembrane helix is predicted to be largely unfolded in these proteins, and particularly for Nur1, Pom152 and Unc84, and maybe less so for Sun2, the linker domains are long enough for a nuclear targeting mechanism alike Heh2 and Heh1 (> 120 residues). Based on these considerations, we infer that these mammalian membrane proteins may be targeted to the INM according to the model we have presented for the transport of Heh1 and Heh2 (Fig. 2).

Heh1 and Heh2 are so far the sole proteins confirmed to feature the “high affinity NLS-linker-TM-domain” sorting signal that is sufficient to give high accumulation at the INM using Kap-facilitated import. This mechanism potentially has three advantages over selective retention. The first is that it potentially allows for fast changes of INM levels of Heh1 and Heh2 e.g., by activation of import through exposure of the h2NLS. There are however no indications that this occurs. The second would be, that the INM levels are solely dependent on the import machinery. Heh1 and Heh2 could thus initiate recruitment of specific INM proteins and serve a role in defining the environment of the INM. Indeed, functions in chromatin organization, regulation of gene expression and genome stability have been described for Heh1.74,76,77 Also relevant to note is that Heh1 and Heh2 contribute at discrete and different steps to the initiation of the assembly of new NPCs.74 The strong accumulation at the INM might be essential to determine the orientation of the NPC in the NE, e.g., Heh1 and Heh2 could mark the INM and assure initiation of NPC formation in the right orientation. The Heh1 and Heh2 membrane proteins may need to be mobile to define these new sites for pore formation. The reversible accumulation of mobile proteins may be a third advantage over a retention mechanism. Lastly, the presence of the linker may alleviate a restriction on the size of the extralumenal domains to be transported, at least if terminal to the linker. Future studies are needed so answers outstanding questions such as: where does the linker pass the NPC scaffold and do the transmembrane segments indeed stay in the pore membrane? What properties of the linker, next to its length, are important: flexibility, charge? What defines the high affinity interaction with Kap60?

**Figure 4.** Eukaryotic proteins with putative NLS-L-TM configuration. The NLS-LTM configurations are indicated for Heh1 (scHeh1) and Heh2 (scHeh2) and predicted for yeast Nur1 (scNur1), human Pom121 (hsPom121), human Sun2 (hsSun2) and the *Caenorhabditis elegans* Unc-84 (ceUnc-84). The transmembrane segments are predicted by the TMHMM server v.2.19 the NLSs are predicted by PSORT II and experimentally determined.35,37,38,41,42,45,74 The length in amino acids of the region spacing the predicted NLS and transmembrane segments is indicated. Indicated between brackets is the number of disordered residues as predicted according to FoldIndex.75

<table>
<thead>
<tr>
<th>Protein</th>
<th>NLS</th>
<th>Putative linker</th>
</tr>
</thead>
<tbody>
<tr>
<td>scHeh1</td>
<td>173-RKKKRK 214-KAKKKRK</td>
<td>235 (198)</td>
</tr>
<tr>
<td>scHeh2</td>
<td>102-KKKRK 124-PKPKKKR KRSSKANK</td>
<td>180 (176)</td>
</tr>
<tr>
<td>scNur1</td>
<td>475-PKPKKK NYHKRR</td>
<td>197 (132)</td>
</tr>
<tr>
<td>hsPom121</td>
<td>321-KKR 339-KRRR 387-KRSR 453-KRIR 504-RKRK</td>
<td>260 (169)</td>
</tr>
<tr>
<td>hsSun2</td>
<td>41-RTLKRK 51-RK</td>
<td>160 (86)</td>
</tr>
<tr>
<td>ceUnc-84</td>
<td>35-KVRRK 170-HRRR</td>
<td>209 (165)</td>
</tr>
</tbody>
</table>
Which proteins feature an h2NLS-L-TM, and is the proposed mechanism for INM accumulation of membrane proteins confined to yeast or also present in humans?

Acknowledgements

We would like to thank P. Lusk, J.K. Laba, R.A. Hapsari, A. Kralt, P. Popken, G.E. Janssens and U. Kutay for discussions and insightful comments. This work was supported by funding from the Netherlands Organization for Scientific Research (VIDI fellowship to L.M.V.; Top-subsidy grant 700.56.302 to B.P.).

References


16. Soullias C, Worman HJ. The amino-terminal domain of the lamin B receptor is a nuclear envelope targeting signal. J Cell Biol 1993; 120:1093-100; PMID:7679672; http://dx.doi.org/10.1083/jcb.120.4.1093.


