FEBS Letters 484 (2000) 118–124 FEBS 24225

Two nuclear localization signals are required for nuclear translocation of nuclear factor 1-A

Masayoshi Imagawa[a,b,*], Reiko Sakaue[a], Atsuhiro Tanabe[a], Shigehiro Osada[a], Tsutomu Nishihara[a]

[a] *Laboratory of Environmental Biochemistry, Graduate School of Pharmaceutical Sciences, Osaka University, 1-6 Yamada-Oka, Suita, Osaka 565-0871, Japan*
[b] *Department of Microbial Chemistry, Faculty of Pharmaceutical Sciences, Nagoya City University, 3-1 Tanabe-dori, Mizuho-ku, Nagoya, Aichi 467-8603, Japan*

Received 26 July 2000; revised 19 September 2000; accepted 29 September 2000

First published online 19 October 2000

Edited by Felix Wieland

Abstract Nuclear factor 1 (NF1) proteins are encoded by at least four genes (NF1-A, B, C, X). Although DNA-binding and the transcription regulation domains of these proteins are well characterized, the nuclear localization signals (NLSs) are still unknown in all NF1s. We have identified two NLSs in NF1-A, and both are required for full translocation to the nucleus, although one of them itself has a partial translocation ability. These two NLSs are conserved in all four NF1s. Interestingly, three isoforms of NF1-A (NF1-A1, A2, A4) have two NLSs and translocate completely to the nucleus. In contrast, NF1-A3 lacks the second NLS and partially stays in the cytoplasm. Since NF1s construct homodimer and heterodimer, these findings indicate the differential regulations of the NF1 translocation.

Key words: Nuclear factor 1; Nuclear localization signal; Transcription factor; Translocation; DNA-binding protein

1. Introduction

Nuclear factor 1 (NF1) was originally identified as a DNA-binding protein that not only stimulates the replication of adenovirus DNA in HeLa cells [1], but also stimulates the transcription of many kinds of genes in mammalian cells [2]. NF1 cDNAs were cloned from several species including humans [3] and rats [4], and the sequence comparison of chicken genes led to the identification of four genes: NF1-A, NF1-B, NF1-C, and NF1-X [5]. In addition, each subtype possesses different isoforms generated by alternative splicing [6–8].

All NF1 proteins contain a highly conserved N-terminal amino acid sequence required for dimerization, DNA binding and the activation of adenovirus DNA replication [9–12]. The NF1 recognition sequence and the DNA-binding specificity were also characterized [8,13].

However, the C-terminal half functions as a transcriptional regulation domain and the sequences are less conserved. The activation domains were identified in the C-terminal region of NF1-B, C and X [8,14,15], whereas the repression domain was mapped in the C-terminal region of NF1-A [16]. Thus, although both the DNA-binding/dimerization domain and the transcription regulation domain were relatively well characterized, the nuclear localization signals (NLSs) are still unclear in all four NF1 subtypes. In the present study, we identified two NLSs in NF1-A using green fluorescent protein as a reporter. These two NLSs are required for full translocation to the nucleus, although each NLS has partial translocation ability.

2. Materials and methods

2.1. Plasmid construction for EGFP–NF1-A fusion proteins

pEGFP-C1 (Clontech Lab. Inc., CA, USA) was used as a green fluorescent protein expression plasmid throughout this study. NF1-A1, A2, A3, and A4 cDNAs in pBluescript were as previously described [7]. These plasmids were digested with *Eco*RI and *Sal*I, and the resultant fragments were subcloned into *Eco*RI/*Sal*I sites in pEGFP-C1. Almost all deletion mutants were made by PCR using appropriate primers with the *Eco*RI site and *Sal*I site for 5′ and 3′, respectively. For NLS1 and NLS2 fragments, the synthesized oligonucleotides were annealed. The in vitro site-directed mutagenesis was performed using a QuikChange site-directed mutagenesis kit (Stratagene, CA, USA) according to the manufacturer's instructions. In this case, full-length NF1-A1 including exons 1–11 in pEGFP-C1 was used for site-directed mutagenesis as a template. All sequences were confirmed using the automated DNA sequencers DSQ 1000 (Shimadzu Corp., Kyoto, Japan) and ABI PRISM 310 (Perkin-Elmer, Applied Biosystems Division, CA, USA).

2.2. Cell culture and DNA transfection

HeLa cells were cultured in minimal essential medium supplemented with 10% fetal bovine serum. For fluorescence microscopy, cells were plated onto a cell disc (Sumitomo Bakelite, Co. Ltd., Tokyo, Japan) in 60 mm dishes at a density of 2×10^5 cells/dish 24 h before transfection. Transfection was carried out by the calcium phosphate co-precipitation technique [17], using 4.5 μg of the EGFP–NF1-A fusion construct.

2.3. Fluorescence microscopy and laser scanning confocal microscopy

Forty hours after transfection, cells were washed with PBS three times, and fixed with 4% paraformaldehyde in 3 ml of PBS for 30 min at room temperature. After PBS washing twice, cells were stained with 2.5 μg/ml DAPI in PBS for 3 min at room temperature, and then washed with PBS. Fluorescence microscopy and photography were performed using BX50-34-FLA and PM30 (Olympas, Tokyo, Japan), respectively. The samples were also observed by a confocal microscopy MRC-1024 (Bio-Rad, Lab., CA, USA).

*Corresponding author. Fax: (81)-52-836-3455.
E-mail: imagawa@phar.nagoya-cu.ac.jp

Abbreviations: NF1, nuclear factor 1; NLS, nuclear localization signal; EGFP, enhanced green fluorescent protein; DAPI, 4,6-diamino-2-phenylindole; PCR, polymerase chain reaction; GST, glutathione S-transferase P; aa, amino acids

0014-5793/00/$20.00
PII: S0014-5793(00)02119-0

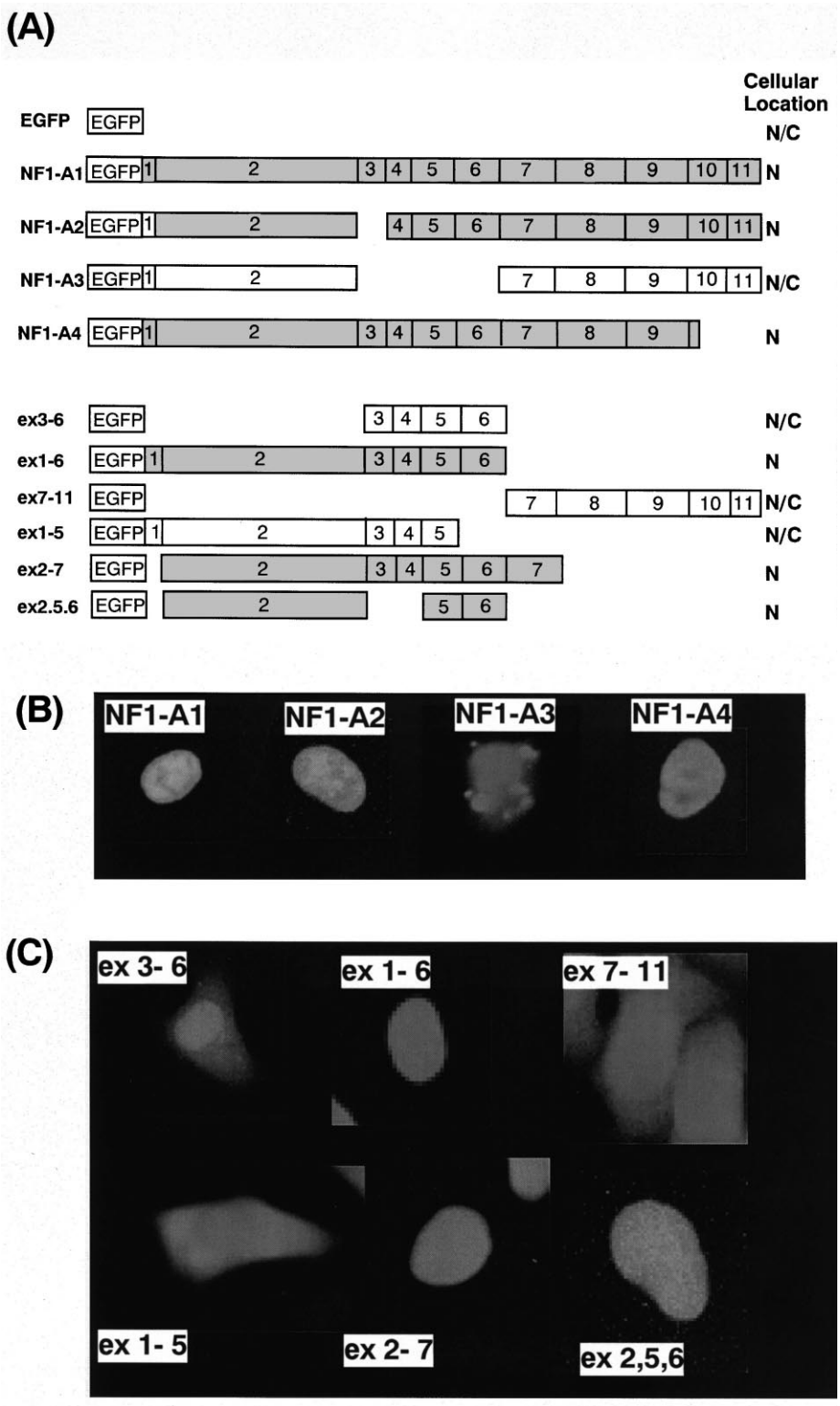

Fig. 1. Subcellular localization of NF1-A isoforms and various deletion mutants fused to EGFP. A: The structure of EGFP–NF1-A constructs with various deletions. The numbers of exons are also shown. NF1-A2 and NF1-A3 are lacking exon 3 and exons 3–6, respectively. NF1-A4 lacks exon 10, leading to a frame shift. B: The subcellular localization of EGFP–NF1-A isoforms transfected into HeLa cells. C: The subcellular localization of various EGFP–NF1-A deletions transfected into HeLa cells. The summary of localizations in nuclear only (N) and both nuclear and cytoplasm (N/C) are indicated in the right panel in A. The constructs with the hatched boxes show nuclear (N) localization alone.

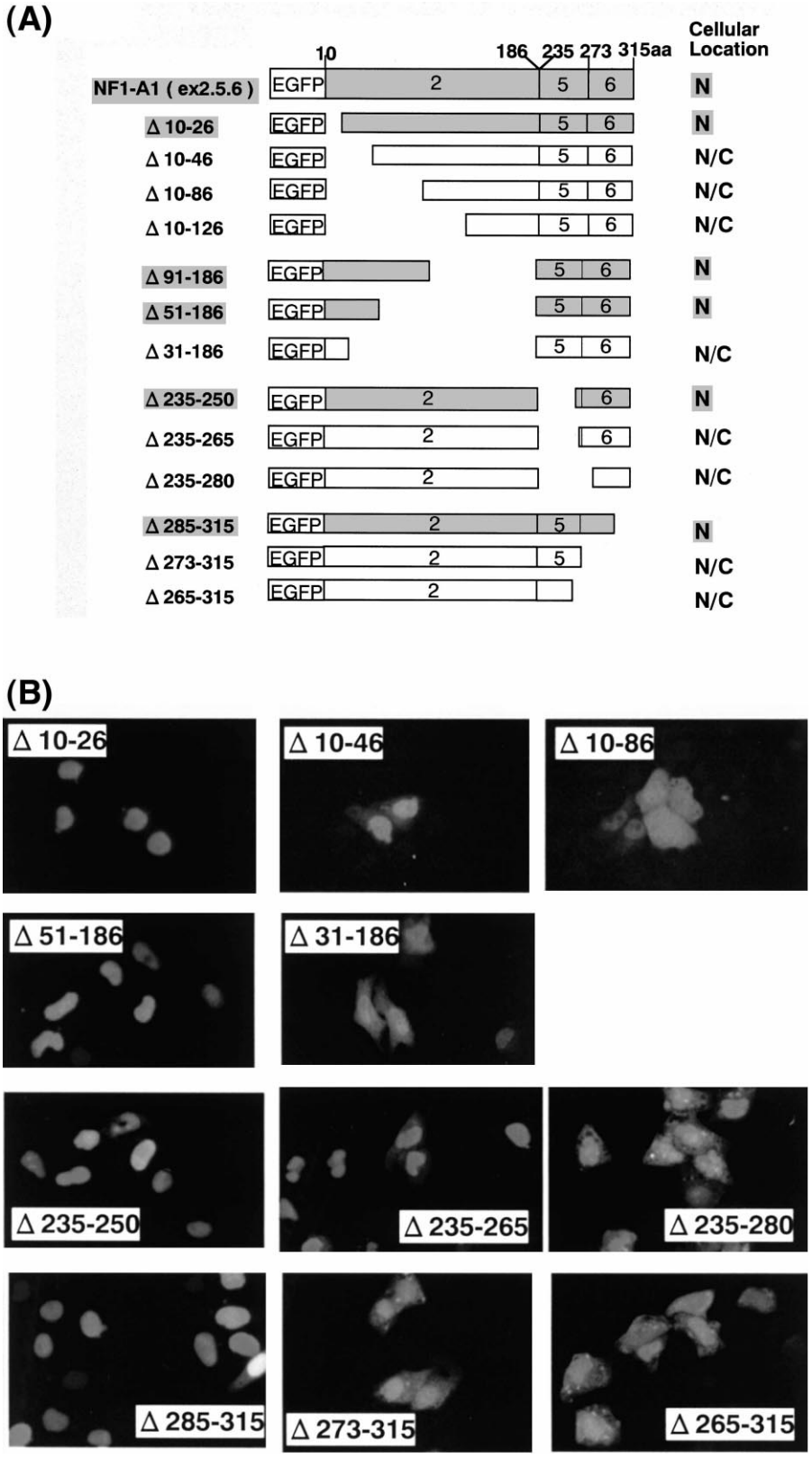

Fig. 2. Subcellular localization of various EGFP–NF1-A1 (exons 2, 5, 6) fusion proteins deleted from the N-terminal or the C-terminal portions. A: The structure of EGFP–NF1-A constructs with various deletions. B: The subcellular localization of various EGFP–NF1-A constructs transfected into HeLa cells. The summary of localizations in nuclear only (N) and both nuclear and cytoplasm (N/C) are indicated in the right panel in A. The constructs with hatched boxes show nuclear (N) localization alone.

3. Results

3.1. NF1-A3 is not fully translocated to nucleus

NF1-A consists of four splicing isoforms [7]. To test the translocation ability of NF1-A, we used green fluorescent protein fused to various NF1-As. Although NF1-A1, A2, and A4 were completely translocated into the nucleus, NF1-A3 was observed both in the nucleus and the cytosol (Fig. 1B). NF1-A2 lacks exon 3, and NF1-A3 lacks exons 3–6. Therefore, we next tested the subcellular localization of exons 3–6. As shown in Fig. 1C, exons 3–6 are not sufficient for the complete translocation, rather exons 2–7 are required. Further deletion analyses indicated that exon 2 and exon 5/6 are required for the full activity (Fig. 1C). It is reported that the proteins up to 60 kDa are able to reach the nucleus by passive diffusion [18]. Therefore, GFP fusion proteins without NLSs seem to distribute in nucleus as well as cytoplasm as reported previously [19,20].

3.2. Two separated regions, of exons 2 and 5/6 are required for sufficient translocation

To map the precise NLSs, we constructed the EGFP gene fused to exons 2, 5 and 6 of NF1-A1, and the various deletion mutants from the 5′ and 3′ ends were also created. By comparison with the findings shown in Fig. 2, the minimum region of exon 2 for full translocation was identified as 24 aa from the 27th to the 50th aa of NF1-A1. The minimum region of exon 5/6 was also determined as 34 aa from the 251st to the 284th aa of NF1-A1, which includes the border of exons 5 and 6. Taken together, these findings indicate that the two separated regions, of exon 2 and exon 5/6 are essential for the nuclear translocation of NF1-A. These two sequences are shown in Fig. 3A. Since both sequences are rich in basic amino acids, the regions rich in basic amino acids were tentatively named NLS1 (15 aa from the 36th to the 50th aa of NF1-A1) and NLS2 (17 aa from the 263rd to the 279th aa of NF1-A1) for exon 2 and exon 5/6, respectively. Interestingly, NLS1 contained a single basic type NLS, and NLS2 contained a bipartite type NLS [21]. In general, one basic type or one bipartite type NLS is sufficient for the translocation of the transcription factors [21]. Therefore, we next constructed NLS1, NLS2, and NLS1/2 chimeric genes fused to EGFP. The NLS1/2 contains additional two amino acids (Val-Asp) between NLS1 and NLS2 due to the *Sal*I site inserted. As shown in Fig. 3B, each NLS itself is not sufficient, and both NLSs are required for full translocation (Fig. 3B). A similar case was reported for SRY and SOX9, the members of the family of the high-mobility group (HMG) domain transcription factors [22]. Although the reason for this difference is not known, it is possible that the conformations of transcription factors are different and both NLSs are needed for the nuclear pore-targeting complex formation [21]. Interestingly, when we used only 6 aa of RRKRLK (termed NLS2-6aa), in which RR and KRLK were 2 aa from N-terminus and 4 aa from C-terminus of NLS2, respectively, instead of NLS2, NLS1/NLS2-6aa and NLS2-6aa gave similar results indicating that

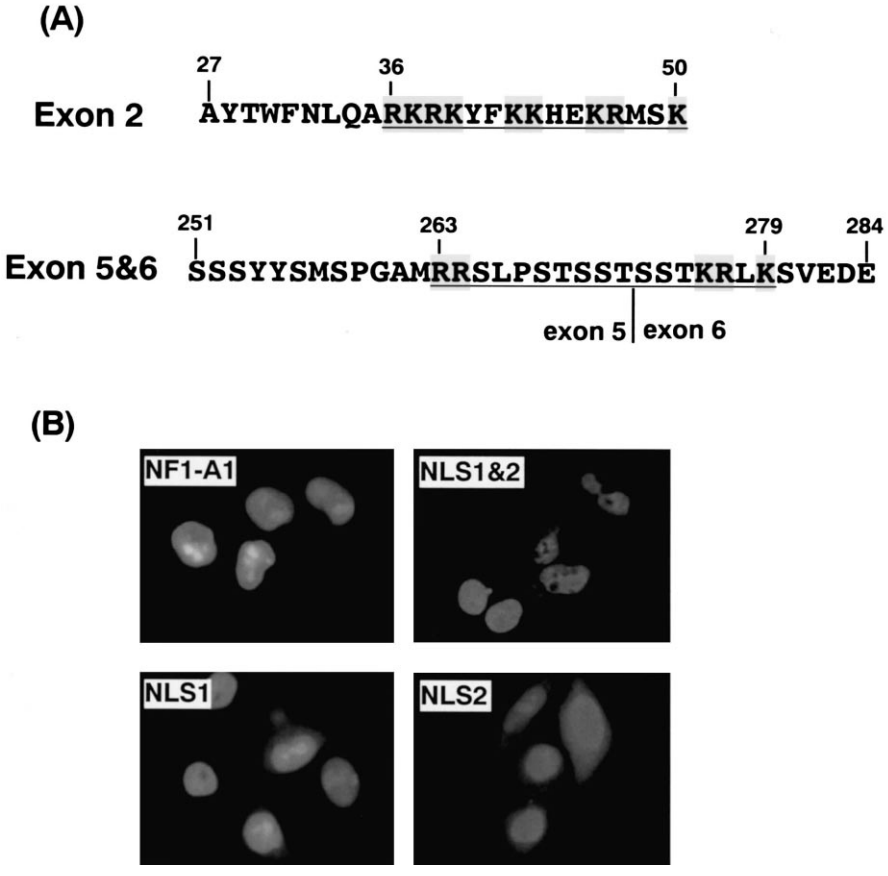

Fig. 3. Two NLSs mapped in NF1-A. A: The sequences of NLSs identified in NF1-A1. The lysine and arginine residues are boxed. The tentative NLS1 in exon 2 and NLS2 in exon 5/6 are underlined. B: The subcellular localization of NLS1 and/or NLS2 with EGFP constructs transfected into HeLa cells.

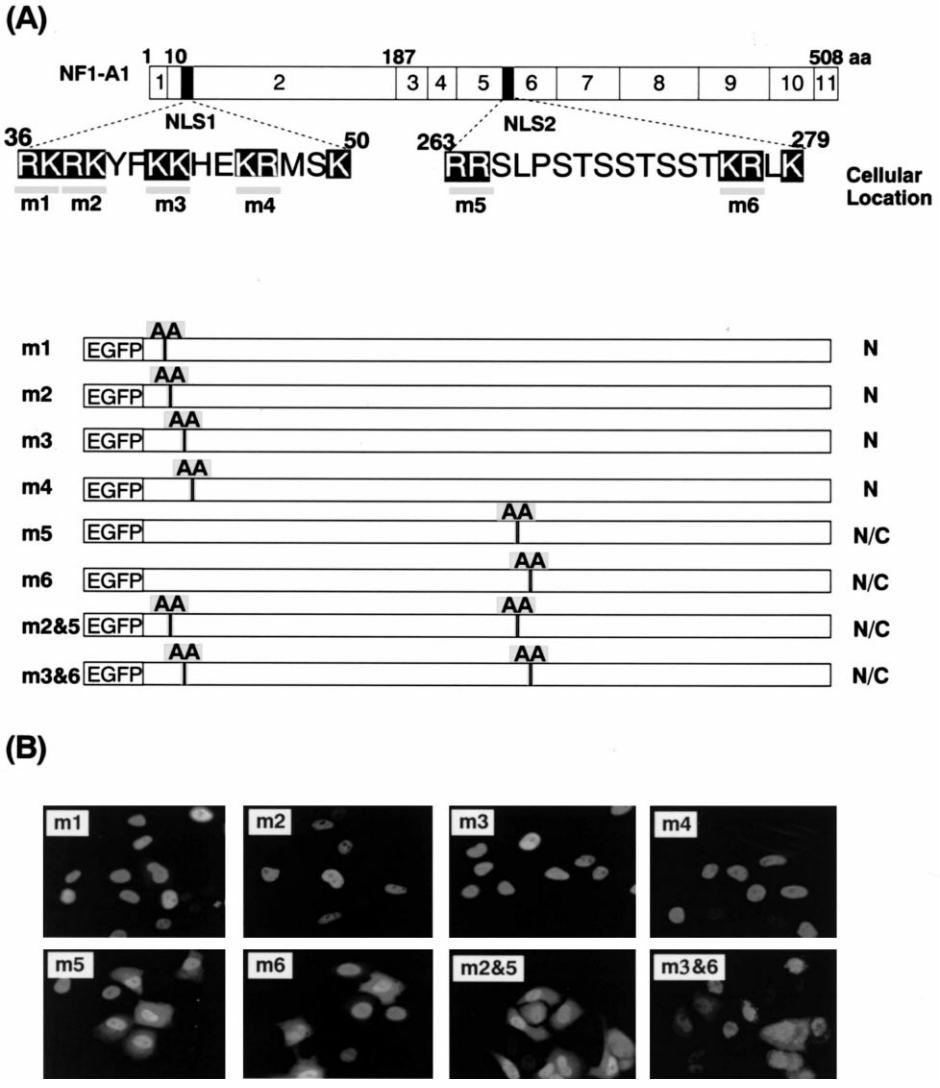

Fig. 4. Subcellular localization of double point mutations and combined double point mutations in NLS1 and NLS2 with EGFP full-length NF1-A1 constructs transfected into HeLa cells. A: The map of NF1-A1 and NLS1 and NLS2. Two basic amino acids were substituted with alanines in each construct. B: The subcellular localization of various EGFP–NF1-A constructs transfected into HeLa cells. The summary of localizations in nuclear only (N) and both nuclear and cytoplasm (N/C) are indicated in the right panel in A.

these basic amino acids have critical roles for nuclear translocation (data not shown).

3.3. *Identification of critical basic amino acids for translocation in NLS1 and NLS2*

In Fig. 3, only a small portion of the basic amino acid rich regions in NLS1 and NLS2 was used for the experiment. To identify the critical basic amino acids for nuclear translocation of NF1-A, we next generated several double point mutations in NLS1 and NLS2 in full-length NF1-A1. In these experiments, leucine and arginine residues were substituted with alanine residues. As shown in Fig. 4, the double point muta-

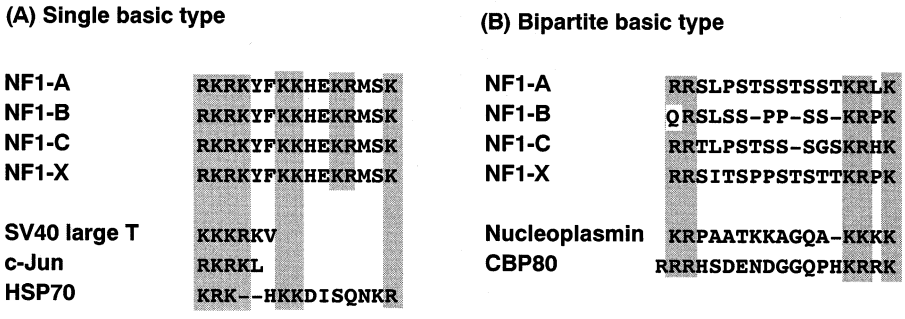

Fig. 5. Sequence comparison of two NLSs mapped near the DNA-binding domain in the NF1 family and of NLSs identified in various genes. A: Single basic type. B: Bipartite basic type. Lysine and arginine residues are enclosed by dark shading.

tions in NLS1 did not affect the translocation ability, although the partial translocation efficiency of NLS1 itself was higher than that of NLS2. However, the mutations in NLS2, both double point mutations, m5 and m6, affected the translocation efficiency, indicating that the basic amino acids in NLS2 are relatively important in cooperation with NLS1. Next, we performed combined mutations in NLS1 and NLS2. Interestingly, the combined mutants of m2 and m5, m2 and m6, m1 and m5 and m1 and m6 showed less translocation to the nucleus compared with m5 or m6. The mutants of m3 and m6 also gave less translocation to the nucleus compared with m6, whereas the mutants of m3 and m5, m4 and m5 and m4 and m6 showed the same results as m5 or m6 (Fig. 4 and data not shown).

4. Discussion

NF1s constitute a family consisting of at least four genes, and function as activators and repressors as well as stimulators for replication [1]. Whereas the DNA-binding domain and the regulation domain for the transcription are relatively well characterized, the NLS is not yet characterized for all four NF1s. The process of nuclear translocation mediated by NLS is divided into at least two steps; targeting to the pores and translocation through the pores [21]. The first step involves the formation of a stable complex of nuclear protein through NLS, termed the nuclear pore-targeting complex, in cytoplasm. There is no single and strict consensus NLS, but there are some general rules for NLSs [23]. These include: (1) they are typically short sequences, usually not more than 8–10 amino acids; (2) they contain a high proportion of positively charged amino acids (lysine and arginine); (3) they are not located at specific sites within the protein; (4) they are not removed following localization; and (5) they can occur more than once in a given protein [23]. Thus, the short sequences rich in basic amino acids seem to be candidates for NLS.

We previously cloned rat NF1-A, B, C, and X and the amino acid sequence alignment revealed that there was high similarity in the DNA-binding domain in the N-terminal portion, especially in the lysine helix region, whereas the transcription regulation domain in the C-terminal portion were less conserved [8]. Although none of the NLSs in all four NF1s were identified, many lysine and arginine residues were found and conserved in the four NF1s (Fig. 5). To identify the NLSs, GFP fusion protein is widely used [18–20]. Therefore, we constructed various kinds of EGFP–NF1-A chimeric genes and mapped the NLSs. Two NLSs in NF1-A were identified in the present study. Interestingly, these two NLSs, NLS1 and NLS2, were similar to single basic type NLS and bipartite basic type NLS, respectively, both NLS1 and NLS2 were required for the full translocation of NF1-A to the nucleus. NLS1 and NLS2 were located near the DNA-binding domain of NF1-A, but these two NLSs were separated by conserved cysteine residues which were important amino acid residues for DNA binding, and were separately by more than 200 amino acids. In general, either the single basic type NLS or bipartite basic type is sufficient for nuclear targeting [23]. For example, the PKKKRKV sequence in SV40 large T antigen was defined as minimal NLS [24]. Two independent NLSs have been identified in some genes, including polyoma large T antigen containing two basic sequences (VSRKRPRPA and PKKARED) [25]. In this case, these two NLSs cooperate to mediate nuclear translocation, and the mutation of either one individually impairs but does not eliminate the ability for nuclear entry [23,25]. In the DNA-binding high-mobility group (HMG) domains of SRY and SOX9, two NLSs were also present, and both are required for complete nuclear entry [22]. These two NLSs are conserved in other HMG domain proteins, in which the bipartite NLS is found in the N-terminal portion and the single basic NLS is localized in the C-terminal portion; the opposite orientation for NF1-A1 [22]. The relationship between NLS1 and NLS2 in NF1-A is still unclear. It is possible that these are close three-dimensionally. The three-dimensional structural analyses may solve this question.

A protein may have two signals, because the interaction with other proteins regulates the nuclear localization ability of transcription factors. Interestingly, since NF1-A3 lacks exons 3–6 [7], this protein has only one NLS, and does not import completely to the nucleus (Fig. 1B). The expression levels of NF1-A3 are relatively low compared with NF1-A1 and NF1-A4 [7]. However, NF1-A itself makes a heterodimer within isoforms as well as a homodimer [11]. Therefore, it is possible that NF1-A3 makes a heterodimer with other isoforms and inhibits the translocation of isoforms to the nucleus. This possibility and the comparison of expression levels of NF1 isoforms remain to be investigated.

Acknowledgements: This work was supported in part by grants from the Ministry of Education, Science, Sports and Culture, Japan, and from the Nissan Science Foundation.

References

[1] Nagata, K., Guggenheimer, R.A., Emoto, T., Lichy, J.H. and Hurwitz, J. (1982) Proc. Natl. Acad. Sci. USA 79, 6438–6442.
[2] Jones, K.A., Kadonaga, J.T., Rosenfeld, P.J., Kelly, T.J. and Tjian, R. (1987) Cell 48, 79–89.
[3] Santoro, C., Mermod, N., Andrews, P.C. and Tjian, R. (1988) Nature 334, 218–224.
[4] Paonessa, G., Gounari, F., Frank, R. and Cortese, R. (1988) EMBO J. 7, 3115–3123.
[5] Rupp, R.A.W., Kruse, U., Multhaup, G., Göbel, U., Beyreuther, K. and Sippel, A.E. (1990) Nucleic Acids Res. 18, 2607–2616.
[6] Kruse, U. and Sippel, A.E. (1994) J. Mol. Biol. 238, 860–865.
[7] Osada, S., Daimon, S., Ikeda, T., Nishihara, T., Yano, K., Yamasaki, M. and Imagawa, M. (1997) J. Biochem. 121, 355–363.
[8] Osada, S., Matsubara, T., Daimon, S., Terazu, Y., Xu, M., Nishihara, T. and Imagawa, M. (1999) Biochem. J. 342, 189–198.
[9] Gounari, F., De Francesco, R., Schmitt, J., van der Vliet, P.C., Cortese, R. and Stunnenberg, H. (1990) EMBO J. 9, 559–566.
[10] Novak, A., Goyal, N. and Gronostajski, R.M. (1992) J. Biol. Chem. 267, 12986–12990.
[11] Kruse, U. and Sippel, A.E. (1994) FEBS Lett. 348, 46–50.
[12] Dekker, J., van Oosterhout, J.A.W.M. and van der Vliet, P.C. (1996) Mol. Cell. Biol. 16, 4073–4080.
[13] Osada, S., Daimon, S., Nishihara, T. and Imagawa, M. (1996) FEBS Lett. 390, 44–46.
[14] Mermod, N., O'Neil, E.A., Kelly, T.J. and Tjian, R. (1989) Cell 58, 741–753.
[15] Roulet, E., Amentero, M.T., Krey, G., Corthesy, B., Dreyer, C., Mermod, N. and Wahli, W. (1995) Mol. Cell. Biol. 15, 5552–5562.
[16] Osada, S., Ikeda, T., Xu, M., Nishihara, T. and Imagawa, M. (1997) Biochem. Biophys. Res. Commun. 238, 744–747.
[17] Chen, C. and Okayama, H. (1987) Mol. Cell. Biol. 7, 2745–2752.
[18] Mattaj, I.W. and Engelmeier, L. (1998) Annu. Rev. Biochem. 67, 265–306.
[19] Hessabi, B., Ziegler, P., Schmidt, I., Hessabi, C. and Walther, R. (1999) Eur. J. Biochem. 263, 170–177.

[20] Moede, T., Leibiger, B., Pour, H.G., Berggren, P. and Leibiger, I.B. (1999) FEBS Lett. 461, 229–234.
[21] Yoneda, Y. (1997) J. Biochem. 121, 811–817.
[22] Sudbeck, P. and Scherer, G. (1997) J. Biol. Chem. 272, 27848–27852.
[23] Garcia-Bustos, J., Heitman, J. and Hal, M.N. (1991) Biochim. Biophys. Acta 1071, 83–101.
[24] Kalderon, D., Roberts, B.L., Richardson, W.D. and Smith, A.E. (1984) Cell 39, 499–509.
[25] Richarson, W.D., Roberts, B.L. and Smith, A.E. (1986) Cell 44, 77–85.