FEBS Letters 416 (1997) 30–34 FEBS 19240

Identification of novel homologues of mouse importin α, the α subunit of the nuclear pore-targeting complex, and their tissue-specific expression

Lyuji Tsuji[a], Toru Takumi[b], Naoko Imamoto[a], Yoshihiro Yoneda[a,*]

[a]*Department of Anatomy and Cell Biology, Osaka University Medical School, 2-2 Yamada-oka, Suita, Osaka 565, Japan*
[b]*Department of Anatomy and Brain Science, Kobe University School of Medicine, 7-5-1 Kusunoki-cho, Chuo-ku, Kobe 650, Japan*

Received 8 July 1997; revised version received 12 August 1997

Abstract Transport of karyophilic proteins into the nucleus is mediated by nuclear localization signals (NLSs) via a multistep process. The karyophiles are recognized by the importin α subunit in the cytoplasm to form a stable complex, termed the nuclear pore-targeting complex (PTAC). To date, three different mammalian α subunits (mSRP1/NPI-1, PTAC58/mPendulin/Rch1 and Qip1) have been identified. In this study, we report the identification of three additional mouse genes homologous to the known α subunits using RT-PCR methodology and show that the mouse α subunits can be classified into at least three subfamilies, α-P, α-Q and α-S families, each composed of closely related members (more than 80% amino acid sequence identity). These three subfamilies, however, have ~50% amino acid identity to one another. Northern blot analysis showed that all were differentially expressed in various mouse tissues. These results suggest that the function of these proteins may be controlled in a tissue-specific manner and that their combinatorial expression may play a role in differentiation and organogenesis.

Key words: Nuclear protein import; Nuclear localization signal; Nuclear pore-targeting complex; Importin α

1. Introduction

Selective nuclear protein import into the nucleus is mediated by short amino acid sequences termed nuclear localization signals (NLSs) [1–5]. This process is divided into multiple, sequential steps, including (a) recognition of an NLS-containing substrate by importin α, (b) the formation of a stable complex, termed the nuclear pore-targeting complex (PTAC) involving importin β, (c) binding of the PTAC to the cytoplasmic face of nuclear pore complex (NPC), (d) the energy-dependent translocation of the complex through the nuclear pore, and (e) the dissociation of the complex on the nucleoplasmic side of the nuclear pores and/or within the nucleus [3–16].

Importin α, the α subunit of PTAC, has a modular structure, which consists of an N-terminal hydrophilic region, a hydrophobic central region composed of 8 repeats, named the arm motifs, and a short hydrophilic C-terminus. A number of related molecules have been identified from a variety of organisms. The N-terminal hydrophilic region contains the 41 amino acid importin β binding domain (IBB domain), which is sufficient for binding importin β and essential for nuclear protein import [17,18]. The amino acid sequence of this domain is strongly basic and highly conserved among importin α homologues. The hydrophobic central region consists of the arm motifs, which were named on the basis of a repetitive structure originally identified in the *Drosophila* segment polarity protein, Armadillo [19], and found in several unrelated proteins that have diverse cellular functions [20].

Two types of mouse importin α, individually referred to as mSRP1 [21] and PTAC58 [13] /mPendulin [21], were previously identified and characterized by means of the digitonin-permeabilized in vitro semi-intact cell system [22]. They have only ~50% amino acid identity to each other, although their functional properties are not clearly distinguished in terms of their ability to carry SV40 T-antigen NLS substrates into the nucleus in the in vitro assay.

Saccharomyces cerevisiae has a single importin α subunit, SRP1, an essential gene which was first identified as a suppressor of temperature-sensitive RNA polymerase I mutations [23]. In contrast, three proteins, Rch1 [24], NPI-1 [25] and Qip1 [26], have been identified as α subunit in humans. Studies involving *Drosophila* suggest that the function of the different types of α subunit may not be identical [27,28]. Moreover, a recent study showed that two types of human α subunit, Rch1 and NPI-1, show different specificities toward a variety of NLS peptides [29]. These data stimulated us to hypothesize the existence of a variety of importin α molecules which may be biologically significant in higher eukaryotes and which may perform unique roles in nuclear protein transport or, possibly, additional functions which have not yet been identified.

The present study reports the identification of three novel mouse proteins homologous to the known mSRP1 and PTAC58/mPendulin. On the basis of sequence similarities, these five mouse proteins can be classified into three major kingdoms, which we refer to as m-importin α-P, α-Q and α-S. In order to avoid confusion of the nomenclature, we have renamed PTAC58/mPendulin and mSRP1 as m-importin α-P1 and m-importin α-S1, respectively. Northern blot analysis indicated that these five molecules show differential expression patterns in tissue-specific manners. These findings suggest that combinatorial expression of these proteins may play some role in cell differentiation and organogenesis.

2. Materials and methods

2.1. cDNA cloning

This procedure was performed as described previously [30]. Briefly, regions conserved among m-importin α-P1, α-S1 and *Xenopus* importin α [10,13,21] were used to design three degenerate oligonucleotide primers for the cloning of mouse importin α-related genes. The sequence of the sense primer was 5′-GA(A/G) CA(A/G) GC(A/G/C/T) GT(A/G/C/T) TGG GC-3′ (the sequence corresponding to nucleotides 538–554 of α-P1) (P58-2). The sequences of the two antisense primers were 5′-GCC CA(A/G/C/T) (A/G)C(A/G/C/T) GC(C/T) TC(C/T) TT-3′ (nucleotides 1199–1183 of α-P1) (P58-12), and 5′-(C/T)TC (A/G/T)AT (C/T)TT (A/G/C/T)TC (A/G/C/T)A(A/G) (A/G/C/T)CC-3′ (nucleotides 1422–1405 of α-P1) (P58-13). Specific murine cDNA, which

*Corresponding author. Fax: (81) (6) 879-3219.
E-mail: yyoneda@anat3.med.osaka-u.ac.jp

PII S0014-5793(97)01092-2

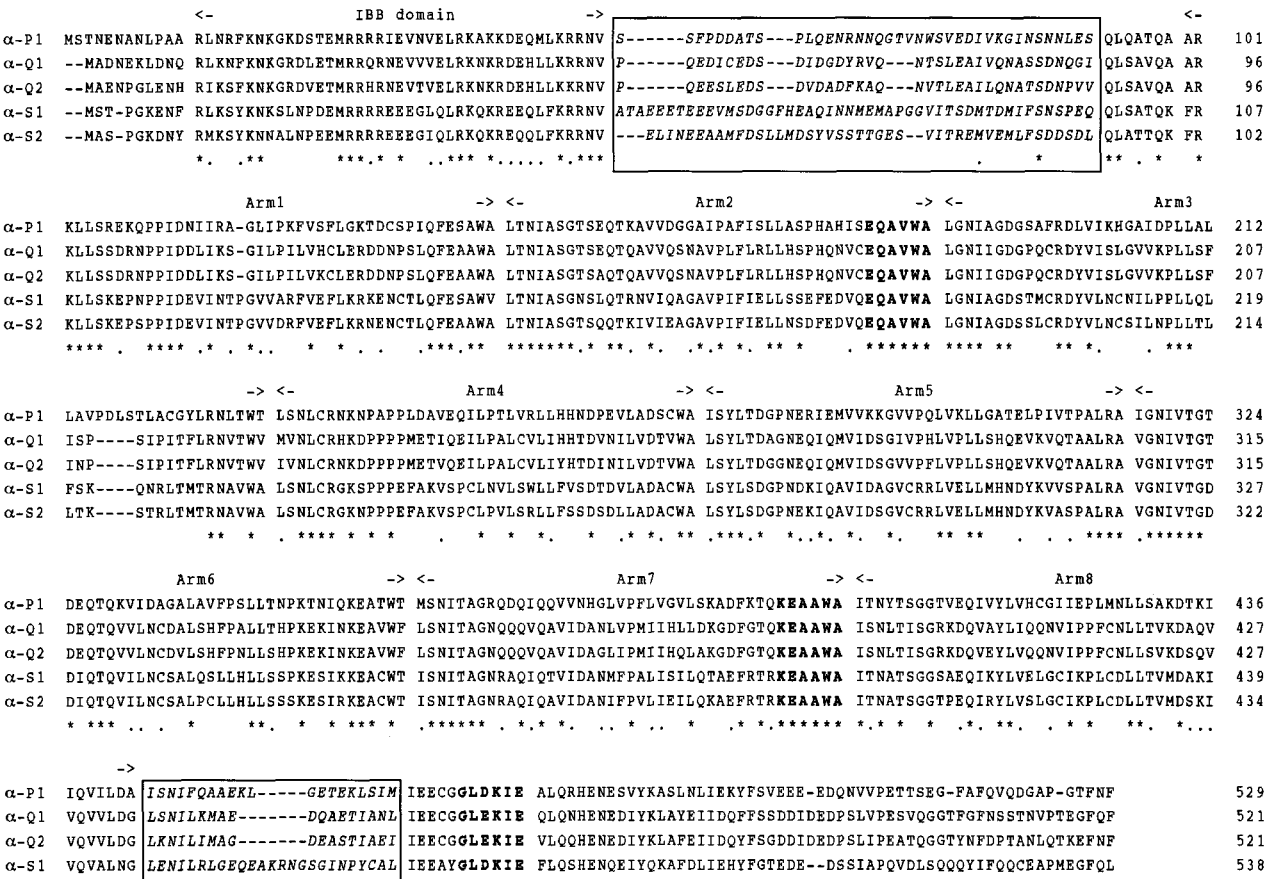

Fig. 1. Multiple sequence alignments among mouse importin α subunits. The amino acid sequences indicated with a single-letter notation are aligned by inserting gaps (-) to achieve maximum homology by the CLUSTAL W (1.60) software with default parameters except that the gap opening penalty of multiple alignment parameters is reduced to 7.00. Asterisks (*) and periods (.) below the alignment indicate positions occupied by identical amino acids and the chemically conserved amino acids among all the sequences, respectively. Variable regions (see text) are boxed. Arm repeats and IBB domain (see text) are indicated above the alignment. Regions used for designing oligonucleotide primers are indicated in boldface. Sequence data are taken from the original papers of the respective cDNA clones, [13] (α-P1), [21] (α-S1). In the figure we have corrected the amino acid sequence of α-S1, namely isoleucine at the 83rd position in place of threonine, originally reported.

was synthesized from 10 μg of total RNA isolated from Ehrlich ascites tumor cells using 100 pmol of P58-13 primer, was used as template for PCR. PCR amplification was performed with P58-2 and P58-12 primers as follows: after a 5 min denaturation at 94°C, the cycling parameters were 94°C for 45 s, 50°C for 1 min, and 72°C for 2 min, for a total of 30 cycles. Amplified products were then ligated to pGEM-T vector (Promega). The clones which contained the amplified fragments of the correct size (about 600 bp) were isolated and sequenced. Using this as a probe, 6×10^4 phage clones of a murine thymus λgt10 cDNA library [31] were hybridized for the isolation of full-sized cDNAs. Both strands of the cDNA sequence were determined using an automated DNA sequencer (LICOR).

2.2. Direct PCR cloning of the missing N-terminal sequence

Two oligonucleotides were designed: the sequence of the antisense oligonucleotides (P10-13) which primes to the 5′ terminal region of incomplete m-importin α-Q2 cDNA was 5′-CCC GAT GCT ATA TTA GTT AGT-3′ and the sequence of the sense oligonucleotides (SK) which primes to the λZAP arm was 5′-CGC TCT AGA ACT AGT GGA TC-3′. The mouse brain uni-ZAP cDNA library (Stratagene), after phenol-chloroform treatment and ethanol precipitation, was used as a template. PCR amplification was performed using AmpliTaq Gold (Perkin Elmer) with P10-13 and SK primers as described previously. The approximately 0.5 kbp fragment obtained was subcloned into the pBluescript II vector (Stratagene) and sequenced. The *Xho*I-*Mun*I fragment obtained from the direct PCR cloning was ligated to incomplete m-importin α-Q2 cDNA.

2.3. Northern blot hybridization

RNAs were isolated from various mouse tissues with TRIzol reagent (BRL). Polyadenylated RNA (2 μg) was prepared using Oligotex-dT30 (Takara), resolved on a 1% agarose gel containing formaldehyde, blotted onto a nylon membrane (Hybond-N^+, Amersham), and hybridized in QuickHyb hybridization solution (Stratagene) according to the manufacturer's recommended protocols. The filter was washed in 0.1×SSC and 0.1% SDS at 60°C and exposed to film (Fuji RX) at −80°C for approximately 18 h.

3. Results and discussion

Regions conserved among m-importin α-P1 (originally named PTAC58/mPendulin), α-S1 (originally named mSRP1) and *Xenopus* importin [10,13,21] were used to design three degenerate oligonucleotide primers for the cloning of importin α-related mouse genes. PCR amplification was performed with P58-2 and P58-12 primers as described in Section 2 using specific murine cDNA derived from Ehrlich ascites tumor cells. The clones, which contained amplified fragments of the expected size (about 600 bp), were sequenced. Four of the total 17 clones were identical to m-importin α-P1, and three to m-importin α-S1. The other 10 cloned fragments,

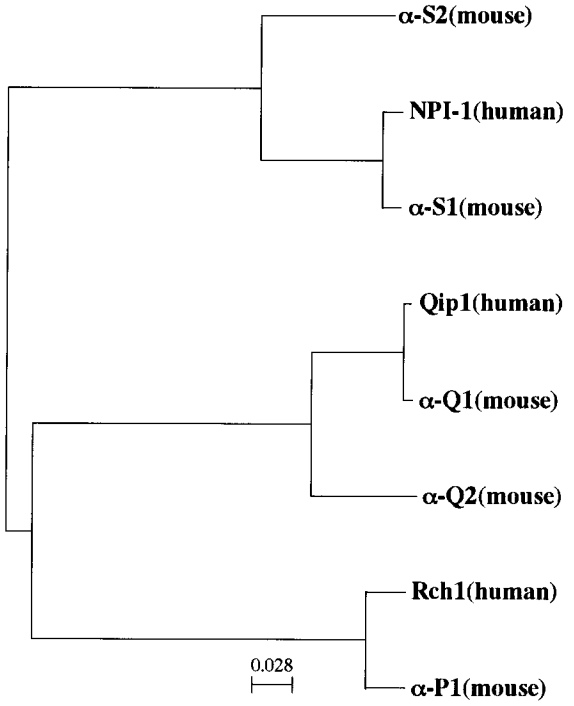

Fig. 2. Phylogenetic tree of mammalian importin α subunits. The tree was constructed using the CLUSTAL W and NJPLOT programs. The branch order and length represent the sequence similarity and divergence, respectively. The scale bar beneath the tree represents a branch length corresponding to a relative difference score of 0.028. Sequence data are taken from [13] (mouse α-P1), [21] (mouse α-S1), [33] (human Rch1), [26] (human Qip1) and [25] (human NPI-1).

however, were not identical but had a high homology to both m-importin α-P1 and m-importin α-S1.

These 10 clones were classified into three groups, based on the sequence similarities. To obtain full-length cDNAs of these three groups, a murine thymus cDNA library was screened using their partial PCR fragments as probes. Three cDNA clones were obtained. However, one of these contained no obvious initial methionine in the sequence. To determine the N-terminal coding and 5′ untranslated sequence of the gene, PCR amplification was utilized for screening the mouse brain uni-ZAP cDNA library as described in Section 2. An approximately 0.5 kbp fragment obtained was sequenced. Since the sequence alignment showed that over 200 bp of the 3′ terminal sequence of the PCR fragment was identical to the 5′ terminal sequence of the incomplete cDNA clone obtained from thymus cDNA library, we concluded that both represented the same gene.

The deduced amino acid sequences of the five independent clones aligned with m-importin α-P1 and m-importin α-S1 are shown in Fig. 1. The relationship of importin α homologues from various species, based on the deduced amino acid sequences, is shown in Table 1. Three novel mouse clones were isolated in this study. m-Importin α-Q1, α-Q2 and α-S2 encode 521, 521 and 533 amino acids with a predicted molecular mass of 57.9 kDa, 57.8 kDa and 59.6 kDa, respectively. Each clone has 98%, 85% and 97% amino acid identity with human Qip1, m-importin α-Q1 and m-importin α-S1, respectively. As expected, the regions used for designing degenerated oligonucleotides, P58-2 and P58-12, are identical among these proteins (Fig. 1). The positions of the arm repeats are also shown in Fig. 1. The IBB domain and arm motifs are well conserved among these five members (see Fig. 1), which suggests that all are involved in nuclear protein transport.

As shown in Fig. 2, the deduced amino acid sequences of importin α isoforms reveal that the mouse isoforms can be classified into three subfamilies, the α-P, α-Q and α-S families. While the members of each subfamily have a high degree (greater than 80%) of amino acid identity to each other, these subfamilies have only about 45% identity with one another.

The basic local alignment search tool (BLAST) [32] analysis showed that a variety of proteins similar to murine m-importin α have been isolated from species other than mouse: NPI-1/hSRP1 and Rch1/hSRP1α [33] in human, importin 60-1 and 60-2 in *Xenopus* [10], pendulin in *Drosophila* [27,28] and SRP1 in yeast. The percentages amino acid sequence identity among all of these α subunits calculated using CLUSTAL W software [34] are shown in Table 1. In addition to mammals, a number of molecules homologous to m-importin α have been isolated from nonmammals as described previously, but their relationship to the mammalian subfamilies has not yet been

Table 1
Percent identity among importin α subunits from a variety of organisms

Species	Mouse					Human			*Xenopus*		*Drosophila*	Yeast
	α-P1	α-Q1	α-Q2	α-S1	α-S2	Rch1	Qip1	NPI-1	60-1	60-2	pendulin	SRP1
α-P1		49	48	44	43	94	48	43	61	59	47	44
α-Q1			85	46	47	49	98	46	45	46	45	45
α-Q2				46	47	48	85	46	47	47	44	45
α-S1					81	44	46	97	43	43	43	47
α-S2						44	46	81	44	43	46	48
Rch1							49	44	62	60	48	44
Qip1								46	46	46	45	45
NPI-1									43	42	43	49
60-1										95	48	43
60-2											48	42
pendulin												40
SRP-1												

Amino acid sequence identity among importin α subunits was calculated using CLUSTAL W software. Sequence data are taken from [13] (mouse α-P1), [21] (mouse α-S1), [33] (human Rch1), [26] (human Qip1), [25] (human NPI-1), [10] (*Xenopus* 60-1 and 60-2), [28] (*Drosophila* pendulin) and [23] (yeast SRP1).

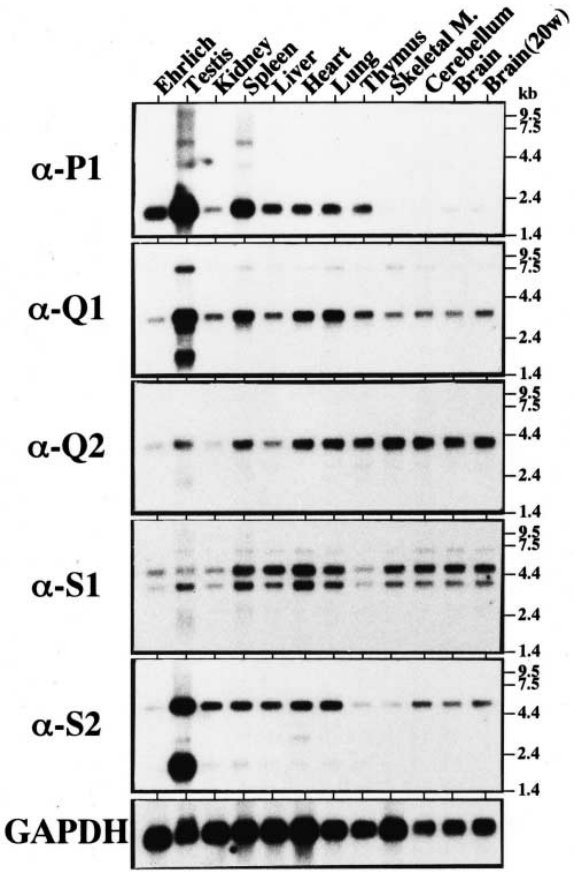

Fig. 3. Northern blot of mouse importin α isoform mRNA. Polyadenylated mRNA (2 μg) from various mouse tissues was analyzed for the presence of transcripts hybridizing to the following cDNA as probes. An approximate 0.7 kbp fragment from the initial methionine to the *Acc*I restriction site in α-P1 cDNA (upper panel), the *Nco*I-*Apa*I fragment in α-Q1 cDNA (approximately 0.8 kbp) (second panel), the *Nco*I-*Acc*I fragment in α-Q2 cDNA (approximately 0.8 kbp) (third panel), an approximately 0.9 kbp fragment from the −144 bp 5′ non-coding region to the *Bsm*I restriction site in α-S1 cDNA (fourth panel), an approximately 0.7 kbp fragment from the initial methionine to the *Apa*I restriction site in α-S2 cDNA (fifth panel), and glyceraldehyde 3-phosphate dehydrogenase as a control (lower panel) were used for probes. The same filter was used for all hybridizations. 'Ehrlich' indicates Ehrlich ascites tumor cells, and 'Brain(20w)' indicates brain (without cerebellum) mRNA derived from a mouse 20 weeks after birth and other tissues derived from mice 5 weeks after birth. The positions of the RNA size marker (BRL) are shown on the right-hand side.

clearly defined. Whether such multiplicity is unique to mammalian systems can only be speculated on, since it has not yet been established if yeast, *Drosophila*, or *Xenopus* possess multiple functional subfamilies. Detailed structure/function analysis will be required to determine which of these corresponds to mammalian subgroups.

Multiple alignment analysis revealed that sequence-variable domains in the region of about 40 amino acids which are flanked by the IBB domain exist and the first arm motif, and the region just after the eighth arm repeat (boxed in Fig. 1). The amino acid sequence and the length of these regions differ greatly among the five subunits. m-Importin α-S1 and m-importin α-S2 have 81% identity over the entire molecule, but their variable regions have only 25% identity. A recent study showed that two types of human α subunit have different specificities to various NLS peptides [29]. Very recently, we found that tyrosine phosphorylated Stat1 interacts with NPI-1 but not with Rch1 (T. Sekimoto, submitted). These results suggest the possibility that closely related subfamily members differ functionally and that these variable regions may play essential roles in a diversity of functions.

In order to better understand the biological significance of the various α subunits, we compared their expression patterns in a variety of mouse tissues by Northern blot analysis (Fig. 3). The data showed significantly different patterns in the mRNA expression of these murine α subunits, suggesting that these proteins may have functional diversity. Although the two members of each subfamily have very high sequence similarity, their expression patterns were not similar. m-Importin α-Q1 and α-Q2 transcripts were detected more or less in all tissues examined (Ehrlich ascites tumor cells, testis, kidney, spleen, liver, heart, lung, thymus, skeletal muscle, cerebellum and brain (without cerebellum)), while multiple-sized m-importin α-Q1 transcripts were highly expressed, especially in testis. Multiple-sized m-importin α-S1 transcripts were detected at low levels in all tissues examined. In contrast, m-importin α-S2 transcripts were only slightly detected in Ehrlich ascites tumor cells, thymus and skeletal muscle, while the single transcript of α-S2 was clearly detected in kidney, spleen, liver, heart and lung, and a high expression of two major transcripts of 5.5 and 1.8 kbp was observed, especially in testis. m-Importin α-P1 transcripts were only slightly detected in brain but not in cerebellum and skeletal muscle, while multiple-sized transcripts were highly expressed in testis and spleen. These results show that m-importin α-Q1, α-Q2, and α-S1 were ubiquitously expressed to a variety of extents, while the expression of m-importin α-P1 and α-S2 was restricted to certain tissues. Furthermore, a recent study showed that at least three structurally and functionally distinct α subunits exist in a human single cell population such as HeLa cells (Y. Miyamoto, submitted). It is important to understand how the expression of each isoform is controlled in each tissue and/or at which developmental stage. The data herein will be helpful in terms of a better understanding of the role and mechanism of the nuclear protein transport in development and organogenesis.

In summary, we isolated three novel mouse genes, m-importin α-Q1, α-Q2 and α-S2, indicating that at least five isoforms of importin α exist in mice. Amino acid sequence analysis revealed that the mouse α subunits are grouped into at least three subfamilies, α-P, α-Q and α-S, each being composed of closely related members. Two members of each subfamily have a high degree of amino acid similarity (more than 80% identity), but Northern blot analysis showed that expression patterns of these two members differ in various mouse tissues. Although it is important to understand which specific type of karyophilic molecules each m-importin α isoform carries into the nucleus in each cell, these data suggest that the combinatorial expression of these proteins in each cell may play a role in cell differentiation and organogenesis.

Acknowledgements: This work was supported by a Grant-in-Aid for Scientific Research on Priority Areas (No. 07282103), a Grant-in-Aid for Scientific Research (B) (No. 08458229), and a Grant-in-Aid for COE Research (No. 07CE2006) from the Japanese Ministry of Education, Science, Sports and Culture, the Nissan Science Foundation and the Naito Foundation. We thank Dr. Y. Tsujimoto for the mouse brain cDNA library and Dr. Y. Minami for GAPDH cDNA.

References

[1] Silver, P.A. (1991) Cell 64, 489–497.
[2] Garcia-Bustos, J., Heitman, J. and Hall, M.N. (1991) Biochim. Biophys. Acta 1071, 83–101.
[3] Dingwall, C. and Laskey, R.A. (1991) Trends Biochem. Sci. 16, 478–481.
[4] Görlich, D. and Mattaj, I.W. (1996) Science 271, 1513–1518.
[5] Nigg, E.A. (1997) Nature 386, 779–787.
[6] Newmeyer, D.D. and Forbes, D.J. (1988) Cell 52, 641–653.
[7] Richardson, W.D., Mills, A.D., Dilworth, S.M., Laskey, R.A. and Dingwall, C. (1988) Cell 52, 655–664.
[8] Adam, S.A. and Gerace, L. (1991) Cell 66, 837–847.
[9] Moore, M.S. and Blobel, G. (1992) Cell 69, 939–950.
[10] Görlich, D., Prehn, S., Laskey, R.A. and Hartmann, E. (1994) Cell 79, 767–778.
[11] Görlich, D., Kostka, S., Kraft, R., Dingwall, C., Laskey, R.A., Hartmann, E. and Prehn, S. (1995) Curr. Biol. 5, 383–392.
[12] Imamoto, N., Tachibana, T., Matsubae, M. and Yoneda, Y. (1995) J. Biol. Chem. 270, 8559–8565.
[13] Imamoto, N., Shimamoto, T., Takao, T., Tachibana, T., Kose, S., Matsubae, M., Sekimoto, T., Shimonishi, Y. and Yoneda, Y. (1995) EMBO J. 14, 3617–3626.
[14] Imamoto, N., Shimamoto, T., Kose, S., Takao, T., Tachibana, T., Matsubae, M., Sekimoto, T., Shimonishi, Y. and Yoneda, Y. (1995) FEBS Lett. 368, 415–419.
[15] Görlich, D., Vogel, F., Mills, A.D., Hartmann, E. and Laskey, R.A. (1995) Nature 377, 246–248.
[16] Moroianu, J., Hijikata, M., Blobel, G. and Radu, A. (1995) Proc. Natl. Acad. Sci. USA 92, 6532–6536.
[17] Görlich, D., Henklein, P., Laskey, R.A. and Hartmann, E. (1996) EMBO J. 15, 1810–1817.
[18] Weis, K., Ryder, U. and Lamond, A.I. (1996) EMBO J. 15, 1818–1825.
[19] Riggleman, B., Wieschaus, E. and Schedl, P. (1989) Genes Dev. 3, 96–113.
[20] Peifer, M., Berg, S. and Reynolds, A.B. (1994) Cell 76, 789–791.
[21] Prieve, M.G., Guttridge, K.L., Munguia, J.E. and Waterman, M.L. (1996) J. Biol. Chem. 271, 7654–7658.
[22] Adam, S.A., Marr, R.S. and Gerace, L. (1990) J. Cell Biol. 111, 807–816.
[23] Yano, R., Oakes, M., Yamaghishi, M., Dodd, J.A. and Nomura, M. (1992) Mol. Cell. Biol. 12, 5640–5651.
[24] Cuomo, C.A., Kirch, S.A., Gyuris, J., Brent, R. and Oettinger, M.A. (1994) Proc. Natl. Acad. Sci. USA 91, 6156–6160.
[25] O'Neill, R.E. and Palese, P. (1995) Virology 206, 116–125.
[26] Seki, T., Tada, S., Katada, T. and Enomoto, T. (1997) Biochem. Biophys. Res. Commun. 234, 48–53.
[27] Török, I., Strand, D., Schmitt, R., Tick, G., Török, T., Kiss, I. and Mechler, B.M. (1995) J. Cell Biol. 129, 1473–1489.
[28] Küssel, P. and Frasch, M. (1995) J. Cell Biol. 129, 1491–1507.
[29] Nadler, S.G., Tritschler, D., Haffar, O.K., Blake, J., Bruce, A.G. and Cleaveland, J.S. (1997) J. Biol. Chem. 272, 4310–4315.
[30] Takumi, T., Ishii, T., Horio, Y., Morishige, K., Takahashi, N., Yamada, M., Yamashita, T., Kiyama, H., Sohmiya, K., Nakanishi, S. and Kurachi, Y. (1995) J. Biol. Chem. 270, 16339–16346.
[31] Ogata, M., Sawada, M., Kosugi, A. and Hamaoka, T. (1994) J. Immunol. 153, 4478–4487.
[32] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) J. Mol. Biol. 215, 403–410.
[33] Weis, K., Mattaj, I.W. and Lamond, A.I. (1995) Science 268, 1049–1053.
[34] Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) Nucleic Acids Res. 22, 4673–4680.