Characterization of nucleolar localization and exclusion signals in terminal deoxynucleotidyltransferase interacting factor 2/estrogen receptor α-binding protein

doi: 10.1080/09168451.2019.1591265(https://doi.org/10.1080/09168451.2019.1591265)

Characterization of nucleolar localization and exclusion signals in terminal deoxynucleotidyltransferase interacting factor 2/estrogen receptor α-binding protein

Takasi Fukada[a,*,†], Shun Shibata[a,*], Toshihiro Ueda[a], Katsuhiko Sasaki[a], Yukiko Shimoida[a], Kaori Senda-Murata[a,b], and Kenji Sugimoto[a,b]

[a]*Graduate School of Life and Environmental Sciences, Osaka Prefecture University, Sakai, Osaka, Japan*

[b]*Live Cell Imaging Institute, Osaka Prefecture University, Sakai, Osaka, Japan*

[*]These authors equally contributed to this study.

[†]Corresponding author. Email: tfukada@bioinfo.osakafu-u.ac.jp

Characterization of nucleolar localization and exclusion signals in terminal deoxynucleotidyltransferase interacting factor 2/estrogen receptor α-binding protein

Abstract

Terminal deoxynucleotidyltransferase interacting factor 2/estrogen receptor α-binding protein (TdIF2/ERBP) is a multifunctional nucleolar protein. The nucleolar localization of TdIF2/ERBP is important for its functions because it promotes ribosomal RNA transcription. However, signal sequences that direct TdIF2/ERBP to the nucleolus are not well characterized. We examined the TdIF2/ERBP sequence using truncation and mutation analyses to determine whether the nucleosome binding and C-terminal domains of TdIF2/ERBP possess nucleolar localization signals (NoLSs). In these domains, four NoLSs that could direct the mCherry protein to the nucleolus were detected. In addition, a short stretch of hydrophobic residues (VLLVL) in the center of TdIF2/ERBP acted as a nucleolar exclusion signal, which reduced the nucleolar accumulation of mCherry–NoLS fusion proteins. These results would contribute to improving the prediction of NoLSs from protein sequences. The short, transferrable localization signals would be valuable tools for understanding the association between localization and functions of nucleolar proteins.

Abbreviations: TdIF2: terminal deoxynucleotidyltransferase interacting factor 2; ERBP: estrogen receptor α-binding protein; EGFP: enhanced green fluorescent protein; NLS: nuclear localization signal; NoLS: nucleolar localization signal; NoES: nucleolar exclusion signal; DAPI: 4′,6-diamidino-2-phenylindole

Keywords: TdIF2; ERBP; nuclear localization signal; nucleolar localization signal; nucleolar exclusion signal

Introduction

The nucleolus, an intranuclear structure lacking a surrounding membrane, is the site of ribosome assembly and is involved in various cellular processes, such as ribonucleoprotein assembly, cell proliferation, and stress response [1,2]. A variety of proteins localize in the nucleolus, many of which dynamically shuttle between the nucleolus and nucleoplasm or even the cytoplasm [3]. Because the nucleolus is not a membrane-bound structure, these proteins are considered to accumulate in it via interactions with nucleolar core components, such as ribosomal DNA; however, detailed mechanisms underlying their retention remain unclear.

Several nucleolar proteins have clusters of basic amino acid residues—nucleolar localization signals (NoLSs) [3]. These sequences are essential for the nucleolar localization of proteins, and if the sequences are transferred to unrelated reporter proteins, they often direct the reporter proteins to the nucleolus. NoLSs resemble nuclear localization signals (NLSs), and the same sequence often acts as an NLS and as an NoLS. Scott *et al.* have suggested that it would be useful to divide these localization signals into three classes: NLS-only, NoLS-only, and joint NoLS–NLS signals [4]. Although they have reported an artificial neural network (ANN) that can predict NoLSs, distinguishing between the three classes remains challenging, with the experimental confirmation of their identity still being necessary.

Terminal deoxynucleotidyltransferase interacting factor 2 (TdIF2), also known as estrogen receptor α-binding protein (ERBP), is a nucleolar protein that binds to and inhibits terminal deoxynucleotidyltransferase (TdT) [5,6]. TdIF2/ERBP

reportedly promotes ribosomal RNA transcription [7]; therefore, the nucleolar localization of TdIF2/ERBP is important for its functions. Fujita *et al.* have annotated residues 631–635 of TdIF2/ERBP as an NLS [5]. Moreover, Koiwai *et al.* have reported that the N- (1–234) and C-terminal (606–756) fragments of TdIF2/ERBP are sorted to the nucleolus [7]; however, they have not annotated the exact sequences responsible for the nucleolar localization of these fragments.

In the present study, we examined the TdIF2/ERBP sequence in detail to identify and characterize its NoLSs. We found that the previously assigned NLS (631–635) is not necessary for the nucleolar localization of the C-terminal fragment (631–756). Instead, a short stretch of basic amino acid residues at the C-terminus (750–754) serves as an NoLS. Two similar clusters of basic residues in the nucleosome binding domain (NBD) were responsible for the nucleolar localization of the N-terminal fragment (1–234), and a short cluster of hydrophobic residues (441–445) antagonized these NoLSs. These results would contribute to improving the prediction of NoLSs from protein sequences. Moreover, these short, transferrable NoLSs and nucleolar exclusion signals (NoESs) would serve as valuable tools for controlling the localization of nucleolar proteins and examining the association between their function and localization.

Materials and methods

Plasmid construction

All oligonucleotides used in the present study are summarized in Table 1. The cDNA fragment of mCherry was amplified by PCR using a set of primers, mCherry-Forward and mCherry-Reverse, and pRSET-B-mCherry (provided by Dr. Roger Tsien, Univ. California, San Diego) as a template. The mammalian expression vector pmCherry-C1

was constructed by inserting the *Nhe*I-*Bgl*II-digested PCR products into the corresponding sites of the vector backbone of pEGFP-C1 (Clontech, Palo Alto, CA, USA). The TdIF2/ERBP cDNA was cloned from a cDNA library derived from estrogen-treated MCF7 cells (OriGene Technologies, Rockville, MD, USA). mCherry-TdIF2 was constructed by inserting a 2.3-kb fragment of the human TdIF2/ERBP cDNA into the *Hind*III-*Sal*I sites of pmCherry-C1. The truncated fragments and point mutants of TdIF2/ERBP were generated using PCR and cloned into pEGFP-C1. Individual nuclear localization/exclusion signals, namely NLS2 (RVTRRR), NLS3 (RKKPK), NLS4 (KKRRK), NLS6 (RKKKK), and NoES (VLLVL), were chemically synthesized and annealed (Table 1) and inserted into the *Bgl* II, *Xho* I, *Hind* III, *Sal* I, or *Bam*H I cutting sites of pmCherry-C1 without additional linkers.

Cell culture and transfection

C3H10T1/2 cells were grown in Dulbecco's modified minimum essential medium containing 10% fetal bovine serum. Plasmid DNA was transfected into the cells using Lipofectamine LTX and Opti-MEM I (Life Technologies Japan, Tokyo, Japan) in accordance with the manufacturer's instructions. The transfected cells were incubated at 37°C under 5% CO_2 for 24 h before fluorescence observation.

Fluorescence microscopy

Cells were fixed using 4% paraformaldehyde in phosphate-buffered saline (PBS) for 10 min and permeabilized using 0.1% Triton X-100 in PBS for 5 min. The cells were briefly stained using 1-μg/mL DAPI in PBS, following which they were observed using an Eclipse E600 fluorescence microscope (Nikon, Tokyo, Japan) and ORCA-ER CCD-camera (Hamamatsu Photonics, Shizuoka, Japan). Fluorescence intensity was measured

by ImageJ software (National Institutes of Health, Bethesda, MD, USA).

Table 1

Results

Previously assigned NLS domain is not necessary for the nucleolar localization of the C-terminal fragment of TdIF2/ERBP

TdIF2/ERBP has been reported to contain four functional domains: NBD, acidic region (AR), DNA-binding domain (DBD), and NLS domain. Because Koiwai *et al.* have reported that either the N- (1–234) or C-terminal (606–756) fragment of TdIF2/ERBP was localized in the nucleolus, there should be multiple NoLSs in the protein. Analysis of the EGFP-fused deletion mutants revealed that the C-terminal fragment (631–756) was localized in the nucleolus (Figure 1). Survey of the sequence of this region (631–756) revealed two NLS-like clusters of basic residues: 631-KKRRK-635 and 750-RKKKK-754. The NLS domain described by Fujita *et al*. included residues 631–635, but not 750–754 [5].

Figure 1

To examine the role of these clusters in the nucleolar localization of TdIF2/ERBP, we replaced two basic residues in each of the clusters with alanine residues. The full-length double mutant EGFP-FL K632A/K633A/K751A/K752A (Figure 2(a)) was predominantly localized to the nucleoplasm as well as was accumulated in the nucleolus in a significant proportion of cells (Figure 2(b)), suggesting that there are additional NLSs and NoLSs in TdIF2/ERBP. To avoid the effect of the N-terminal fragment, we created EGFP-fused C-terminal fragments (631–756) with and without the NLS mutations (Figure 2(a)). The EGFP-631–756/K632A/K633A fragment was localized in the nucleolus, suggesting that the NLS at 631–635 is not indispensable for the nucleolar localization of this fragment (Figure 2(b)). The EGFP-631–756/K751A/K752A and 631–756/K632A/K633A/K751A/K752A

Figure 2

fragments showed little accumulation in the nucleolus (Figure 2(b)). These results suggested that the sequence 750–754 has NoLS activity.

NBD is involved in the nucleolar localization of TdIF2/ERBP but is antagonized by a nucleolar exclusion signal (NoES)

We created a series of deletion mutants of the N-terminal fragment (1–505) of TdIF2/ERBP to identify NoLSs in this region (Figure 3). The EGFP-fused fragment 1–100 was localized in the nucleoplasm, but the fragment 1–133 was accumulated in the nucleolus, thereby suggesting that NBD is involved in nucleolar localization. Because the fragment 134–440 was localized in the cytoplasm, the region 1–100 should have an NLS-only signal(s). Extending the C-terminus of the deletion mutants beyond residue 440 reduced the nucleolar localization of the fragments. The minimum sequence associated with this antagonizing effect was 441-VLLVL-445, and we named this hydrophobic sequence the nucleolar exclusion signal (NoES).

Figure 3

TdIF2/ERBP includes at least four NoLSs that direct mCherry protein to the nucleolus, and NoES partially compromises NoLS activity

To identify more NoLSs, we scanned the full-length TdIF2/ERBP sequence for NLS by PSORT II [8] because most of the reported NoLSs were rather similar or identical to NLS consensus sequences. We found six putative NLSs in TdIF2/ERBP (NLS1–6; Figure 4(a)). NLS4 was a part of the NLS domain reported by Fujita *et al.* [5]. Moreover, we scanned the sequence using NoD ANN developed by Scott *et al.* to predict NoLSs [4]. NoD predicted four NoLSs, including NLS1, NLS4, NLS5, and NLS6, in the TdIF2/ERBP sequence.

Because NLS1 is included in the fragment 1–100 that was localized in the

nucleoplasm (Figure 3), NLS1 is possibly an NLS-only signal rather than an NoLS. NLS5 resided in the fragment 631–756that was localized in the nucleoplasm when NLS4 and 6 were mutated (Figure 2), suggesting that NLS5 is an NLS-only signal. On this basis, we fused the mCherry gene to synthetic DNA encoding NLS2, NLS3, NLS4, NLS6, and NoES, and transiently expressed them in C3H10T1/2 cells to observe the localization of the fusion protein. Furthermore, the fragment 101–134 of TdIF2/ERBP was used to observe the effect of NLS2, NLS3, and NoES in the endogenous context. For quantitative analysis, we measured the average fluorescence intensity in the nucleolus and nucleoplasm and used the ratio (nucleolus/nucleoplasm) as a nucleolar localization index.

Figure 4

All tested NLSs directed mCherry protein to the nucleolus (Figure 4), suggesting that these sequences are joint NoLS–NLSs that allow the protein to enter the nucleus and induce accumulation in the nucleolus. On fusing two NLSs to mCherry, they synergistically enhanced the nucleolar localization of the fusion protein. Although the addition of NoES reduced this nucleolar localization (Figure 5), most of the fusion protein was retained in the nucleus (Figure 4(b)).

Figure 5

Discussion

In the present study, we examined the TdIF2/ERBP protein sequence and found four NoLSs (NLS2, NLS3, NLS4, and NLS6) and one NoES. Our findings demonstrated that NLS1 and NLS5 are possibly NLS-only signals rather than NoLSs. The relative activity of NoLSs to direct mCherry protein to the nucleolus appeared to be in the order NLS3 and NLS6 > NLS2 and NLS4, as measured by the fluorescence intensity ratio (nucleolus/nucleoplasm). Notably, NLS4, which was the only NLS annotated on TdIF2/ERBP by Fujita *et al*. [5], was not necessary for the nucleolar localization of the

C-terminal fragment (631–756) provided that NLS6 was intact, suggesting that NLS4 is a relatively weak NoLS.

No simple consensus sequence is known for NoLS, but Scott *et al*. have reported that an ANN can predict NoLSs from protein sequences [4]. Their ANN predicted four putative NoLSs—NLS1, NLS4, NLS5, and NLS6—in the TdIF2/ERBP sequence. However, our results suggested that NLS1 and NLS5 are possibly NLS-only signals, whereas NLS2 and NLS3 are NoLS. These discrepancies indicate that distinguishing between NoLS and NLS-only signals without experimental analysis remains challenging.

Martin *et al*. have reported that arginine hexamer peptide, but not lysine hexamer, was localized to the nucleolus. They found that if another NLS existed nearby, four arginine residues were sufficient for the nucleolar localization of GFP [9]. They have concluded that the high isoelectric point (pI > 12.6) and the distribution pattern of the positive charge on the molecular surface of the NLS domain were important to distinguish between NoLS and NLS-only signals. Although the cumulative effect of NoLSs (Figure 4) would support their conclusion, it is not the sole principle of the nucleolar localization of protein because the individual NoLS of TdIF2/ERBP could localize mCherry protein to the nucleolus without extra NLS. Further research is required to resolve this issue.

In the center of the TdIF2/ERBP protein, we found an NoES that reduced the nucleolar accumulation of the protein. Unlike the nuclear export signal (NES), NoES showed little effect on the nucleoplasmic localization of TdIF2/SRBP. Meng *et al*. have reported similar signals in two nucleolar proteins, GNL3L and Ngp1, and named them nucleoplasmic localization signals (NpLSs) [10]. However, they assigned NpLS activity to relatively large domains (285 and 272 residues for GNL3L and Ngp1, respectively),

whereas the minimum sequence of NoES was only five residues. We found no NoES-like sequence (*i.e.*, five successive hydrophobic residues) in the reported NpLSs. These findings suggest that NoES and NpLS are functionally similar but structurally different signals. The existence of functionally similar signals in multiple nucleolar proteins suggests the importance of these NoLS-antagonizing signals in regulating the localization and function of nucleolar proteins.

The leucine-rich feature of the NoES may suggest a relationship to the leucine-rich NESs. Leucine-rich NESs are well characterized and mostly comprise one of the three consensus sequences: [Φ-X2,3-Φ-X2,3-Φ-X-Φ] (class 1), [Φ-X-Φ-X2-Φ-X-Φ] (class 2), and [Φ-X2-Φ-X3-Φ-X2-Φ] (class 3), where Φ represents large hydrophobic residues and X2,3 any two or three amino acids [11]. However, the NoES sequence (VLLVL) does not conform to any of these consensus sequence classes because it does not have sufficient spacing between the hydrophobic residues. Moreover, an NESmapper [11] scan of the TdIF2/ERBP sequence failed to identify the NoES as a NES. These facts suggest that the NoES is not a variant of leucine-rich NES, although we cannot completely dismiss the possibility that the NoES works as a weak NES.

In summary, we found four NoLSs and one NoES in TdIF2/ERBP. Our findings will contribute to improving the prediction of NoLSs from protein sequences. Because the NoES is only five residues long, it could easily be inserted into other nucleolar proteins with the expectation of reducing their accumulation in the nucleolus. This method would be valuable for studying the association between the localization and function of nucleolar proteins.

Author contribution

Kenji Sugimoto designed the research. Shun Shibata, Toshihiro Ueda, Katsuhiko Sasaki, Yukiko Shimoida, and Kaori Senda-Murata performed the experiments and analyzed data. Takshi Fukada analyzed data and wrote the manuscript.

References

[1] Martindill DMJ, Riley PR. Cell cycle switch to endocycle: the nucleolus lends a hand. Cell Cycle. 2008; 7: 17-23.

[2] Vlatkovic N, Boyd MT, Rubbi CP. Nucleolar control of p53: a cellular Achilles' heel and a target for cancer therapy. Cell Mol Life Sci. 2014; 71: 771-791.

[3] Emmott E, Hiscox JA. Nucleolar targeting: the hub of the matter. EMBO Rep. 2009; 10: 231-238

[4] Scott MS, Boisvert FM, McDowall MD, Lamond AI, Barton GJ. Characterization and prediction of protein nucleolar localization sequences. Nucleic Acids Res. 2010; 38: 7388-7399.

[5] Fujita K, Shimazaki N, Ohta Y, et al. Terminal deoxynucleotidyltransferase forms a ternary complex with a novel chromatin remodeling protein with 82 kDa and core histone. Genes Cells. 2003; 8: 559-571.

[6] Bu H, Kashireddy P, Chang J, et al. ERBP, a novel estrogen receptor binding protein enhancing the activity of estrogen receptor. Biochem Biophys Res Commun. 2004; 317: 54-59.

[7] Koiwai K, Noma S, Takahashi Y, et al. TdIF2 is a nucleolar protein that promotes rRNA gene promoter activity. Genes Cells. 2011; 16: 748-764.

[8] Nakai K, Horton P. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem Sci. 1999; 24: 34-36.

[9] Martin RM, et al. Principles of protein targeting to the nucleolus. Nucleus. 2015; 6: 314-325.

[10] Meng L, Zhu Q, Tsai RYL. Nucleolar Trafficking of Nucleostemin Family Proteins: Common versus Protein-Specific Mechanisms. Mol Cell Biol. 2007; 27: 8670-8682.

[11] Kosugi S, et al. NESmapper: Accurate prediction of leucine-rich nuclear export signals using activity-based profiles. PLOS Computational Biol. 2014; 10: e1003841.

Figure legends

Figure 1. Subcellular localization of the C-terminal fragments of TdIF2/ERBP. (a) Schematic diagram of truncated fragments of TdIF2/ERBP. NBD: nucleosome binding domain; AR: acidic region; DBD: DNA-binding domain; NLS: nuclear localization signal. (b) The truncated fragments of TdIF2/ERBP were fused with EGFP, transiently expressed in C3H10T1/2 cells, and observed using fluorescence microscopy. We observed at least 50 cells for each fragment and the localization pattern in all the cells was similar to those shown.

Figure 2. Subcellular localization of TdIF2/ERBP fragments with NLS mutations. (a) Schematic diagram of potential NLSs in the C-terminal domain of TdIF2/ERBP and mutant fragments. Solid boxes: wildtype NLS; open boxes: mutated NLS. (b) Wildtype or mutant fragments of TdIF2/ERBP were fused with EGFP, transiently expressed in C3H10T1/2 cells, and observed using fluorescence microscopy. We observed at least 50 cells for each fragment and the localization pattern in all the cells was similar to those shown. However, because the full-length double mutant (full-length K632A/K633A/K751A/K752A) showed two different localization patterns (nucleolus or nucleoplasm), the proportion of cells showing each pattern is indicated.

Figure 3. Subcellular localization of the N-terminal fragments of TdIF2/ERBP. (a) Schematic diagram of truncated fragments of TdIF2/ERBP and the proportions of cells that showed the nucleolar, nucleoplasmic, and cytoplasmic localization of the EGFP-fused fragments in C3H10T1/2 cells. No: nucleolus; Nu: nucleoplasm; Cyt: cytoplasm. (b) Representative cell images showing EGFP-fused N-terminal fragments of

TdIF2/ERBP.

Figure 4. Effects of the putative NLSs and NoES on mCherry protein localization. (a) Locations and sequences of the putative NLSs and NoES in TdIF2/ERBP. Basic residues are underlined. (b) Schematic drawing showing the structures of mCherry-fused NLSs and NoES. mCherry-fused TdIF2/ERBP fragment (101–134), which includes NLS2 and NLS3, is shown. The numbers indicate the proportions of cells that showed the nucleolar localization of the mCherry fusion proteins, and the ratios of average fluorescence intensities (nucleolus/nucleoplasm) in C3H10T1/2 cells. Moreover, representative cell images indicating the localization of mCherry fusion proteins are shown. For each construct, we observed at least 50 cells for the cell numbers (%), and analyzed 10 cells for fluorescence intensity.

Figure 5. Antagonistic activity of the NoES to the nucleolar accumulation of mCherry–NoLS fusion proteins. Ratios of average fluorescence intensities (nucleolus/nucleoplasm) in C3H10T1/2 cells are indicated for the fusion proteins with (blank bars) NoES or without (solid bars) NoES. Error bars: standard deviation (n = 10).