Review Letter

Active transport of proteins into the nucleus

Philipp Wagner, Jeannette Kunz, Antonius Koller and Michael N. Hall

Department of Biochemistry, Biocenter of the University of Basel, CH-4056 Basel, Switzerland

Received 4 September 1990

Nuclear proteins are actively and posttranslationally transported across the nuclear envelope. This transport is a highly selective process that can be divided into two steps, receptor-binding followed by translocation through the nuclear envelope. Receptor-binding is mediated by nuclear localization signals that have been identified in many nuclear proteins. Translocation is energy-dependent and occurs through the nuclear pore complex.

Nuclear localization signal; Nuclear pore complex

1. INTRODUCTION

The nuclear envelope sequesters the genome and its activities within a unique biochemical environment, the nucleus. The nuclear envelope consists of two lipid bilayers, the outer and inner nuclear membranes, separated by a perinuclear cisternal space. The perinuclear space is continuous with the lumen of the endoplasmic reticulum (ER). The outer nuclear membrane and the ER membrane are also continuous and functionally similar in that both contain ribosomes on their cytoplasmic surfaces. The nucleoplasmic surface of the inner membrane is associated with the nuclear lamina, a fibrous network that supports the nuclear envelope.

Pores traverse the nuclear envelope at sites where the inner and outer membranes are fused, thereby providing a link between the cytoplasm and the interior of the nucleus. Each nuclear pore is a water-filled channel within a large proteinaceous nuclear pore complex (NPC). The NPC mediates active transport of proteins across the nuclear envelope, and, in addition, allows passive diffusion of small proteins and other molecules. Here we review the salient features of active protein transport into the nucleus. The reader is referred to other more extensive reviews for additional information on passive [1] and active [2–6] transport and on the structure of the nucleus [7–12].

2. NUCLEAR PORE COMPLEX (NPC)

In 1876, using the light microscope as a means of observation, Hertwig first suggested the existence of nuclear pores. Observing a punctate pattern on nuclear envelopes, he proposed the existence of channels connecting the cytoplasm and nucleoplasm. In 1949, Callan [13] observed the ultrastructure of nuclear pores with the electron microscope. Wischnitzer [14] proposed the first structural r el of the nuclear pore in 1958. Recently, more detailed information on the structure of the NPC has been obtained using high resolution electron microscopy combined with image processing [15–17]. About an axis perpendicular to the plane of the nuclear envelope, the NPC has an 8-fold rotational symmetry formed by two coaxial rings connected to a central plug by a spoke assembly. One ring lies on each side of the nuclear envelope, facing the cytoplasm or the nucleoplasm. The cytoplasmic ring is heavier than the nucleoplasmic ring, creating asymmetry along the axis parallel to the nuclear envelope. The spoke assembly probably contains 24 spokes; 8 inner spokes extending outward from the central plug (inner spokes) joining 8 outer spokes extending inward from each of the two rings. The central plug or transporter is thought to contain the transport channel. Occasionally, fibrils can be seen extending from the pore into either the cytoplasm or the nucleoplasm, possibly acting as tracks through the pore. The overall size of the NPC is 120 nm in diameter by 70 nm. Radial arms extending out from the NPC, possibly to anchor the two coaxial rings to the membrane, have also been proposed. With these radial arms the NPC would have an overall diameter of 145 nm.

Mass determination of the NPC has yielded a molecular mass of ~124 MDa [17]. Assuming that an average pore protein has a molecular mass of 200 kDa, approximately 500 proteins (possibly over a 100 different protein species) are required to assemble an NPC. To date, relatively few NPC proteins have been

Correspondence address: M.N. Hall, Department of Biochemistry, Biocenter of the University of Basel, CH-4056 Basel, Switzerland

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

identified. One, gp210 (or gp190), is an integral membrane glycoprotein containing *N*-linked high mannose-type carbohydrate and has been proposed to anchor the pore complex within the nuclear envelope [18–20]. Among the handful of other proteins thought to be part of the NPC [21–27] many bear an unusual carbohydrate modification consisting of single *O*-linked *N*-acetylglucosamine (GlcNAc) residues ([28–30]; for review see [12]).

3. NUCLEAR LOCALIZATION SIGNAL (NLS)

Upon injecting BSA and histones into *Xenopus* oocytes, Gurdon [31] found that histones, but not BSA, accumulate in the nucleus. He concluded that protein uptake into nuclei is a selective process. In 1975, Bonner [32] injected either nuclear contents or cytoplasm of *Xenopus* oocytes into recipient oocytes and found that nuclear, but not cytoplasmic, proteins accumulate in the nucleus; however, he also found that cytoplasmic proteins smaller than 20 kDa equilibrate between the nucleus and cytoplasm within 24 h, suggesting that small proteins can freely diffuse into the nucleus. In a variation of Gurdon's original experiment, De Robertis et al. [33] suggested in 1978 that nuclear proteins contain in their molecular structure a signal that enables them to accumulate in the nucleus. The first direct evidence for such a signal came in 1982 from experiments employing nucleoplasmin as a nuclear import substrate. Dingwall et al. [34] found that nucleoplasmin rapidly accumulates in the nucleus when injected into oocytes, but nucleoplasmin derivatives lacking the C-terminal end are not transported to the nucleus. This work demonstrated that sequences present within a mature protein are sufficient to direct nuclear localization. The first nuclear targeting signal whose sequence was defined is from the yeast nuclear protein MATα2. By constructing hybrid proteins containing β-galactosidase fused to different portions of MATα2, Hall et al.[35] showed that a segment of MATα2 as small as 13 amino acids can direct β-galactosidase to the nucleus. Within these 13 amino acids is a possible consensus signal sequence consisting of 5 animo acids, Lys^3-Ile-Pro-Ile-Lys [35,36]. Independently and shortly thereafter, Kalderon et al. [37,38] and Lanford and Butel [39] described an NLS in SV40 large T antigen. This sequence, Pro-Lys-Lys^{128}-Lys-Arg-Lys-Val, is sufficient to target linked non-nuclear proteins to the nucleus [38]. Mutant large T antigen, in which Lys^{128} is replaced by threonine or asparagine, does not accumulate in the nucleus [37,39]. NLSs have now been identified in a large number of nuclear proteins (for a comprehensive list see [6]).

In 1986, Richardson et al. [40] found that the polyoma virus large T antigen contains two independent nuclear signal sequences. When both signals are deleted, the protein is no longer directed to the nucleus, whereas deletion of only one of the two signals confers a much less dramatic localization defect. Both sequences resemble the nuclear targeting signal of SV40 large T antigen, containing basic amino acids and a proline residue. Subsequent studies in which two NLSs have been found in several nuclear proteins (see [6]) suggest that two signals could be a general feature of nuclear protein import. The role for a dual signal is not yet clear. The two signals could be functionally equivalent and act additively or could be functionally distinct (for review see [6]).

There is no single consensus among the many NLSs that have been identified to date. However, there are some general descriptive rules. Nuclear localization signals (1) are typically short sequences, usually not more than 8 to 10 amino acids; (2) contain a high proportion of positively charged amino acids (lysine and arginine) often associated with a proline; (3) can reside in any exposed region of a nuclear protein; (4) are not removed following localization; and (5) can occur more than once in a given protein.

4. MECHANISM OF TRANSPORT INTO THE NUCLEUS

The molecular machinery that translocates proteins across the nuclear envelope is part of the nuclear pore complex (NPC), as incisively demonstrated in 1984 by Feldherr et al.[41]. Feldherr et al. examining by electron microscopy the nuclear accumulation of colloidal gold particles coated with the nuclear protein nucleoplasmin, observed that at early times after injection into the cytoplasm, the particles accumulated at nuclear pores. They concluded that the pores are the major, if not the exclusive, sites for transport; there is no evidence supporting a model in which exchange occurs directly across the nuclear membranes. Additional evidence for transport through the NPC has subsequently come from numerous studies in which NPC-binding agents, such as antibodies [42,43] and WGA [44–48], are shown to inhibit nuclear import.

How are proteins transported across the nuclear envelope? Although this question is not yet solved, the process is known to consist of at least two steps [36,49,50]: (1) binding to a receptor either on the nuclear envelope or in the cytoplasm; and (2) a subsequent energy-dependent translocation through the nuclear pore complex.

4.1. *Step one: receptor binding*

A large number of studies demonstrating the saturability and specificity of nuclear protein import argue that the import process is receptor-mediated [51]. For example, Goldfarb et al. [52] showed that BSA conjugated to the NLS of SV40 large T antigen is imported into nuclei with saturable kinetics (approximate K_m of 1.8 μM and a V_{max} of 6.4 $\times$ 10^9 molecules/cell per

min). By coinjecting the BSA-large T NLS conjugate with free NLS peptide, uptake of the conjugate could be inhibited. In contrast, a peptide containing an import deficient signal did not inhibit import. These data suggest that nuclear import is mediated by a receptor that interacts with the NLS.

Where is the receptor? There are at least three possibilities which are currently being considered. First, the receptor is located at the NPC where it binds and transports protein into the nucleus. Second, the receptor resides in the cytoplasm where it binds a nuclear protein, shuttles it to the NPC, and then releases it to be transported into the nucleus by a separate machinery. Third, the receptor is located in the cytoplasm, nucleoplasm and nuclear envelope such that it binds a nuclear protein in the cytoplasm; a receptor-nuclear protein complex then binds the NPC and is transported into the nucleus, where the nuclear protein is released and the receptor is recycled.

Evidence for receptors near or in the NPC is provided by studies in which presumed import intermediates are visualized directly [36,49,50,53,54]. Employing colloidal gold particles coated with nuclear proteins, it has been shown that under ATP-depleted conditions, nuclear proteins bind to the NPC, suggesting that the receptor is located on the NPC or on the fibrils which emanate from the NPC. Additional evidence for a NLS-binding receptor in the NPC has come from a study by Yoneda et al. [43]. Reasoning that a receptor for a positively charged signal such as the large T antigen NLS (Lys-Lys-Lys-Arg-Lys) might have a complementary stretch of negatively charged amino acids, Yoneda et al. raised antibodies against a peptide containing the sequence Asp-Asp-Asp-Glu-Asp. The antibodies both recognize antigens in the NPC and inhibit nuclear import, suggesting that an antigen in the NPC is a receptor for the large T antigen NLS.

Four groups have described nuclear (presumably exclusively) NLS-binding proteins. Lee and Melese [55] and Silver et al. [56], identified yeast NLS-binding proteins in the nucleus by ligand blotting. Silver et al. found two proteins of 70 and 59 kDa. The biochemical behavior of these proteins suggests that they associate with nuclei via protein-protein interactions. Lee and Melese identified a 67 kDa protein that is tightly associated with the nuclear envelope, either with the NPC or the putative yeast nuclear lamina. The 70 kDa protein found by Silver et al. and the 67 kDa found by Lee and Melese could be the same protein. Benditt et al. [57] identified 56, 57, 65 and 74 kDa signal-binding proteins from purified rat nuclear envelopes by crosslinking. Li and Thomas [58], also by crosslinking, have described a 66 kDa nucleoplasmic NLS-binding protein from human cells.

A receptor found exclusively at the NPC is difficult to reconcile with the lack of a consensus NLS. One might imagine that many different types of signals would require a large number of receptors on a limited number of pore complexes. This problem becomes particularly acute when considering that all pores appear to be competent to import a given protein [41]. One solution to this problem is to invoke cytoplasmic receptors that serve as adaptors between the different types of signals and a common binding site on the nucleus. Evidence for non-nuclear receptors has been reported by several groups. Newmeyer and Forbes [59] have described a frog cytoplasmic factor, NIF-1, that stimulates nuclear protein import in vitro. Breeuwer and Goldfarb [60] found that histone H1 is retained in the cytoplasm at 4°C, whereas non-nuclear proteins of similar size diffuse into the nucleus; the cytoplasmic retention of H1 is overcome by injection of excess H1. This saturable cytoplasmic retention supports the notion that H1 is localized to the nucleus by a cytoplasmic receptor that prevents diffusion at low temperature. Yamasaki et al. [61] have described proteins from rat liver cells that bind an NLS peptide; two of these proteins, 100 and 70 kDa, are cytoplasmic, while the remaining two proteins, 140 and 55 kDa, are loosely nucleus-associated. These proteins could bind NLS-bearing proteins in the cytoplasm and shuttle them to a receptor residing at the NPC. Finally, Adam et al. [62] have also found in rat liver cell extract two proteins that bind an NLS-bearing protein. These proteins, of 60 and 70 kDa, reside in the cytosol, nuclear envelope and nucleoplasm, supporting the multistep model for carrier-mediated nuclear protein import in which a recycling receptor carries a passenger protein all the way from the cytoplasm to the nucleoplasm. Which of the many NLS-binding proteins described to date are physiologically relevant to the import process and the validity of any of the three models on receptor function described above remain to be determined.

4.2. *Step two: translocation through the nuclear pore complex*

The second step in the transport of nuclear proteins from the cytoplasm to the nucleoplasm is translocation through the nuclear pore. This step is clearly the active step in the import process, requiring both physiological temperature and ATP. At low temperature or in the absence of ATP, import substrates bind to the nuclear periphery but are not translocated across the nuclear envelope [49,50]. The requirement for physiological temperature has been demonstrated in vivo [47,50,63] and in vitro [48,64-69]. The ATP requirement has also been investigated in a number of systems. Nuclear protein import is abolished following depletion of ATP with apyrase in cell-free systems [49,64,65,67–70] or with deoxyglucose and sodium azide in cultured cells [50]. Furthermore, a non-hydrolyzable ATP analogue does not support import [68].

The actual translocation step itself might consist of

several steps, as revealed by EM visualization of NLS-coated colloidal gold particles bound to nuclear pore complexes [53,54] and of transport-related configurations of the pore transporter [71]. A nuclear protein appears first to bind to the periphery of the transporter and then to move through an opening in the center of the transporter. The transport channel perhaps opens concomitantly with the binding and movement of an import substrate such that only certain proteins are allowed to pass through the channel. Remarkably, nucleoplasmin-coated gold particles with a diameter as large as 20 nm can pass through the nuclear pore [41], even though the pore aqueous channel has a diameter of only 9 nm [72]. Since the pore can accomodate large, rigid gold particles, protein deformation or unfolding must not be a prerequisite for nuclear import. The above observations taken together suggest that the pore may work as an ATP-driven iris diaphragm that opens in response to a signal recognition event.

5. CONCLUSION

The study of nuclear protein localization is now entering its third phase. The first phase, the early years, consisted mostly of morphological and physiological studies on the nuclear envelope, and was responsible primarily for establishing the nuclear pore complex as a channel through which molecules can diffuse between the cytoplasm and the nucleoplasm. The second phase was ushered in during the early 1980s by the discovery of nuclear localization signals and the realization that nuclear import involves active transport in addition to passive diffusion. The third phase is now beginning to identify the cellular components with which NLSs interact and should eventually elucidate the actual molecular mechanism by which proteins are actively transported into the nucleus.

Acknowledgements: We thank Joe Heitman and Jose Garcia-Bustos for comments on the manuscript. The laboratory, in which some of the experiments reviewed here were performed, was supported by a grant from the Swiss National Science Foundation to M.N.H.

REFERENCES

[1] Bonner, W.M. (1978) in: The Cell Nucleus, (Busch, H. ed.) pp. 97-148, Academic Press, New York.
[2] Dingwall, C. and Laskey, R.A. (1986) Annu. Rev. Cell Biol. 2, 367-390.
[3] Peters, R. (1986) Biochim. Biophys. Acta 864, 305-359.
[4] Goldfarb, D.S. (1989) Curr. Op. Cell Biol. 1, 441-446.
[5] Roberts, B. (1989) Biochim. Biophys. Acta 1008, 263-280.
[6] Garcia-Bustos, J.F., Heitman, J. and Hall, M.N. (1991) Biochim. Biophys. Acta (in press).
[7] Maul, G.G. (1977) in: International Review of Cytology, (Bourne, G.H., Danielli, J.F. and Jeon, K.W. eds) pp.. 75-186, Academic Press, New York.
[8] Harris, J.R. (1978) Biochim. Biophys. Acta 515, 55-104.
[9] Franke, W.W., Scheer, U., Krohne, G. and Jarasch, E.-D. (1981) J. Cell Biol. 91, 39s-50s.
[10] Newport, J.W. and Forbes, D.J. (1987) Annu. Rev. Biochem. 56, 535-565.
[11] Gerace, L. and Burke, B. (1988) Annu. Rev. Cell Biol. 4, 335-374.
[12] Hart, G.W., Haltiwanger, R.S., Holt, G.D. and Kelly, W.G. (1989) Annu. Rev. Biochem. 58, 841-874.
[13] Callan, H.G., Randall, J.R. and Tomlin, S.G. (1949) Nature 163, 280-283.
[14] Wischnitzer, S. (1958) J. Ultrastruc. Res. 1, 201-222.
[15] Unwin, P.N.T. and Milligan, R.A. (1982) J. Cell Biol. 93, 63-75.
[16] Akey, C.W. (1989) J. Cell Biol. 109, 955-970.
[17] Reichelt, R., Holzenburg, A., Buhle, E.L., Jarnik, M., Engel, A. and Aebi, U. (1990) J. Cell Biol. 110, 883-894.
[18] Gerace, L., Ottaviano, Y. and Kondor-Koch, C. (1982) J. Cell Biol. 95, 826-837.
[19] Wozniak, R.W., Bartnik, E. and Blobel, G. (1989) J. Cell Biol. 108, 2083-2092.
[20] Greber, U.F., Senior, A. and Gerace, L. (1990) EMBO J. 9, 1495-1502.
[21] Davis, L.I. and Blobel, G. (1986) Cell 45, 699-709.
[22] Park, M.K., D'Onofrio, M., Willingham, M.C. and Hanover, J.A. (1987) Proc. Natl. Acad. Sci. USA 84, 6462-6466.
[23] Snow, C.M., Senior, A. and Gerace, L. (1987) J. Cell Biol. 104, 1143-1157.
[24] Finlay, D.R. and Forbes, D.J. (1990) Cell 60, 17-29.
[25] Davis, L.I. and Fink, G. (1990) Cell 61, 965-978.
[26] Nehrbass, U., Kern, H., Mutvei, A., Horstman, H., Marshallsay, B. and Hurt, E.C. (1990) Cell 61, 979-989.
[27] Starr, C.M., D'Onofrio, M., Park, M.K. and Hanover, J.A. (1990) J. Cell. Biol. 110, 1861-1871.
[28] Davis, L.I. and Blobel, G. (1987) Proc. Natl. Acad. Sci. USA 84, 7552-7556.
[29] Hanover, J.A., Cohen, C.K., Willingham, M.C. and Park, M.K. (1987) J. Biol. Chem. 262, 9887-9894.
[30] Holt, G.D., Snow, C.M., Senior, A., Haltiwanger, R.S., Gerace, L. and Hart, G.W. (1987) J. Cell Biol. 104, 1157-1164.
[31] Gurdon, J.B. (1970) Proc. R. Soc. B 176, 303-314.
[32] Bonner, W.M. (1975) J. Cell Biol. 64, 431-437.
[33] De Robertis, E.M., Longthorne, R.F. and Gurdon, J.B. (1978) Nature 272, 254-256.
[34] Dingwall, C., Sharnick, S.V. and Laskey, R.A. (1982) Cell 30, 449-458.
[35] Hall, M.N., Hereford, L. and Herskowitz, I. (1984) Cell 36, 1057-1065.
[36] Hall, M.N., Craik, C. and Hiraoka, Y. (1990) Proc. Natl. Acad. Sci. USA 87, 6954-6958.
[37] Kalderon, D., Richardson, W.D., Markham, A.F. and Smith, A.E. (1984) Nature 311, 33-38.
[38] Kalderon, D., Roberts, B.L., Richardson, W.D. and Smith, A.E. (1984) Cell 39, 499-509.
[39] Lanford, R.E. and Butel, J.S. (1984) Cell 37, 801-813.
[40] Richardson, W.D., Roberts, B.L. and Smith, A.E. (1986) Cell 44, 77-85.
[41] Feldherr, C.M., Kallenbach, E. and Shultz, N. (1984) J. Cell Biol. 99, 2216-2222.
[42] Featherstone, C., Darby, M.K. and Gerace, L. (1988) J. Cell. Biol. 107, 1289-1297.
[43] Yoneda,Y., Imamoto-Sonobe, N., Matsuoka, Y., Iwamoto, R., Kiho, Y. and Uchida, T. (1988) Science 242, 275-278.
[44] Finlay, D.R., Newmeyer, D.D., Price, T.M. and Forbes, D.J. (1987) J. Cell Biol. 104, 189-200.
[45] Yoneda, Y., Imamoto-Sonobe, N., Yamaizumi, M. and Uchida, T. (1987) Exp. Cell Res. 173, 586-595.
[46] Dabauvalle, M.-C., Shulz, B., Scheer, U. and Peters, R. (1988) Exp. Cell Res. 174, 191-196.

[47] Wolff, B., Willingham, M.C. and Hanover, J.A. (1988) Exp. Cell Res. 178, 318–334.
[48] Shirakawa, F. and Mizel, S.B. (1989) Mol. Cell. Biol. 9, 2424–2430.
[49] Newmeyer, D.D. and Forbes, D.J. (1988) Cell 52, 641–653.
[50] Richardson, W.D., Mills, A.D., Dilworth, S.M., Laskey, R.A. and Dingwall, C. (1988) Cell 52, 655–664.
[51] Hall, M.N. and Fried, S.R. (1988) in: Molecular Biology of Intracellular Protein Sorting and Organelle Assembly, (Bradshaw, R.A., McAlister-Henn, L. and Douglas, M.G. eds) pp. 187–192, Alan R. Liss, New York.
[52] Goldfarb, D.S., Gariepty, J., Schoolnik, G. and Kornberg, R.D. (1986) Nature 322, 641–644.
[53] Akey, C.W. and Goldfarb, D.S. (1989) J. Cell Biol. 109, 971–998.
[54] Stewart, M., Whytock, S. and Mills, A.D. (1990) J. Mol. Biol. 213, 575–582.
[55] Lee, W.-C. and Melese, T. (1989) Proc. Natl. Acad. Sci. USA 86, 8808–8812.
[56] Silver, P., Sadler, I. and Osborne, M.A. (1989) J. Cell Biol. 109, 983–989.
[57] Benditt, J.O., Meyer, C., Fasold, H., Barnard, F.C. and Riedel, N. (1989) Proc. Natl. Acad. Sci. USA 86, 9327–9331.
[58] Li, R. and Thomas, J.O. (1989) J. Cell Biol. 109, 2623–2632.
[59] Newmeyer, D.D. and Forbes, D.J. (1990) J. Cell Biol. 110, 547–557.
[60] Breeuwer, M. and Goldfarb, D.S. (1990) Cell 60, 999–1008.
[61] Yamasaki, L., Kanda, P. and Lanford, R.E. (1989) Mol. Cell. Biol. 9, 3028–3036.
[62] Adam, S.A., Lobl, T.J., Mitchell, M.A. and Gerace, L. (1989) Nature 337, 276–279.
[63] Borer, R.A., Lehner, C.F., Eppenberger, H.M. and Nigg, E.A. (1989) Cell 56, 379–390.
[64] Newmeyer, D.D., Finlay, D.R. and Forbes, D.J. (1986) J. Cell Biol. 103, 2091–2102.
[65] Markland, W., Smith, A.E. and Roberts, B.L. (1987) Mol. Cell Biol 7, 4255–4265.
[66] Imamoto-Sonobe, N., Yoneda, Y., Iwamoto, R., Sugawa, H. and Uchida, T. (1988) Proc. Natl. Acad. Sci USA 85, 3426–3430.
[67] Kalinich, J.F. and Douglas, M.G. (1989) J. Biol. Chem. 264, 17979–17989.
[68] Garcia-Bustos, J.F., Wagner, P. and Hall, M.N. (1990) Exp. Cell Res. (in press).
[69] Parnaik, V.K. and Kennady, P.K. (1990) Mol. Cell Biol. 10, 1287–1292.
[70] Newmeyer, D.D., Lucocq, J.M., Burglin, T.R. and De Robertis, E.M. (1986) EMBO J 5, 501–510.
[71] Akey, C.W. (1990) Biophys. J. 58, 341–355.
[72] Paine, P.L., Moore, L.C. and Horowitz, S.B. (1975) Nature 254, 109–114.