The Nuclear Localization Signal of Lymphoid Enhancer Factor-1 Is Recognized by Two Differentially Expressed Srp1-Nuclear Localization Sequence Receptor Proteins*

Mary G. Priewe, Katherine L. Guttridge, Jesus E. Munguia, and Marian L. Waterman‡

From the Department of Microbiology and Molecular Genetics, University of California, Irvine, California 92717-4025

Proteins are directed to the nucleus by their nuclear localization sequences (NLSs) in a multistep process. The first step, which is to dock the NLS-containing protein to the nuclear pore, is carried out in part by a recently identified NLS receptor named Srp1/importin-α. Using the high mobility group (HMG) DNA binding domain of human lymphoid enhancer factor-1 (hLEF-1) as bait in a yeast two-hybrid screen, we have identified two different mouse Srp1 proteins (pendulin/importin-α and mSrp1) that each bind to a 9-amino acid sequence in hLEF-1 called the B box. We show that the B box of hLEF-1, a region essential for high affinity DNA binding, is also necessary and sufficient for nuclear localization, lending support to the model that NLSs can function both in nuclear transport and DNA binding. Pendulin and mSrp1 are the mouse homologues of hRch1/hSrp1/importin-α and hSrp1/karyopherin α/NPI-1, respectively, and show considerable sequence divergence from each other. We find a surprising and significant difference in the expression pattern of pendulin and mSrp1 mRNA, suggesting that these two Srp1 proteins are distinguishable in function as well as sequence.

Analysis of the nuclear localization of transcription regulatory factors has largely focused on defining the amino acid sequences within transcription factor proteins that direct nuclear entry. A universal consensus sequence that signals nuclear localization (NLS) has not been found; nevertheless, short regions, rich in basic residues are a feature common to most identified NLSs (1, 2). It is also common for the NLS of transcription factors to be in regions that overlap or are near the DNA binding domain, prompting speculation that (a) nuclear occupancy by a protein is due to both nuclear import and retention by DNA binding, (b) NLS sequences contribute to DNA binding, or (c) there is coordinate regulation of the two functions due to their proximity (3). It is known that the NLS directs transport through nuclear pores via a receptor-mediated, multistep process (4–6). Recently, a cytosolic NLS receptor protein in Xenopus named importin-α was found to carry out the first step along with the Xenopus importin-β subunit (also known as p97 or karyopherin β) by docking NLS-containing proteins at the nuclear pore (7). Xenopus importin-α is homologous to yeast Srp1p, an essential gene first identified as a suppressor of RNA polymerase I mutations in Saccharomyces cerevisiae (8). Immunofluorescence studies localized ySrp1p mainly to the nuclear pore, and genetic and biochemical evidence established a tight association of ySrp1p with known protein components of the pore (8, 9). More recently, work with human Srp1α (hSrp1α; also known as hRch1) and its mouse homolog pendulin/importin-α, and hSrp1/karyopherin α/NPI-1, a different human Srp1 protein, has confirmed in vivo and in vitro that Srp1 proteins and p97/karyopherin β function as cytoplasmic NLS receptors that usher NLS-containing proteins to the cytoplasmic side of the nuclear pore, where additional GTP-requiring factors complete the second step of translocation through the pore (10–12). Finally, Gorlich et al. (13) have shown that nuclear pore binding occurs mainly via Xenopus importin-β and that importin-α translocates through the pore with its NLS substrate to accumulate in the nucleus.

The lymphocyte transcription factor hLEF-1 (for human lymphoid enhancer factor) is a member of the high mobility group (HMG) family of DNA-binding proteins (14, 15). hLEF-1 binds DNA through a single HMG homologous region (HMG box) near the carboxyl terminus resulting in a 130° bend in the substrate DNA (16). Previous studies to delimit the DNA binding region of hLEF-1 determined the HMG box to be necessary for specific sequence recognition and a COOH-terminal 9-amino acid region rich in basic residues called the B box, to be necessary for high affinity binding to the DNA target (17). The HMG and B boxes together are sufficient for independent, high affinity recognition of LEF-1 binding sites, as well as DNA bending that is indistinguishable from full-length hLEF-1 protein (17–18). We show here that in addition to its role in DNA binding, the B box also functions as a NLS and is bound by at least two different Srp1 nuclear transport proteins.

EXPERIMENTAL PROCEDURES

Yeast Two-hybrid Screening—The yeast two-hybrid screening procedure was performed as described by Durfee et al. (19). The yeast strain (Y190), cDNA plasmid library (mouse thymocyte cDNA into pACT1) were a generous gift of Dr. S. Elledge (Baylor College of Medicine, Houston, TX). The coding region for the HMG DNA binding domain of hLEF-1 (amino acid 297–399, written as hLEF-1(297–399)) was inserted in the pAS-1 plasmid 3′ of the yeast Gal4 DNA binding domain. Other HMG DNA binding domains (hSRY-(48–162), hTCF-1-(152–254), and ySte-11-(12–96)) were similarly constructed for use as control baits. Plasmid constructs and positive cDNA clones were sequenced using specific oligonucleotide primers and the Sequenase kit. Nucleotide sequence analysis and amino acid comparisons were per-
formed using Lasergene software for the Macintosh (DNASTAR Inc., Madison, WI).

In Vitro Binding Assays—Pendulin and mSrp1 clones (A and H, respectively) were inserted in-frame, 3′ of the glutathione S-transferase gene in pGEX-3X (Pharmacia Biotech Inc.), and recombinant protein obtained by overexpression in isopropyl 1-thio-β-D-galactopyranoside-treated DH5α (0.2 mg/ml, 4 h, 37°C). Cells from 1 liter of culture were lysed by sonication in 20 ml of lysis buffer (1% PBS, 1% Triton X-100), and clarified extract applied to a 1-ml glutathione-Sepharose column. The column was washed with lysis buffer (20 column volumes), followed by a wash with column buffer (20 column volumes; 50 mM Tris-HCl (pH 8.0), 1 mM dithiothreitol, 1 mM EDTA, 5% glycerol, 50 mM ammonium sulfate). GST-fusion proteins were eluted in 5 ml of column buffer supplemented with 0.2 x glutathione. Protein was quantified by SDS-PAGE and Coomassie staining or by the Bradford assay. GST fusion protein was detected with either GFP antisera (1:1000, Clontech) or affinity-purified hLEF-1 antibody (1:1000; Ref. 17), followed by incubation with FITC-conjugated anti-rabbit antibody for 1 h at room temperature. Cells were fixed with 3% formaldehyde, 1 x PBS for 30 min at room temperature, washed and incubated with primary antibody diluted into 3% BSA, 1 x PBS for 1 h at room temperature, followed by a wash and incubation with FITC-conjugated anti-rabbit antibody for 1 h at room temperature.

Northern Analysis—Total RNA was purified from mouse tissues using RNA STAT-60 reagent according to the manufacturer’s suggestions (Tel-Test, Friendswood, TX). Northern analysis of the RNA samples was performed as described (23). Specific activities of the random-prime labeled probes (see Fig. 2) averaged 5.0 x 106 cpm/µg of DNA.

RESULTS AND DISCUSSION

To identify proteins that bind directly to HMG DNA binding domains, we performed a yeast two-hybrid screen of a mouse thymocyte cDNA library fused with the yeast Gal4 activation domain (amino acids 768 – 881), using the HMG DNA binding domain of hLEF-1 (amino acids 297 – 399) fused to the DNA binding domain of Gal4 (amino acids 1-147) as bait. Of 1.4 x 109 independent clones analyzed, 10 gave strong positive signals in yeast expressing unrelated baits such as the protein OM2 or a reverse transcriptase from the yeast Ty3 retrotransposon. In addition, we observed no interaction, as judged by growth on medium lacking histidine and expression of β-galactosidase activity, with the DNA binding domains of three other HMG family members: yeast Ste11, human TCF-1, or human SRY. BLAST analysis identified 8 of the 10 clones to be Srp1 proteins. Seven clones encoded the same protein and were nearly identical at the nucleotide level to a GenBank entry called mouse pendulin (accession no. U12270 (1995); a Srp1 protein first identified in Drosophila as pendulin or Oho-31 (24, 25). Recently, a second BLAST search identified a new GenBank entry called mouse importin-α that also had near sequence identity with the seven clones (accession no. D55720 (1995); Ref. 12). Since the sequence variations between pendulin, importin-α, and the seven yeast two hybrid clones described here are minor, each is likely to encode the same Srp1 protein; to simplify discussion, we will refer to our clones as pendulin. The eighth clone is a distinct gene identical at the amino acid level to mouse mSrp1 (mSrp1), a variant Srp1 protein more closely related to yeast Srp1p (26).

Srp1 protein can be divided into an NH2-terminal hydrophilic region, a hydrophobic central region composed of 8–10 degenerate repeats called “arm” repeats, and a short hydrophilic COOH terminus (27). Fig. 1 shows the amino acid sequence of the two Srp1 variants identified in our screen and the positions at which each isolated clone begins relative to the positions of the arm repeats (Fig. 1, panels A and B). Arm repeats are so named on the basis of their similarity to an amino acid repeat structure first identified in the Drosophila segment polarity protein Armadillo. Although the precise function of multiple, degenerate arm repeats in Srp1 proteins is not known, it is possible that they bind an array of different NLS signals. Indeed, hSrp1α, the human homologue of pendulin, has been shown to directly interact with two different types of NLS sequences (10). An exact determination of the number and borders of arm repeats in Srp1 proteins awaits a detailed structure/function analysis, and for that reason they have been variously assigned by different investigators. Fig. 1 shows an alignment of arm repeats in pendulin and mSrp1 with borders similar to those described by Peifer et al. (28) and Torok et al. (25), as well as a ninth, degenerate arm repeat that has been noted in Drosophila pendulin (24). All of the Srp1 cDNA isolates were partial clones, yet all were able to interact with hLEF-1 in vivo, suggesting that at least the first three arm repeats of pendulin and the first two arm repeats of mSrp1 are dispensable for binding to the hLEF-1 bait (Fig. 1). Multiple alignment analysis of the arm repeats within pendulin show a great deal of diversity in repeat length because more spacing must be introduced for an optimal alignment (Fig. 1, panel A). This is not the case for mSrp1, which requires minimal changes in spacing for alignment (Fig. 1, panel B).

Amino acid sequence comparisons of pendulin and mSrp1 with other Srp1 proteins show that mouse pendulin (or importin-α) is the homologue of human Rch1/NIPL-1/karyopherin α hSrp1α (94% sequence identity) and mSrp1 the homologue of human hSrp1 (97%) (24). Surprisingly, pendulin is no more similar to mSrp1 than it is to yeast Srp1p (45% identity to each), revealing a natural division of the mammalian Srp1 proteins into two subtypes (Fig. 1C). To date, only one type of Srp1 protein has been identified in Drosophila (as pendulin or Oho-31) and Xenopus (importin-α) and each is more similar in amino acid sequence and arm repeat spacing to the pendulin/
hRch1/hSrp1 class than to the mSrp1 class, which is more closely related to yeast Srp1p (7, 24, 25). The existence of two Srp1 subtypes in mouse and human and the distinct arrangement of their respective arm repeat structures suggests that each subtype may perform unique functions in nuclear transport or additional functions not yet identified. Whether these additional functions are unique to mammalian systems is only speculative, since it is not yet established whether yeast, Drosophila, or Xenopus have one or two Srp1 subtypes.

The reported tissue-specific and developmental stage-specific pattern of expression for Drosophila pendulin (24, 25) prompted us to compare the expression pattern of pendulin and mSrp1 by Northern analysis (Fig. 2). A 4.3-kilobase RNA hybridizing to a mSrp1 probe appears at a low level in all tissues examined (lung, liver, spleen, thymus, heart, brain, cerebellum, uterus, and kidney), an expression pattern expected of a protein involved in an essential cellular function. In contrast, expression of the pendulin gene is highly variable; a single 2.4-kilobase RNA species appears in all of the tissues surveyed at a low level, but is very highly expressed in thymus, spleen, and heart. This great difference between pendulin and mSrp1 expression patterns is a distinguishing characteristic that supports a grouping of mammalian Srp1 proteins into two subtypes. It also suggests that pendulin performs an additional or unique function in tissues that express high levels of this protein. In Drosophila, pendulin/Oho-31 was identified as a tumor suppressor gene, as its absence causes over-proliferation of the hematopoietic organ (24, 25). Drosophila pendulin also localizes to the nucleus precisely at the G2/M stage of the cell cycle. It will be interesting to compare the nuclear localization patterns of the pendulin and mSrp1 subtypes relative to cell cycle progression and determine whether this might be yet another distinguishing characteristic indicative of separate functions in the cell.

In order to compare the binding specificities of pendulin and mSrp1 for hLEF-1 and to map the hLEF-1 domains involved in binding, we established an in vitro binding assay with GST fused to pendulin (clone A) or mSrp1 (clone H). Recombinant GST fusion proteins were bound to glutathione-Sepharose beads and incubated with in vitro translated, 35S-labeled
hLEF-1 deletion mutants or control proteins (Fig. 3). Neither the DNA binding domain of Gal4 (data not shown) nor a hLEF-1 deletion mutant missing the entire HMG DNA binding domain (lane 1) was precipitated with either GST-pendulin-bound or GST-mSrp1-bound glutathione-Sepharose. In contrast, protein fragments containing the HMG DNA binding domain were specifically precipitated with both GST-pendulin and GST-mSrp1 (lanes 2, 4, 5, and 8) but not with GST alone bound to beads (data not shown). This confirmed that the DNA binding domain of hLEF-1 was specifically associating with Srp1 protein and not GST or glutathione-Sepharose. To precisely map the region of the HMG DNA binding domain recognized by pendulin and mSrp1, we examined a set of deletions surrounding the HMG and B boxes. Precipitation of the DNA binding domain with GST-pendulin was most efficient when all residues of the B box were present (KKKKRK; lane 8). Although in the experiment shown there is little apparent binding of mSrp1 to the same hLEF-1 deletion mutant in lane 8, we observed in two additional repeat experiments that mSrp1 bound as well to this fragment as to the hLEF-1 deletion mutants in lanes 4 and 5. Continued deletion of the B box to 6 (KKKKR; lane 9) and then 4 residues (KK; lane 11) decreased binding of both to pendulin by 5-fold, but did little to disrupt binding by mSrp1. A mutant B box in which two lysines were substituted for two arginines appeared to bind better to pendulin than its wild type counterpart (KKRRK; lane 10). Finally, deletion of the B box to 2 residues (KK; lane 12) dramatically disrupted interaction with pendulin and mSrp1 15-fold. A residual amount of binding (4%) was detectable with hLEF-1 fragments that did not contain the B box, suggesting that residues in adjacent regions might also contribute to binding (panel A, compare lanes 8, 13, and 14).

One particularly interesting feature of hLEF-1 is that its HMG DNA binding domain is almost identical in amino acid sequence (98% amino acid identity) to that of hTCF-1, a second HMG lymphocyte transcription factor that binds and bends identical DNA sequences in promoters and enhancers of T cell-specific genes (29, 30). As mentioned above, neither pendulin nor mSrp1 appeared to interact with the hTCF-1 control...
bait in yeast. A lack of interaction with hTCF-1 is surprising because the amino acid sequence of the B box differs by only 2 residues (KKKRRSREK), one of which is a conservative change from lysine to arginine. We prepared in vitro translated Gal4/hLEF-1 (297–399) and Gal4/hTCF-1 (152–294) fusion protein and tested each for binding to pendulin and mSrpl in vitro.

Consistent with our observations in yeast, pendulin bound poorly to the HMG DNA binding domain of hTCF-1 (Fig. 3A, lanes 2 and 3), showing a 4-fold drop in Gal4/hTCF-1 recovered with GST-pendulin beads compared to Gal4/hLEF-1; in other experiments we have observed differences as great as 20-fold. In contrast, mSrpl binding was significantly weaker and equal for Gal4/hLEF-1 and Gal4/hTCF-1, a result we do not observe in yeast (Fig. 3B, lanes 2 and 3). Whether this weaker, equal binding reflects the abundance of truncated translation products in the binding reaction, truncated mSrpl protein, or is indicative of a true difference in the recognition of hLEF-1 by pendulin and mSrpl, will be determined by further study with full-length proteins. That pendulin appears to interact with hLEF-1 and not hTCF-1 is interesting and may reveal a function for these two functions.

HMG DNA binding domain eliminates nuclear localization (GPf/hLEF-1(297–373)). Again, although intrinsic green fluorescence is not detectable because of greatly reduced nuclear localization, GFP/hLEF-1 (297–373) was detected in the cytoplasm by immunofluorescence using either affinity-purified hLEF-1 or GFP antisera. Finally, fusion of sequences encoding only the B box (GPf/hLEF-1(374–384)) caused exclusive localization of GFP to the nucleus. These data show that the B box, which binds pendulin and mSrpl, is necessary and sufficient for nuclear localization and is most likely the major NLS of hLEF-1.

In summary, we have shown that the B box, a 9-amino acid motif within the hLEF-1 HMG DNA binding domain, functions as an NLS to direct nuclear localization and binds to two different Srpl NLS receptor proteins (one of which is highly expressed in thymus, heart, and spleen). It is significant that this same motif also plays an important role in DNA binding by contacting the sugar-phosphate backbone of the DNA recognition site (18). We note that Srpl transport proteins move through nuclear pores with their NLS substrates (13), and suggest that direct binding of the B box both to NLS receptor proteins and to DNA may indicate coordinate regulation of these two functions.

Acknowledgments—We thank M. Nomura, S. Sandmeyer, R. Sandri-Goldie, E. Ehrenfeld, and H. Mangalam for critique and helpful discussions. We are especially grateful to H. Mangalam for expert advice and help with computer graphics and image analysis.

REFERENCES

1. Osborne, M. A., and Silver, P. A. (1993) Annu. Rev. Biochem. 62, 219–254
2. Dingwall, C., and Laskey, R. A. (1991) Trends Biochem. Sci. 16, 478–481
3. LaCasse, E. C., and Lefebvre, Y. A. (1995) Nucleic Acids Res. 23, 1647–1656
4. Akey, C. W., and Goldfarb, D. S. (1989) J. Cell Biol. 109, 971–1008
5. Nomura, M., Berg, S., and Reynolds, A. B. (1994) Cell 75, 641–653
6. Weis, K., Mattaj, I. W., and Lamond, A. I. (1995) J. Cell Biol. 126, 619–630
7. Weis, K., Mattaj, I. W., and Lamond, A. I. (1995) Science 268, 1049–1053
8. Mikiyama, J., Hidaka, G., and Radu, A. (1995) Proc. Natl. Acad. Sci. 92, 2008–2011
9. Mikiyama, J., Shimamoto, T., Takao, T., Tachibana, T., Kose, S., Matsubae, M., Somooka, T., Shimohishii, Y., and Yoneda, Y. (1995) EMBO J. 14, 3617–3626
10. Goodrich, D., Vogel, F., Mills, A. D., Hartmann, E., and Laskey, R. A. (1995) Nature 377, 246–248
11. Waterman, M. L., Fischer, W. H., and Jones, K. A. (1994) Genes & Dev. 5, 656–669
12. Travis, A., Amsterdam, A., Belanger, C., and Grosschedl, R. (1991) Genes & Dev. 5, 880–894
13. Giese, K., Cox, J., and Grosschedl, R. (1992) Cell 69, 185–196
14. Carlson, P., Waterman, M., and Jones, K. (1993) Genes & Dev. 7, 2418–2430
15. Love, J. J., Li, X., Case, D. A., Giese, K., Grosschedl, R., and Wright, P. E. (1995) Nature 376, 793–795
16. Dorfue, T., Biecherer, K., Chen, P.-L., Ye, S.-H., Yang, Y., Kilburn, A. E., Lee, W.-H., and Eledge, S. J. (1993) Genes & Dev. 7, 555–569
17. Azuma, T., Tabb, M. M., Vu, L., and Nomura, M. (1995) Proc. Natl. Acad. Sci. in press
18. Heim, R., Cubitt, A. B., and Tsien, R. Y. (1995) Nature 373, 663–664
19. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994) Science 263, 802–805
20. Virca, G. D., Northemann, W., Shiels, B. R., Widera, G., and Broome, S. (1990) BioTechniques 8, 370–371
21. Kussel, P., and Frasch, M. (1995) J. Cell Biol. 129, 1491–1507
22. Cortes, P., Ye, Z. S., and Baltimore, D. (1994) Proc. Natl. Acad. Sci. 91, 7633–7637
23. Yano, R., Oakes, M., Yamagishi, M., Dodd, J. A., and Nomura, M. (1994) Nature 370, 6880–6884
24. Peifer, M., Berg, S., and Reynolds, A. B. (1994) Cell 76, 789–791
25. van Houte, L., van Oers, A., van de Wetering, M., Dooijes, D., Kapiént, R., and Clevers, H. (1993) J. Biol. Chem. 268, 18083–18087
26. Giese, K., Amsterdam, A., and Grosschedl, R. (1991) Genes & Dev. 5, 2567–2578

Downloaded from http://www.jbc.org/ by guest on July 27, 2018
The Nuclear Localization Signal of Lymphoid Enhancer Factor-1 Is Recognized by Two Differentially Expressed Srp1-Nuclear Localization Sequence Receptor Proteins

Mary G. Prieve, Katherine L. Guttridge, Jesus E. Munguia and Marian L. Waterman

J. Biol. Chem. 1996, 271:7654-7658.
doi: 10.1074/jbc.271.13.7654

Access the most updated version of this article at http://www.jbc.org/content/271/13/7654

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 29 references, 16 of which can be accessed free at http://www.jbc.org/content/271/13/7654.full.html#ref-list-1