The conserved basic residues and the charged amino acid residues at the α-helix of the zinc finger motif regulate the nuclear transport activity of triple C_2H_2 zinc finger proteins

Chih-Ying Lin, Lih-Yuan Lin*

Institute of Molecular and Cellular Biology and Department of Life Science, National Tsing Hua University, Hsinchu, Taiwan, ROC

* lylin@life.nthu.edu.tw

Abstract

Zinc finger (ZF) motifs on proteins are frequently recognized as a structure for DNA binding. Accumulated reports indicate that ZF motifs contain nuclear localization signal (NLS) to facilitate the transport of ZF proteins into nucleus. We investigated the critical factors that facilitate the nuclear transport of triple C_2H_2 ZF proteins. Three conserved basic residues (hot spots) were identified among the ZF sequences of triple C_2H_2 ZF proteins that reportedly have NLS function. Additional basic residues can be found on the α-helix of the ZFs. Using the ZF domain (ZFD) of Egr-1 as a template, various mutants were constructed and expressed in cells. The nuclear transport activity of various mutants was estimated by analyzing the proportion of protein localized in the nucleus. Mutation at any hot spot of the Egr-1 ZFs reduced the nuclear transport activity. Changes of the basic residues at the α-helical region of the second ZF (ZF2) of the Egr-1 ZFD abolished the NLS activity. However, this activity can be restored by substituting the acidic residues at the homologous positions of ZF1 or ZF3 with basic residues. The restored activity dropped again when the hot spots at ZF1 or the basic residues in the α-helix of ZF3 were mutated. The variations in nuclear transport activity are linked directly to the binding activity of the ZF proteins with importins. This study was extended to other triple C_2H_2 ZF proteins. SP1 and KLF families, similar to Egr-1, have charged amino acid residues at the second (α2) and the third (α3) positions of the α-helix. Replacing the amino acids at α2 and α3 with acidic residues reduced the NLS activity of the SP1 and KLF6 ZFD. The reduced activity can be restored by substituting the α3 with histidine at any SP1 and KLF6 ZFD. The results show again the interchangeable role of ZFs and charge residues in the α-helix in regulating the NLS activity of triple C_2H_2 ZF proteins.

Introduction

Zinc finger domain (ZFD) is among the most common DNA binding structures found in eukaryotic transcription factors. A variety of ZF motifs have been identified [1–3]. The most typical structure has two cysteines and two histidines (C_2H_2) to interact with a zinc ion. Approximately 700 genes encoding C_2H_2 ZFD have been identified in the human genome. They represent the most abundant DNA binding proteins in cells [4]. C_2H_2 ZFD was identified initially in the transcription factor TFIIIA of *Xenopus laevis* [5] and presumed to present only in eukaryotic cells. Later studies indicate that proteins with C_2H_2 ZFD also exist in prokaryotic cells, suggesting a structural conservation in evolution [6]. Besides DNA binding, the C_2H_2 motif of the ZF proteins participates also in RNA binding [7, 8] and protein-protein interactions [9].

The C_2H_2 ZF motif has a highly conserved structure. It contains 28–30 amino acid residues with a consensus sequence of ψ-X-C-X_{2-4}-C-X_3-ψ-X_5-ψ-X_2-H-X_{3-4}-H (ψ stands for hydrophobic and X represents any amino acid residues) which folds into a ßßα secondary structure. The zinc ion binds to the cysteine and histidine residues to form a unique steric configuration [10–13]. The α-helix of the motif inserts into the major groove of DNA. Each C_2H_2 ZF motif covers three base pairs of DNA generating a specific recognition between the ZF protein and a DNA sequence [14, 15].

ZF proteins usually have tandem repeats of ZFs connected by a consensus Thr-Gly-Glu-Lys-Pro sequence. Based on the number and the distribution of the C_2H_2 ZF, proteins can be categorized as (a) single, (b) triple, (c) multiple adjacent and (d) separated paired C_2H_2 ZF proteins [16]. The C_2H_2 motif on the single C_2H_2 ZF protein has to coordinate with other region of the protein for DNA binding [17]. The other types of C_2H_2 ZF proteins can interact directly with DNA. For triple C_2H_2 ZF proteins, such as Egr-1 (Zif268), Krüppel-like Factor (KLF) or SP1, all the ZF domains are involved in DNA binding and even protein-protein interaction [18–20]. However, the role of ZF may vary in multiple adjacent C_2H_2 ZF proteins. For example, ZF1-ZF4 of Miz-1 are responsible for interacting with SMAD3 while the rest of the ZF motifs participate in binding the target promoter [21]. OAZ has 30 ZF motifs. Only ZF2-ZF5 interacts with specific DNA sequence, whereas all other ZF motifs can interact with proteins [22]. These findings demonstrate the diverse functions of ZFs.

Proteins are synthesized in the cytoplasm. For transcription factors that localize mainly in the nucleus, they have to be actively transported into the nucleus. These proteins have a requisite nuclear localization signal (NLS) that recognize by carrier proteins (importin or karyopherin) for the process of translocating through the nuclear pore [23, 24]. Classical NLS consists of consecutive basic amino acid residues occur in one (monopartite) [25] or two (bipartite) fragments separated by 10–12 residues [26, 27]. Proteins carrying classical NLS can be recognized by Importinα/ß [28–30]. Importinα (Impα) binds the cargo protein while Importinß (Impß) interacts with Impα. The complex then interacts with nucleoporins at the cytoplasmic side of the nuclear pore complex and transports the cargo protein into the nucleus [31–34].

Recent studies indicate that the C_2H_2 ZFD can serve as an NLS to assist the nuclear entrance of ZF proteins. Unlike proteins with the classical NLS that contains consecutive basic residues, the transport mechanism for ZF proteins is not clear. However, the basic residues in the C_2H_2 ZF motifs are considered to associate functionally with nuclear transport. Noticeably, most proteins with a single C_2H_2 ZF motif have difficulty in translocating into the nucleus. At least two C_2H_2 ZF motifs are required in a ZFD to serve as NLS [35–40]. Evidence from recent studies suggest that residues from a few adjacent ZF motifs of a triple or multiple C_2H_2 ZF protein [36, 41–45] can serve as NLS. Specific structural characteristics and/or residues in these

ZF motifs are recognized by the carrier proteins [42, 46]. Besides binding with Impα [43, 47], C_2H_2 ZF proteins can bind directly with Impβ or β-like importins for transport [35, 48]. C_2H_2 ZF motifs are highly conserved in structure and even basic amino acid residues distribution. Since the number and the role of C_2H_2 ZF vary in ZF proteins, the function of each ZF cannot be distinguished easily. In this study, we covalently joined the ZFD of triple C_2H_2 ZF proteins to green fluorescent proteins (GFP) and investigated the nuclear transport activity of the fusion proteins. Residues on the ZFD that modulate the nuclear transport process are characterized.

Materials and methods

Cell culture and chemicals

Chinese hamster ovary (CHO) K1 cells (ATCC® CCL-61™) were cultured at 37 ˚C in McCoy's 5A medium with 10% heat-inactivated fetal bovine serum (FBS), 0.22% sodium bicarbonate, 100 units/ml ampicillin, and 100 mg/ml streptomycin in 5% CO_2, 95% air and 100% humidity. Reagent grade chemicals were purchased from Sigma unless specified. Cell culture reagents were purchased from Invitrogen/Gibco. DNA polymerase (KOD-plus) was obtained from TOYOBO. Lipofectamine™ 2000 was obtained from Invitrogen. TALON® Metal Affinity resin was provided by Clontech. Glutathione Sepharose 4B beads and secondary antibodies (anti-mouse and anti-rabbit IgGs) were from GE Healthcare. Protein and DNA molecular weight markers were from SMOBIO. Anti-GFP (ab290) antibodies were purchased from Abcam. Anti-His (C-term) and anti-GST (sc-138) antibodies were from Invitrogen and Santa Cruz Biotechnology, respectively.

Plasmid constructions

GST-importinα1 [49] plasmid was kindly provided by Dr. Yoshihiro Yoneda of Osaka University. Human EGR-1 (residues 338–418), SP1 (residues 626–708) and KLF6 (residues 200–282) were amplified by polymerase chain reactions (PCR) using human cDNA as templates. The sequences of the gene-specific primers are listed in S1 Table. The PCR products were restricted with appropriate enzymes and the fragments were cloned into peGFP-C1 vectors to generate the various GFP-ZFD fusion proteins. Importinβ1 expressed with an N-terminal His-tag (His-Impβ1) was constructed similarly with 2 sets of primers as indicated in S1 Table and inserted into the pCDNA4 vectors. Site-directed mutagenesis was performed by PCR-based Quick-Change method (Stratagene), and the primer sequences are listed in S2 Table. All mutations were confirmed by DNA sequencing.

Subcellular localization analysis

CHO K1 cells (8×10^5) were seeded on coverslip in 24-well dish for 24 h. Transfections were conducted by using Lipofectamine™ 2000 with 0.8 μg of plasmid DNA according to the manufacturer's instructions. Cells were transfected for 24 h then rinsed with ice-cold phosphate-buffered saline (PBS) (0.14 M NaCl, 2.7 mM KCl, 1.5 mM KH_2PO_4, 8 mM Na_2HPO_4, pH 7.4), and fixed with 4% para-formaldehyde in PBS at 4 ˚C for 30 min. After washing three times with PBS, cells were stained with Hochast 33258 (100 ng/ml) in PBST (PBS with 0.3% Triton® X-100) at room temperature for 15 min. Cells were washed three times with PBS, then mounting medium (90% glycerol, 5% n-propyl gallate) was applied to the coverslips for fluorescence microscopy. The subcellular localization of the GFP-fusion proteins was analyzed with Axio-Vision Rel. 4.8 (Zeiss). The florescence signals from each cell were quantified as pixel values from the cytoplasm (C) or nucleus (N). GFP-fusion proteins were considered to localize

exclusively in nucleus when greater than 90% of the fluorescence signals (C/N pixel values < 10%) came from the nucleus. Results were generated from 30 cells randomly picked from each sample. Statistical analysis was conducted using Student's t test.

His-tag pull-down assay

CHO K1 cells were co-transfected with GFP-tagged proteins and impβ1-His for 24 h. Cells were lysed and cell lysates were added to microcentrifuge tubes containing 20 μl of cobalt TALON® Metal Affinity resin (Clontech) and 1 ml binding buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, 0.4% NP-40, 10% glycerol). Tubes were rotated at 4 ˚C for 2 h. The resins were then washed once with binding buffer, and followed sequentially with binding buffer containing increasing concentrations of imidazole. The wash solutions were collected for immunoblotting.

GST pull-down assay

GST-impα1 was expressed in *E. coli* at 20˚C for 14 h by inducing with 1 mM IPTG. The cells were lysed by sonication in extraction buffer (PBS with 0.75 mg/ml lysozyme, 0.5 mM PMSF), and the cell extracts were incubated with 20 μl of Glutathione Sepharose 4B beads and 1 ml binding buffer (20 mM HEPES, 110 mM potassium acetate, 5 mM sodium acetate, 2 mM magnesium acetate, 250 mM sucrose, 0.5 mM EGTA, 2 mM DTT, pH 7.9) for 2 h at 4 ˚C. The buffer was removed and the beads were further incubated with protein extracts prepared from CHO K1 cells for 2 h at 4 ˚C. Beads were washed with binding buffer three times, then boiled in SDS buffer and the supernatants were collected for immunoblotting.

Preparation of whole cell lysates

Cells were harvested by centrifugation at 12,000 x g and resuspended in five volumes of extraction buffer (20 mM HEPES, pH 7.9, 400 mM NaCl, 0.5 mM PMSF, 50 mM NaF, 0.5 mM Na_3VO_4, 2 μg/ml aprotinin, 5 μg/ml leupeptin, 1 μg/ml pepstatin, and 0.5% NP-40). Tubes were incubated on ice for 10 min then vortexed vigorously for 5 sec. The vortexing step was repeated four times, and cell debris were removed by centrifugation at 12,000 x g for 20 min. The supernatants were collected for immnoblotting.

Immunoblotting analysis

Proteins in the supernatants were separated on SDS-polyacrylamide gels (5% stacking and 8% separating gels) then transferred electrophoretically onto PVDF membranes (GE Healthcare) in a transfer cell (Bio-Rad). The membranes were pre-hybridized in TBST buffer (150 mM NaCl, 10 mM Tris, pH 8.0, 0.1% Tween-20) with 5% skim milk for 1 h, and then incubated with primary antibodies in 5% skim milk at 4 ˚C overnight. After washing three times with TBST buffer, the membranes were submerged in 5% skim milk containing horseradish peroxidase-conjugated secondary antibodies at room temperature for 1 h. After washing three times with TBST buffer, the membranes were rinsed once with TBS buffer (TBST buffer without Tween-20) for 5 min, and developed by the ECL system (Amersham). Visualization was processed by ImageQuantTM LAS 4000 mini imager (GE Healthcare) and the intensities of the bands were quantified with the UN-SCAN-IT gel analysis software (version 6).

Statistical analysis

Statistical analyses were conducted using the results of three independent experiments. The differences between treatments were determined by using two-tailed Student's t test.

Results

Reportedly, ZFD can serve as an NLS and plays a role in nuclear transport. We investigate here the factors that affect the NLS role of the ZFD using ZF proteins with three adjacent C_2H_2 ZF motifs. The ZFD sequence of several triple C_2H_2 ZF proteins (Egr-1 [38], KLF1 [44], KLF6 [41], KLF8 [50], and SP1 [46]) that reportedly act as NLS were aligned. The residues that modulate the nuclear transport activity are expected to conserve in these ZF motifs. Fig 1A shows the sequence alignments. The residue numbers denoted in Fig 1A refer to the positions of the amino acid residues in the ZFD, not those of the whole proteins. Since basic residues are
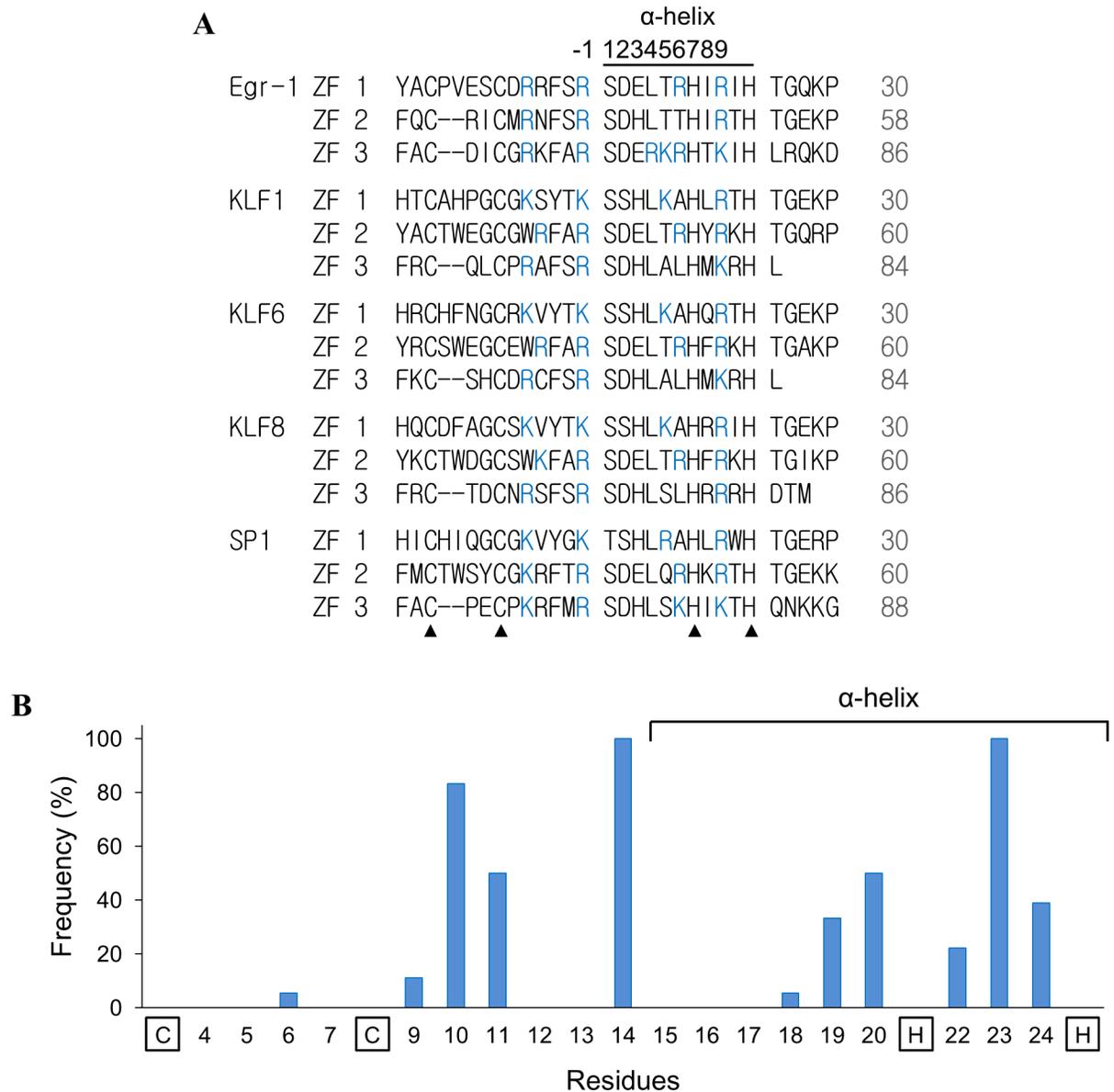
Fig 1. Sequence alignment and analysis of the ZFs from triple C_2H_2 ZF proteins with ZFD that reportedly associate with nuclear transport activity. (A) The ZF sequences of human Egr-1 (NP_001955.1), Sp1 (NP_612482.2), KLF1 (NP_006554.1), KLF6 (NP_001291.3), and KLF8 (NP_009181.2) were aligned. Solid triangles indicate the conserved Cys and His residue in the ZFs. (B) The frequency of basic residues appeared in each position of the ZF according to the proteins listed in (A).

recognized to involve in nuclear transport, the frequency of a basic amino acid occurs in a specific position is shown (Fig 1B). Noticeably, except the Zn-interacting residues (Cys and His), occurrences of basic residues are rather conserved. The "hot spots" (higher than 80% frequency) at homologous positions of 10, 14 and 23 of each ZF motif (numbered according to the ZF-1 of Erg-1) can be easily detected. Structurally, the α-helical region of the ZF motif encompasses residues 15 to 25 [51], five out of the eight hot and high frequency spots occur in this region.

Despite the high similarity in the distribution of basic residues among ZFs, the contribution of each ZF towards nuclear transport may not be equivalent. We investigated the intrinsic factors that modulate the activity of the ZFs. The ZFD of Egr-1 was used as a model for this study. Egr-1 does not have a classical NLS and thus its ZFD is most likely responsible for its nuclear transport.

The sequence of Egr-1 ZFD was amplified from human cDNA and fused with GFP. The construct was used as a template to generate other mutants. To investigate the significance of the "hot spots" in nuclear transport, we replaced individually the three conserved basic residues in every ZF of Egr-1 with alanine. The mutated constructs were transfected and expressed in CHO K1 cells. The proportion of GFP-fusion proteins localize exclusively in the nucleus were recorded. Representative microscopic images in this work are included as S1 Fig. As shown in Fig 2A, GFP by itself does not have nuclear transport activity. The wild type (WT) Egr-1 ZFD fused to GFP presented mainly in the nucleus, indicating the Egr-1 ZFD is responsible for translocating the fusion proteins into the nucleus. Mutation at the hot spots of Egr-1 ZF1 (R10A, R14A and R23A) significantly reduced the nuclear transport activity. Reportedly, ZF1 has limited involvement in the nuclear transport of Egr-1 [38]. Our results indicate that ZF1 does contribute to the nuclear transport activity of Egr-1 ZFD.

The homologous hot spots at the ZF2 and ZF3 of Egr-1 were also mutated. Substitution at any one hot spot of ZF2 (R38A, R42A and R51A) showed a marked reduction in the nuclear transport activity (Fig 2B). The reduction level is more significant than that of ZF1. The same effect was found when similar mutations were generated at ZF3 (R66A, R70A and K79A, Fig 2C). Results from these experiments imply that removing the conserved basic residues on the ZFs can effectively inhibit the nuclear transport activity of the Egr-1 ZFD. In addition, the mutations on ZF2 and ZF3 are more effective than those on ZF1 for this inhibition process.

Further sequence analysis reveals that the residues homologous to R34 and H45 of the Egr-1 ZF2 are acidic in ZF1 (E6 and E17) and ZF3 (D62 and E73). R34 is on the linker connecting the 2 β-sheets while H45 is within the α-helix. These acidic residues might interfere with the activity for nuclear entrance. Thus R34 and H45 on ZF2 were replaced with acidic residues individually (R34D and H45E) or concurrently (R34D/H45E). The resulting fusion proteins all have reduced nuclear transport activity. The mutation at H45, which is within the α-helical region, is even more effective than the replacement at the R34 position. Moreover, the effect seems to be synergic since the R34D/H45E double mutant has the lowest nuclear transport activity (Fig 3A). The results imply that basic residues help while acidic residues inhibit the nuclear transport of ZFD proteins. In supporting this hypothesis, the only acidic residue (which occupies the 2nd position on the α-helix) in ZF2 was replaced with arginine (D44R). The resulting fusion protein allocates exclusively in the nucleus (Fig 3A).

As mentioned above, the amino acid residues in ZF1 (E6 and E17) and ZF3 (D62 and E73) homologous to R34 and H45 of ZF2 are acidic and the R34D/H45E double mutant has drastically reduced nuclear transport activity (Fig 3A). We substituted these acidic residues in ZF1 with Arg or His. Fig 3B shows fusion proteins carrying either the E6R or E17H substitution have similar nuclear transport activity as the wild type Egr-1 ZFD. Substitution of E6R, E17H or both replacements to the ZF2 R34D/H45E double mutant partially or completely restored

Fig 2. Effect of conserved basic residues in ZFs on the nuclear transport activity of Egr-1 ZFD. Egr-1 ZFD was fused with GFP. The conserved basic residues (hot spots) in ZF1 (A), ZF2 (B) and ZF3 (C) of the fusion proteins were individually mutated to alanine. The plasmids were transfected into CHO K1 cells for protein expression. Egr-1 ZF sequences are shown with the mutated residues in red. Solid triangles indicate the conserved Cys and His residue in the ZFs. The proportion of cells having the GFP fusion protein entirely localized in the nucleus was determined. Representative microscopic images are included as S1 Fig. Each value represents a mean ± standard deviation of three independent experiments. Asterisk indicates significant difference as compared to that of the wild type (WT) fusion protein. *: $p < 0.05$; **: $p < 0.01$.

https://doi.org/10.1371/journal.pone.0191971.g002

the nuclear transport activity of the fusion proteins. Conversely, replacing the corresponding acidic residues in ZF3 (D62 and E73) with basic residues showed an enhanced nuclear transport activity of the fusion proteins (Fig 3C). Substitution of E73H or both D62R and E73H to the ZF2 R34D/H45E double mutant partially restored or even enhanced the nuclear transport activity of the Egr-1 GFP-ZFD. The results demonstrated that all three ZFs on Egr-1 contribute interactively to the NLS activity of the protein.

The hot spots of ZF1 play a less effectively role in modulating the nuclear transport as compared to that of ZF2 and ZF3 (Fig 2). Since the quadruple mutant (E6R/E17H/R34D/H45E) has restored the nuclear transport activity of the GFP-ZFD fusion protein (Fig 3B), the role of the hot spots in ZF1 of the quadruple mutant was examined. Fig 3D shows that the quadruple mutant carrying additional substitutions at the hot spots of ZF1 has its restored nuclear transport activity diminished. Similar results were found when the hot spots at ZF3 of the quadruple mutant (R34D/H45E/D62R/E73H) were mutated (Fig 3E). Substitution of the hot spot residue to alanine reduced the enhanced nuclear transport activity to a similar or even lower activity than that of the wild type protein. These results show the impact of charge residues in modulating the NLS activity of the ZFD.

The α-helical region of the ZF interacts with the major groove of target DNA. Three basic residues in tandem (^{74}RKR^{76}, α4-α6 of the helix) can be found in the ZF3 of Egr-1. Mutation of these three residues to alanine individually resulted in a dramatic reduction of the nuclear transport activity (Fig 4A). The effect was most prominent with the mutation at α6. Since quadruple mutant (R34D/H45E/D62R/E73H) has an enhanced nuclear transport activity (Fig 3C, 6th column), and E73 is located right next to the ^{74}RKR^{76} sequence, the role of ^{74}RKR^{76} was examined individually. Mutation at any of the basic residue significantly reduced the activity of the quadruple (R34D/H45E/D62R/E73H) mutant. However, the reduction was not as significant as that of wild type (Fig 4B). The protein levels of the wild type fusion proteins and mutants in the cells are similar. The reduction in transport efficiency is not due to the lack of fusion proteins carrying multiple mutations.

The homologous sequence of ^{74}RKR^{76} at ZF2 is ^{46}LTT^{48} and does not carry any positively charged side chains. The L46R and T48R mutation enhanced the nuclear localization activity of the Egr-1 ZFD (Fig 4C). Apparently, a positive charge at the 4th to 6th position of the α-helix on ZF2 can improve the nuclear transport activity of Erg-1. The results indicate the impact of basic residues at hot spots and α-helix for the nuclear transport activity of ZFs. Whether this mutation affects the DNA binding activity or specificity of Erg-1 remains to be investigated.

Entrance of proteins through nuclear pore complex requires the assistance of importins. Reportedly, several C$_2$H$_2$ ZF proteins interact with importins for nuclear translocation [43, 47, 48, 50]. Since the charge residues in the Egr-1 ZFD affect the NLS activity, we examined whether the effect is derived from the interaction of ZFs with importins. A GST pulldown assay was utilized to examine the interactions. GST-fused Importinα1 (GST-Impα1) was expressed in *E. coli*, and GFP-fused Egr-1 ZFD was synthesized in CHO K1 cells. Cell extracts were prepared and mixed to conduct the assay. As shown in Fig 5A, Egr-1 ZFD can interact with Impα1. The interaction reduced significantly for the R34D/H45E mutant, indicating the presence of acidic residues at these positions affect the interaction. However, this interaction was restored for the E6R/E17H/R34D/H45E and the R34D/H45E/D62R/E73H mutants. A higher binding capacity was noted for R34D/H45E/D62R/E73H which corresponds to the higher nuclear transport activity for this mutated protein (Fig 3E, 3rd column). These results suggest the contribution of the charged residues on ZFs in nuclear transport.

Since Impα1 usually interacts with Impβ in the nuclear transport process. The interaction between ZFD and Impβ1 was also analyzed. Impβ1 was fused with 6 histidine residues
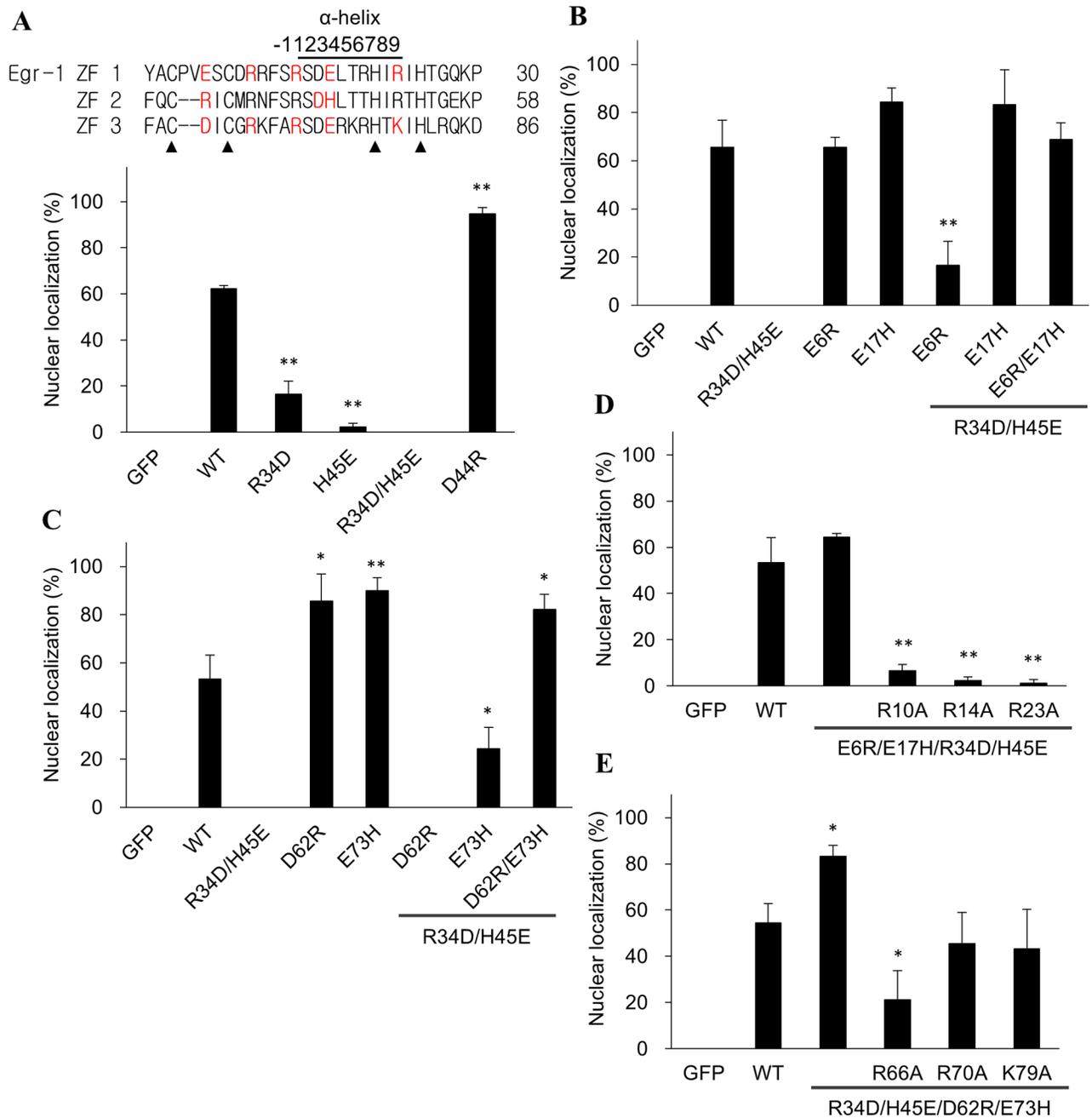
Fig 3. Effect of charge residues in ZFs on the nuclear transport activity of Egr-1 ZFD. Egr-1 ZFD was fused with GFP. (A) Basic (R34 and H45) or acidic (D44) residues in the ZF2 were mutated to oppositely charged residues individually or concurrently. (B) The GFP fusion protein was mutated firstly at ZF2 (R34D/H45E). Additional mutations were performed on ZF1 (E6R, E17H and E6R/E17H). Point mutation (E6R and E17H at ZF1) for the wild type (WT) fusion protein was also conducted. (C) The acidic residue in the ZF3 of the R34D/H45E mutant was further substituted with basic residues (D62R, E73H and D62R/E73H). Point mutation (D62R and E73H at ZF3) for the WT fusion protein was also conducted. (D) The hot spots in the ZF1 or the ZF3 (E) of the E6R/E17H/R34D/H45E mutant protein were further substituted with alanine. Egr-1 ZF sequences are shown on the top panel with the mutated residues in red. Solid triangles indicate the conserved Cys and His residue in the ZFs. The plasmids were transfected into CHO K1 cells for protein expression. The proportion of cells having the GFP fusion protein entirely localized in the nucleus was determined. Representative microscopic images are included as S1 Fig. Each value represents a mean ± standard deviation of three independent samples. Asterisk indicates significant difference as compared to that of the wild type (WT) fusion protein. *: $p < 0.05$; **: $p < 0.01$.
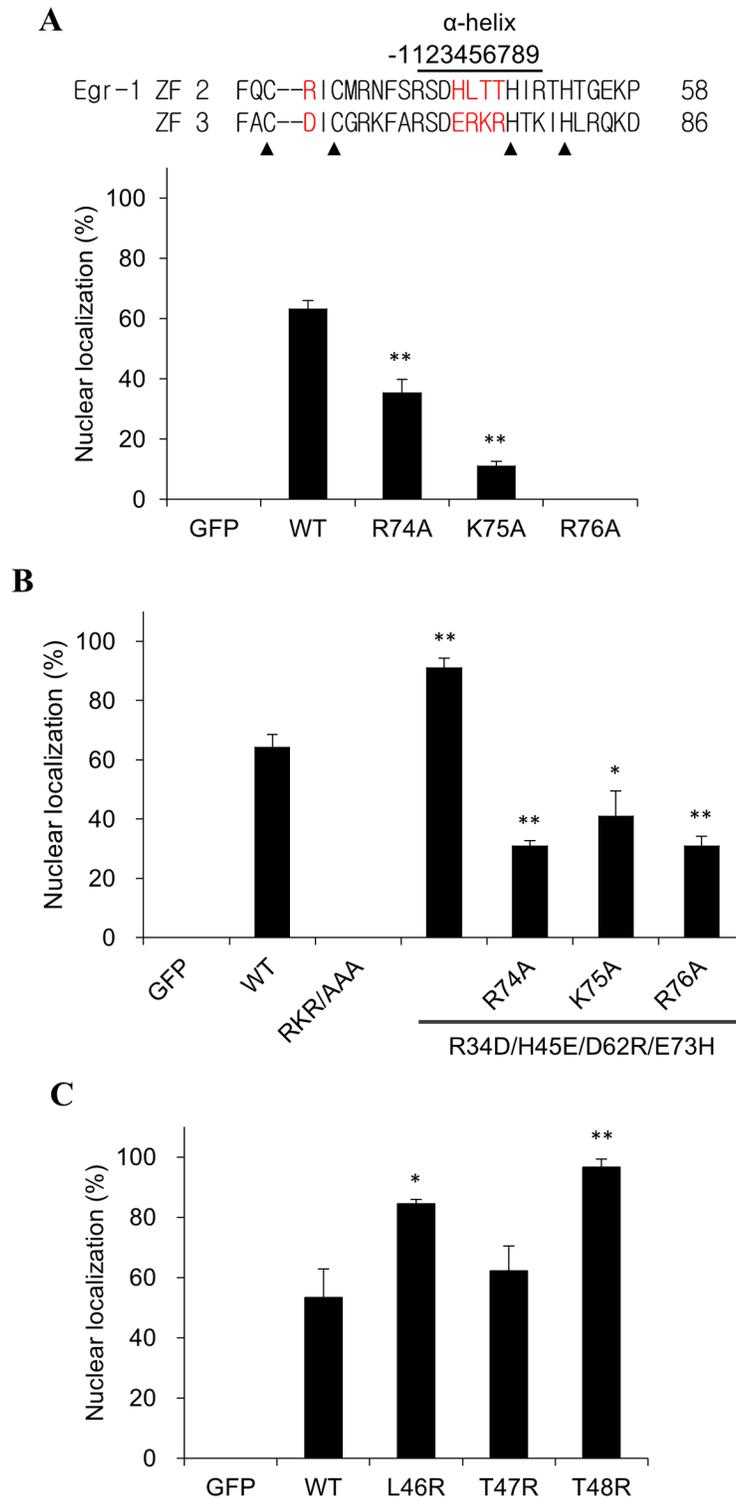
https://doi.org/10.1371/journal.pone.0191971.g003

Fig 4. Effect of basic residues at the α-helical region of ZFs on the nuclear transport activity of Egr-1 ZFD. Egr-1 ZFD was fused with GFP. The basic residues at the α4-α6 (^{74}RKR^{76}) region in the ZF3 of Egr-1 ZFD were substituted with alanine individually (A) or concurrently (RKR/AAA) (B). The hot spot at ZF3 of the R34D/H45E/D62R/E72H mutant protein was mutated to alanine individually. (C) The homologous sequence of ^{74}RKR^{76} at ZF2 (46LTT48) in WT Egr-1 ZFD was mutated to arginine individually. Egr-1 ZF sequences are shown on the top panel with the mutated residues in red. Solid triangles indicate the conserved Cys and His residue in the ZFs. The plasmids were transfected

into CHO K1 cells for protein expression. The proportion of cells having the GFP fusion protein entirely localized in the nucleus was determined. Representative microscopic images are listed as S1 Fig. Each value represents a mean ± standard deviation of three independent samples. Asterisk indicates significant difference as compared to that of the wild type (WT) fusion protein. *: p < 0.05; **: p < 0.01.

(Impβ1-His) and co-expressed with the GFP-fused Egr-1 ZFD constructs in CHO K1 cells. The interaction can be revealed by cobalt beads pull-down assay. Fig 5B shows the interaction of Egr-1 ZFD and Impβ1. A similar result as that of Fig 5A was obtained. The R34D/H45E mutant, but not the E6R/E17H/R34D/H45E and the R34D/H45E/D62R/E73E mutants, showed a decrease of interaction with Impβ1 as compared to that of wild type Egr-1 ZFD. The result implies that both importins are required for the nuclear transport of the Egr-1 ZFD.

The above results show that charge residues in the ZFD of triple C_2H_2 ZF proteins are critically associated with nuclear transport. We examined the distribution and proportion of charge residues in the ZFs. Sixty-one triple C_2H_2 ZF proteins (183 ZF motifs) were identified and retrieved from the Uniprot database [52]. The hot spots are conserved in these ZF proteins (Fig 6A), indicating the significance of the hot spots in nuclear transport. Besides, a marked proportion of charge residue locates at the α-helix of the ZFs. Acidic residues are mainly at the α2 and α3, while the basic residues are found at the α3, α5 and α6 positions. Noticeably, the basic residue at α3 is exclusively His while Arg and Lys residues are mainly at α5 and α6. This finding indicates that the charge residues in the hot spots and the α-helix of ZF motif are important for the NLS function.

Further analysis of the acidic residue in the α-helix reveals that 96% (81/84) of α2 and 85% (46/55) of α3 residue is Asp and Glu, respectively. These DE residues (α2-α3) present in the ZF1 and ZF3 of Egr-1 ZFD (Fig 1). Mutation at the α3 position of ZF2 (H45E) greatly reduced the NLS activity (Fig 3A). On the contrary, replacing the α2 with basic residue (D44R) significantly enhanced the nuclear transport activity of the Egr-1 ZFD (Fig 3A). We examined the occurrence of DE at α2-α3 of the triple C_2H_2 ZF proteins and found that this combination can only be found in the ZFs of Egr, SP1 and KLF protein families. Further analysis of residues in the α-helix of those ZFs showed a similar pattern as that depicts in Fig 6A (Fig 6B). Noticeably, the α2-α3 combination is consistently SH (ZF1), DE (ZF2) and DH (ZF3) for SP1 and KLF families. Although a high proportion of His residue presents at α3, changing SH to DE (S16D/H17E) or DH to DE (H75E) in KLF6 did not alter the nuclear transport activity (Fig 6C). Concurrent substitution of the α2-α3 to DE (S16D/H17E/H75E) resulted in a marked reduction of the nuclear transport activity. However, further substitution of DE at ZF2 of KLF6 to DH (S16D/H17E/E47H/H75E), the same α2-α3 charge distribution as that of Egr-1 ZFD, restored the activity (Fig 6C).

Experiments utilizing SP1 ZFD to conduct the experiment generate similar results; the nuclear transport activity dropped only when the DE sequence presents in the α2-α3 of every ZF (Fig 6D). The activity reduced further when α6 (K78) of SP1 ZF3 was mutated (S2 Fig). Applying GST pull down assay to examine the interaction of the KLF6/SP1 ZFD with Impα1 reveals that the interaction reduced with the DE residues presented at the α2-α3 of every ZF (S3 Fig). The interaction restored when the DE at ZF2 was further substituted with DH. Besides the acidic residues, substitution of basic residue (R50A) at α6 of the S16D/H17E/E47H/H75E mutant decreased the restored NLS activity again for KLF6 or SP1 ZFDs (Panels A and B in S4 Fig). These results indicate the influence of charge residues at the α-helix of ZF in regulating the nuclear transport activity.
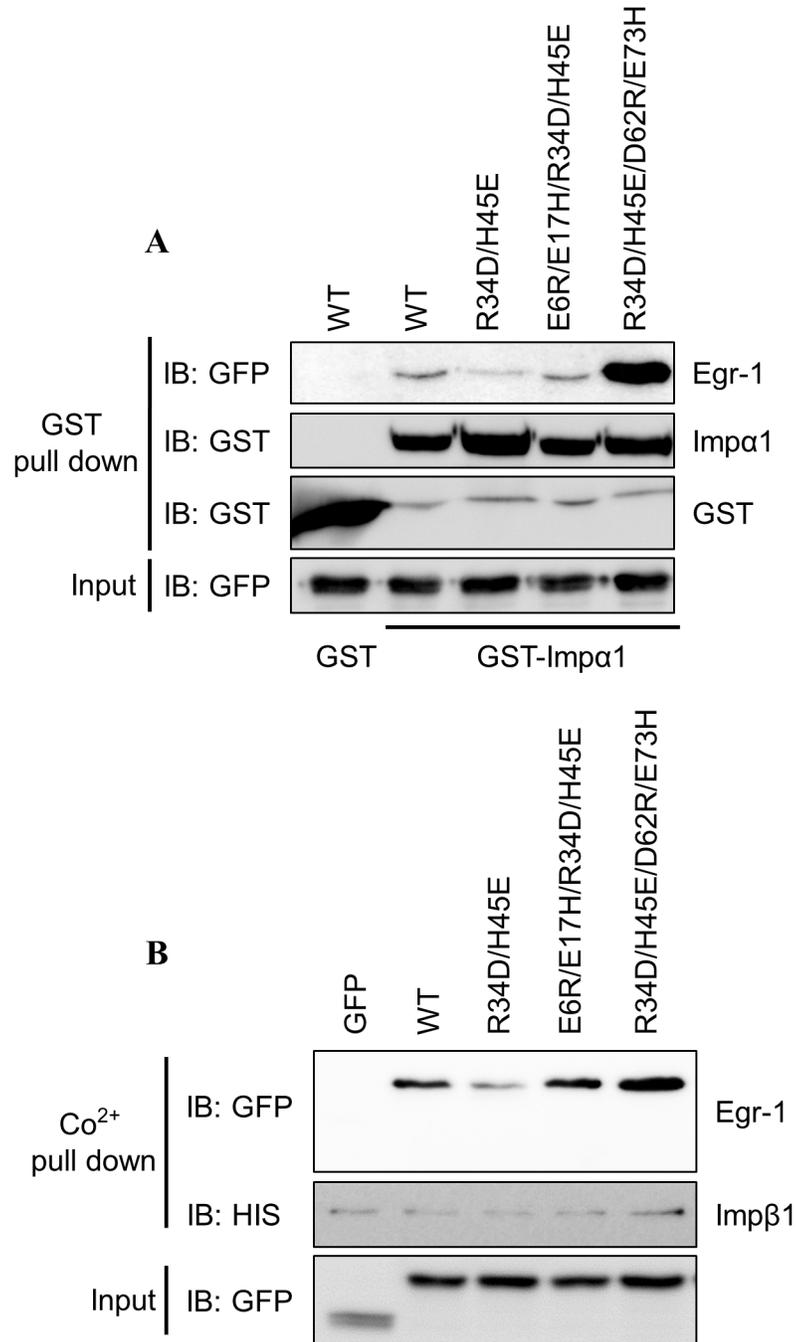
Fig 5. The binding affinity of nuclear transport proteins with Egr-1 GFP and the mutated proteins. (A) Cell extracts containing GFP-Egr-1 ZFD fusion protein or its mutants were incubated with extracts expressing GST-Impα1. The interactions between the fusion proteins and GST-Impα1 were analyzed by the GST pull-down assay. (B) GFP-Egr-1 ZFD fusion protein or its mutants were co-expressed with His-tagged Impβ1. The interactions between the fusion proteins and Impβ1 were analyzed by the His-tag pull-down assay.

Discussion

Accumulated studies indicate that ZFD can be an NLS to facilitate nuclear translocation. However, the transport mechanism remains unclear. The configurations of the C_2H_2 ZFs are
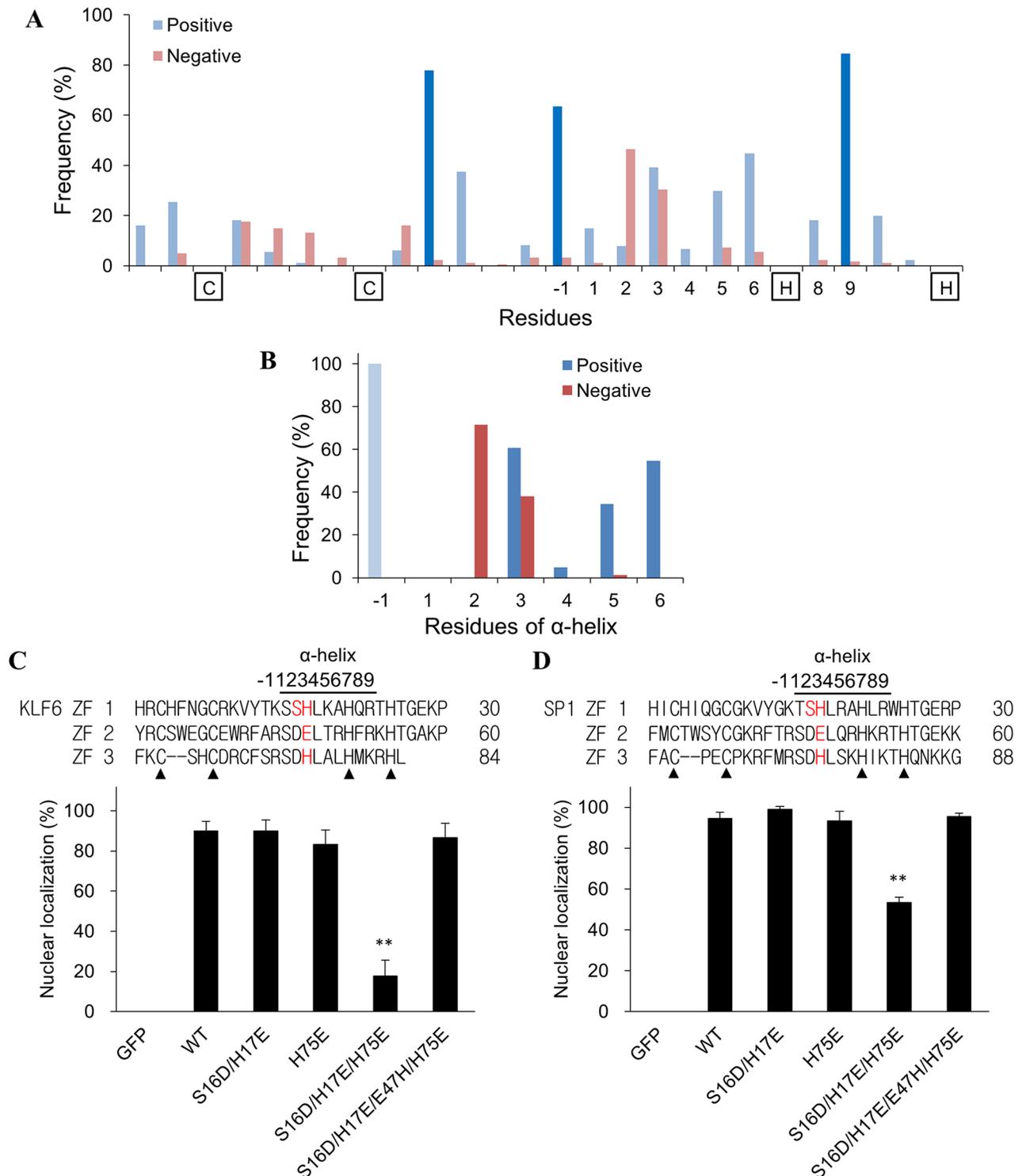
Fig 6. Analysis of acidic residues in the α-helical region of triple C₂H₂ ZFD and their role in nuclear transport. (A) The ZF sequences of 61 triple C₂H₂ ZF proteins (183 ZFs) were retrieved from the Uniprot database. The frequency of charge residue distribution in each position was analyzed. The basic and acidic residues are highlighted in blue and red colors, respectively. (B) The sequences of the α-helical region of the Eger/Sp1/KLF family were analyzed. The frequencies of occurrence of basic and acidic residues at each position are represented in blue and red colors, respectively. GFP was fused with KLF6 (C) or SP1 (D) ZFD. The α2 and/or α3 of the α-helix of KLF6 or SP1 ZFD were mutated to acidic residues. The construct with acidic residues (DE at α2-α3) for all three ZFs (S16D/H17E/H75E) was further mutated at E47 (S16D/H17E/E47H/H75E). Egr-1 ZF sequences were shown on the top panel with the mutated residues in red. Solid triangles indicate the conserved Cys and His residue in the ZFs. The sequence of the KLF6 or SP1

ZFD is shown in the upper panel of each figure with the mutated residues in red. Solid triangles indicate the conserved Cys and His residue in the ZFs. The plasmids were transfected into CHO K1 cells for protein expression. The proportion of cells having the GFP fusion protein entirely localized in the nucleus was determined. Representative microscopic images are listed as S1 Fig. Each value represents a mean ± standard deviation of three independent samples. Asterisk indicates significant difference as compared to that of the wild type (WT) fusion protein. **: $p < 0.01$.

similar and conserved amino acid residues can be identified among their sequences. Results from this study show that charged residues in the triple C_2H_2 ZFD are associated with nuclear transport activity. Although the distribution of basic residues is conserved in the ZFs (Figs 1A and 6A), not every ZF contributes equally to the NLS activity. We showed in this study that the conserved basic residues (hot spots) and the charge residues in the α-helical region modulate the NLS activity of the triple C_2H_2 ZF proteins.

Distribution of the charge residues at the α-helical region varies among the C_2H_2 ZFs. For the Egr/SP1/KLF families, the charge residues present within the second (α2) to the sixth (α6) position of the α-helix affects the NLS activity of the ZFs. The acidic residue at the α2 and α3 positions may reduce the activity while the basic residue at α5 and α6 enhance nuclear entry. The α-helix constitutes a small region but exposed on the surface of the ZF. Since we showed clearly that the ZFs interact with importins and the interactions determine the activity of nuclear entry, the locations of the charge residues should be at sites that can be in contact with the carrier proteins.

The classical NLS has a cluster of basic residues which is the major feature recognized by the transporter proteins. ZFs do not have this recognition motif. However, the ZFD can sterically align the basic residues that are distributed at different ZFs to form a pattern to be recognized and interacted with the transporter proteins. Probably, this is the reason why a single C_2H_2 ZF is not sufficient to support the activity of nuclear entry. It should be noted that not every ZF in the multiple C_2H_2 ZFs is required to participate in the NLS activity. A variety of studies indicate that only 3 to 4 consecutive ZFs in a multiple C_2H_2 ZF proteins are involved in the NLS activity [35, 36, 53]. Since the classical NLS is composed of 4–5 basic residues, a few adjacent ZFs are sufficient to generate a monopartite- or bipartite-like construct for nuclear transport.

Results from our study also reveal that the NLS role of ZF is interchangeable. Egr-1 ZF2 plays a significant role in NLS activity. Mutation at ZF2 abolished the nuclear entry (Fig 3A). However, changing the characteristics of charge residue distribution in ZF1 to mimic that of ZF2 (E6R/E17H/R34D/H45E) restored the NLS activity (Fig 3B). The "functional" ZFs for nuclear transport switched from ZF2 to ZF1. For SP1/KLF family, at least one basic residue is required to present at the α3 position of the ZF to gain the full NLS activity; this basic residue can be in any of the three ZFs (Fig 6). Since changes in Egr-1 and SP1/KLF6 can alter the importin binding ability (Fig 5 and S2 Fig), the results indicate that the contact sites between ZFD and importin are interchangeable. The importin accesses the ZFD and interacts with the NLS region. Mutation at the NLS interface can either alters the interaction or generates a new site for the interaction. Alternatively, importin recognizes a broad region of the ZFD and a limited number of interactions are sufficient to form and stabilize the protein complex for nuclear transport.

ZFD is recognized initially as a DNA binding domain. Since accumulated results demonstrate that ZFD can serve as an NLS for nuclear transport, some residues in the triple C_2H_2 ZFD may have dual functions. The interaction of Egr-1 or SP1 ZFD with the DNA of target promoter has been explored [54, 55]. Examining the protein-DNA contact sites reveals that the first and the second hot spot of every ZF are involved in DNA binding. In addition, DE

residues in the α-helix (α2-α3) of ZF1 are required for Egr-1 to bind target promoter [56]. The DE residues in the ZF2 of SP1 recognize specific DNA in the target promoter [54]. Results from these studies indicate the significance and the dual roles of the charge residues in nuclear transport and DNA binding activities.

The three-dimensional (3D) structure of a triple C_2H_2 ZFD and importin interaction has not been reported. However, the crystal structure of Snail1 ZFD complexed with Impβ1 has been solved [48]. Snail1 has 4 ZFs and the ZFD provides the NLS activity. Each of the ZF in Snail1 is in contact with Impβ1 and thus participates in nuclear transport. All of the amino acid residues involved in the NLS activity locates at the hot spots and the α-helical region of each ZF. These residues aligned to form an interface as revealed by the 3D structure. The hot spots at the ZFs of Snail1 are conserved with those of triple C_2H_2 ZFD. Noticeably, the second and the third hot spot of ZF2, the first and the third hot spot of ZF3 and the second hot spot of ZF4 are part of the NLS and interact with the acidic residues on Impβ1. Other contact sites include α1 to α6 of the helices although not every residue is basic [48]. Even though the configuration and substrate interacting sites between Snail1 and triple C_2H_2 ZFD are different, the structural analysis of Snail1 implicates the critical role of hot spots and residues in the α-helical region of the protein in the nuclear transport of ZFs. It is reasonable to expect that residues involved in the nuclear transport of the triple C_2H_2 ZFs aligned at the same interface and they are accessible to the transport proteins. Since our results also indicate that the contributions among ZFs in nuclear transport can be interchanged, the interacting sites between ZFs and importins can vary. Certainly, we can exclude the possibility that there are multiple recognition sites between ZFs and importins and not all the recognition sites have to be interacted with importins to form a stable complex for nuclear transportation.

The SP1 and KLF6 have a ZFD with strong nuclear entry activity (Fig 6C and 6D). Both proteins have an acidic residue at the α3 of ZF2 but a basic residue at both ZF1 and ZF3. Mutation at the α3 of either the ZF1 or ZF3 to acidic residue did not alter the NLS activity. Egr-1 only has a basic residue at the α3 (H45) among the ZFs. Substitution of the amino acid with an acidic residue (H45E) drastically reduced nuclear transport activity (Fig 3A). Switching the basic residue at ZF2 with the acidic residue at ZF1 did not affect the NLS activity (Fig 3B). Possibly, at least one basic residue at the α3 of any ZF is sufficient to maintain the NLS activity of the ZFD.

Recently, C_2H_2 ZFD is broadly used in the application of cell engineering [57, 58]. Besides capable of specific DNA-binding, ZFD has cell penetration activity [59]. Study showed that classical NLS can serve as a cell penetration peptide via the action of basic amino acid residues [60]. The positively charged residues in ZFD might also play a similar role [61]. As discussed above, the hot spot and basic residues in the α-helical region of ZFs are involved in nuclear transport and DNA-binding activities. These residues have been suggested to associate with cell penetration [61]. Therefore, the basic residues on ZFD may have overlapping functions in cell permeability, nuclear transport and DNA binding. These findings can provide information for the design of functional tools in cell engineering.

In summary, we show herein the critical factors modulating the NLS activity of triple C_2H_2 ZF proteins. The conserved basic residues and the charge residues at the α-helical region of the ZFs regulate the nuclear transport activity. Presently, the conclusion is drawn from the results derived from triple C_2H_2 ZF proteins, especially on the Egr/SP1/KLF families. Mutations at the conserved hot spots and the basic residues in the α-helix reduce the nuclear transport activity. Whether this conclusion can be applied to other multiple ZF proteins remains to be investigated.

Supporting information

S1 Table. Primer sequences for fusion protein construction.
(DOCX)

S2 Table. Primer sequences used in ZFD mutagenesis.
(DOCX)

S1 Fig. Representative microscopic images in this study. The nuclei were stained with DAPI (blue). Scale bars, 10 μm.
(PDF)

S2 Fig. Effect of Lys78 in the α-helical region of ZF3 on the nuclear transport activity of SP1 ZFD. The Lys78 (α6) in the α-helical region of SP1 were mutated to alanine with S16D/H17E or S16D/H17E/H75E. The plasmids were transfected into CHO K1 cells for protein expression. The proportion of cells having the GFP fusion protein entirely localized in the nucleus was determined. Each value represents a mean ± standard deviation of three independent samples. Asterisk indicates significant difference as compared to that of the wild type (WT) fusion protein. *: $p < 0.05$; **: $p < 0.01$. SP1 ZF sequences are shown in the top panel with the mutated residues in red. Solid triangles indicate the conserved Cys and His residue in the ZFs. Representative microscopic images are shown in the middle panel. Scale bars, 10 μm.
(TIF)

S3 Fig. The binding affinity of Impα1 with KLF6 and SP1. Cell extracts containing wild type and mutated GFP-KLF6 ZFD (A) or GFP-SP1 ZFD (B) fusion protein were incubated with extracts expressing GST-Impα1. The interactions between the fusion proteins and GST-Impα1 were analyzed by the GST pull-down assay.
(TIF)

S4 Fig. Effect of charge residue in the α-helical region of ZF on the nuclear transport activity of KLF6 and SP1 ZFD. The α2-α3 residues in the α-helical region of KLF6 (A) or SP1 (B) were mutated (S16D/H17E/E47H/H75E) to mimic that of Egr-1 (DE, DH and DE at α2-α3 of ZF1, ZF2 and ZF3, respectively). The basic residue at α6 of the ZF2 was further mutated (R50A). The plasmids were transfected into CHO K1 cells for protein expression. The proportion of cells having the GFP fusion protein entirely localized in the nucleus was determined. Each value represents a mean ± standard deviation of three independent samples. Asterisk indicates significant difference as compared to that of the wild type (WT) fusion protein. **: $p < 0.01$. The sequence of the KLF6 or SP1 ZFD is shown in the upper panel of each figure with the mutated residues in red. Solid triangles indicate the conserved Cys and His residue in the ZFs. Representative microscopic images are shown in the left panel of each figure. Scale bars, 10 μm.
(TIF)

Acknowledgments

Author Contributions

Conceptualization: Chih-Ying Lin, Lih-Yuan Lin.

Funding acquisition: Lih-Yuan Lin.

Investigation: Chih-Ying Lin.

Methodology: Chih-Ying Lin.

Resources: Lih-Yuan Lin.

Software: Chih-Ying Lin.

Supervision: Lih-Yuan Lin.

Validation: Lih-Yuan Lin.

Writing – original draft: Chih-Ying Lin.

Writing – review & editing: Lih-Yuan Lin.

References

1. Klug A, Schwabe JW. Protein motifs 5. Zinc fingers. FASEB J. 1995; 9(8):597–604. Epub 1995/05/01. PMID: 7768350.

2. Mackay JP, Crossley M. Zinc fingers are sticking together. Trends Biochem Sci. 1998; 23(1):1–4. Epub 1998/03/28.PMID: 9478126.

3. Sanchez-Garcia I, Rabbitts TH. The LIM domain: a new structural motif found in zinc-finger-like proteins. Trends Genet. 1994; 10(9):315–20. Epub 1994/09/01. PMID: 7974745.

4. Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM. A census of human transcription factors: function, expression and evolution. Nat Rev Genet. 2009; 10(4):252–63. Epub 2009/03/11. https://doi.org/10.1038/nrg2538 PMID: 19274049.

5. Miller J, McLachlan AD, Klug A. Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. EMBO J. 1985; 4(6):1609–14. Epub 1985/06/01. PMID: 4040853.

6. Bouhouche N, Syvanen M, Kado CI. The origin of prokaryotic C2H2 zinc finger regulators. Trends Microbiol. 2000; 8(2):77–81. Epub 2000/02/09. PMID: 10664601.

7. Hall TM. Multiple modes of RNA recognition by zinc finger proteins. Curr Opin Struct Biol. 2005; 15 (3):367–73. Epub 2005/06/21. https://doi.org/10.1016/j.sbi.2005.04.004 PMID: 15963892.

8. Brown RS. Zinc finger proteins: getting a grip on RNA. Curr Opin Struct Biol. 2005; 15(1):94–8. Epub 2005/02/19. https://doi.org/10.1016/j.sbi.2005.01.006 PMID: 15718139.

9. Brayer KJ, Segal DJ. Keep your fingers off my DNA: protein-protein interactions mediated by C2H2 zinc finger domains. Cell Biochem Biophys. 2008; 50(3):111–31. Epub 2008/02/07. https://doi.org/10.1007/s12013-008-9008-5 PMID: 18253864.

10. Rubin GM, Yandell MD, Wortman JR, Gabor Miklos GL, Nelson CR, Hariharan IK, et al. Comparative genomics of the eukaryotes. Science. 2000; 287(5461):2204–15. Epub 2000/03/24. PMID: 10731134.

11. Lee MS, Gippert GP, Soman KV, Case DA, Wright PE. Three-dimensional solution structure of a single zinc finger DNA-binding domain. Science. 1989; 245(4918):635–7. Epub 1989/08/11. PMID: 2503871.

12. Stubbs L, Sun Y, Caetano-Anolles D. Function and Evolution of C2H2 Zinc Finger Arrays. Subcell Biochem. 2011; 52:75–94. Epub 2011/05/11. https://doi.org/10.1007/978-90-481-9069-0_4 PMID: 21557079.

13. Laity JH, Dyson HJ, Wright PE. DNA-induced alpha-helix capping in conserved linker sequences is a determinant of binding affinity in Cys(2)-His(2) zinc fingers. J Mol Biol. 2000; 295(4):719–27. Epub 2000/02/05. https://doi.org/10.1006/jmbi.1999.3406 PMID: 10656784.

14. Wolfe SA, Nekludova L, Pabo CO. DNA recognition by Cys2His2 zinc finger proteins. Annu Rev Biophys Biomol Struct. 2000; 29:183–212. Epub 2000/08/15. https://doi.org/10.1146/annurev.biophys.29.1.183 PMID: 10940247.

15. Gupta A, Christensen RG, Bell HA, Goodwin M, Patel RY, Pandey M, et al. An improved predictive recognition model for Cys(2)-His(2) zinc finger proteins. Nucleic Acids Res. 2014; 42(8):4800–12. Epub 2014/02/14. https://doi.org/10.1093/nar/gku132 PMID: 24523353.

16. Iuchi S. Three classes of C2H2 zinc finger proteins. Cell Mol Life Sci. 2001; 58(4):625–35. Epub 2001/05/22. https://doi.org/10.1007/PL00000885 PMID: 11361095.

17. Omichinski JG, Pedone PV, Felsenfeld G, Gronenborn AM, Clore GM. The solution structure of a specific GAGA factor-DNA complex reveals a modular binding mode. Nat Struct Biol. 1997; 4(2):122–32. Epub 1997/02/01. PMID: 9033593.

18. Chapman NR, Perkins ND. Inhibition of the RelA(p65) NF-kappaB subunit by Egr-1. J Biol Chem. 2000; 275(7):4719–25. Epub 2000/02/15. PMID: 10671503.

19. Seto E, Lewis B, Shenk T. Interaction between transcription factors Sp1 and YY1. Nature. 1993; 365 (6445):462–4. Epub 1993/09/30. https://doi.org/10.1038/365462a0 PMID: 8003102.

20. Lu H, Wang X, Li T, Urvalek AM, Yu L, Li J, et al. Identification of poly (ADP-ribose) polymerase-1 (PARP-1) as a novel Kruppel-like factor 8-interacting and -regulating protein. J Biol Chem. 2011; 286 (23):20335–44. Epub 2011/04/27. https://doi.org/10.1074/jbc.M110.215632 PMID: 21518760.

21. Bernard D, Bedard M, Bilodeau J, Lavigne P. Structural and dynamical characterization of the Miz-1 zinc fingers 5–8 by solution-state NMR. J Biomol NMR. 2013; 57(2):103–16. Epub 2013/08/27. https://doi.org/10.1007/s10858-013-9770-6 PMID: 23975355.

22. Brayer KJ, Kulshreshtha S, Segal DJ. The protein-binding potential of C2H2 zinc finger domains. Cell Biochem Biophys. 2008; 51(1):9–19. Epub 2008/02/21. https://doi.org/10.1007/s12013-008-9007-6 PMID: 18286240.

23. Poon IK, Jans DA. Regulation of nuclear transport: central role in development and transformation? Traffic. 2005; 6(3):173–86. Epub 2005/02/11. https://doi.org/10.1111/j.1600-0854.2005.00268.x PMID: 15702986.

24. Weis K. Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle. Cell. 2003; 112(4):441–51. Epub 2003/02/26. PMID: 12600309.

25. Kalderon D, Roberts BL, Richardson WD, Smith AE. A short amino acid sequence able to specify nuclear location. Cell. 1984; 39(3 Pt 2):499–509. Epub 1984/12/01. PMID: 6096007.

26. Dingwall C, Robbins J, Dilworth SM, Roberts B, Richardson WD. The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. J Cell Biol. 1988; 107(3):841–9. Epub 1988/09/01. PMID: 3417784.

27. Kosugi S, Hasebe M, Matsumura N, Takashima H, Miyamoto-Sato E, Tomita M, et al. Six classes of nuclear localization signals specific to different binding grooves of importin alpha. J Biol Chem. 2009; 284(1):478–85. https://doi.org/10.1074/jbc.M807017200 PMID: 19001369.

28. Yasuhara N, Oka M, Yoneda Y. The role of the nuclear transport system in cell differentiation. Semin Cell Dev Biol. 2009; 20(5):590–9. Epub 2009/05/26. https://doi.org/10.1016/j.semcdb.2009.05.003 PMID: 19465141.

29. Kelley JB, Talley AM, Spencer A, Gioeli D, Paschal BM. Karyopherin alpha7 (KPNA7), a divergent member of the importin alpha family of nuclear import receptors. BMC Cell Biol. 2010; 11:63. Epub 2010/08/13. https://doi.org/10.1186/1471-2121-11-63 PMID: 20701745.

30. Chook YM, Suel KE. Nuclear import by karyopherin-betas: recognition and inhibition. Biochim Biophys Acta. 2011; 1813(9):1593–606. Epub 2010/10/30. https://doi.org/10.1016/j.bbamcr.2010.10.014 PMID: 21029754.

31. Gorlich D, Kostka S, Kraft R, Dingwall C, Laskey RA, Hartmann E, et al. Two different subunits of importin cooperate to recognize nuclear localization signals and bind them to the nuclear envelope. Curr Biol. 1995; 5(4):383–92. Epub 1995/04/01. PMID: 7627554.

32. Conti E, Uy M, Leighton L, Blobel G, Kuriyan J. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. Cell. 1998; 94(2):193–204. Epub 1998/08/08. PMID: 9695948.

33. Kobe B. Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin alpha. Nat Struct Biol. 1999; 6(4):388–97. Epub 1999/04/14. https://doi.org/10.1038/7625 PMID: 10201409.

34. Bayliss R, Littlewood T, Stewart M. Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. Cell. 2000; 102(1):99–108. Epub 2000/08/10. PMID: 10929717.

35. Saijou E, Itoh T, Kim KW, Iemura S, Natsume T, Miyajima A. Nucleocytoplasmic shuttling of the zinc finger protein EZI Is mediated by importin-7-dependent nuclear import and CRM1-independent export mechanisms. J Biol Chem. 2007; 282(44):32327–37. https://doi.org/10.1074/jbc.M706793200 PMID: 17848547.

36. Hatayama M, Tomizawa T, Sakai-Kato K, Bouvagnet P, Kose S, Imamoto N, et al. Functional and structural basis of the nuclear localization signal in the ZIC3 zinc finger domain. Hum Mol Genet. 2008; 17 (22):3459–73. https://doi.org/10.1093/hmg/ddn239 PMID: 18716025.

37. Huang SM, Huang SP, Wang SL, Liu PY. Importin alpha1 is involved in the nuclear localization of Zac1 and the induction of p21WAF1/CIP1 by Zac1. Biochem J. 2007; 402(2):359–66. Epub 2006/11/18. https://doi.org/10.1042/BJ20061295 PMID: 17109628.

38. Matheny C, Day ML, Milbrandt J. The nuclear localization signal of NGFI-A is located within the zinc finger DNA binding domain. J Biol Chem. 1994; 269(11):8176–81. PMID: 8132543.

39. Quadrini KJ, Bieker JJ. Kruppel-like zinc fingers bind to nuclear import proteins and are required for efficient nuclear localization of erythroid Kruppel-like factor. J Biol Chem. 2002; 277(35):32243–52. https://doi.org/10.1074/jbc.M205677200 PMID: 12072445.

40. Li JZ, Chen X, Gong XL, Hu HY, Shi D, Lu YM, et al. Identification of a functional nuclear localization signal mediating nuclear import of the zinc finger transcription factor ZNF24. PLoS One. 2013; 8(11): e79910. Epub 2013/11/14. https://doi.org/10.1371/journal.pone.0079910 PMID: 24224020.

41. Rodriguez E, Aburjania N, Priedigkeit NM, DiFeo A, Martignetti JA. Nucleo-cytoplasmic localization domains regulate Kruppel-like factor 6 (KLF6) protein stability and tumor suppressor function. PLoS One. 2010; 5(9). https://doi.org/10.1371/journal.pone.0012639 PMID: 20844588.

42. Mingot JM, Vega S, Maestro B, Sanz JM, Nieto MA. Characterization of Snail nuclear import pathways as representatives of C2H2 zinc finger transcription factors. J Cell Sci. 2009; 122(Pt 9):1452–60. https://doi.org/10.1242/jcs.041749 PMID: 19386897.

43. Depping R, Schindler SG, Jacobi C, Kirschner KM, Scholz H. Nuclear transport of Wilms' tumour protein Wt1 involves importins alpha and beta. Cell Physiol Biochem. 2012; 29(1–2):223–32. https://doi.org/10.1159/000337603 PMID: 22415091.

44. Pandya K, Townes TM. Basic residues within the Kruppel zinc finger DNA binding domains are the critical nuclear localization determinants of EKLF/KLF-1. J Biol Chem. 2002; 277(18):16304–12. https://doi.org/10.1074/jbc.M200866200 PMID: 11844803.

45. Fernandez-Martinez J, Brown CV, Diez E, Tilburn J, Arst HN Jr., Penalva MA, et al. Overlap of nuclear localisation signal and specific DNA-binding residues within the zinc finger domain of PacC. J Mol Biol. 2003; 334(4):667–84. Epub 2003/11/26. PMID: 14636595.

46. Ito T, Azumano M, Uwatoko C, Itoh K, Kuwahara J. Role of zinc finger structure in nuclear localization of transcription factor Sp1. Biochem Biophys Res Commun. 2009; 380(1):28–32. Epub 2009/01/14. https://doi.org/10.1016/j.bbrc.2008.12.165 PMID: 19138671.

47. Ito T, Kitamura H, Uwatoko C, Azumano M, Itoh K, Kuwahara J. Interaction of Sp1 zinc finger with transport factor in the nuclear localization of transcription factor Sp1. Biochem Biophys Res Commun. 2010; 403(2):161–6. https://doi.org/10.1016/j.bbrc.2010.10.036 PMID: 20946882.

48. Choi S, Yamashita E, Yasuhara N, Song J, Son SY, Won YH, et al. Structural basis for the selective nuclear import of the C2H2 zinc-finger protein Snail by importin beta. Acta Crystallogr D Biol Crystallogr. 2014; 70(Pt 4):1050–60. PMID: 24699649.

49. Yoneda Y. Nucleocytoplasmic protein traffic and its significance to cell function. Genes Cells. 2000; 5 (10):777–87. PMID: 11029654.

50. Mehta TS, Lu H, Wang X, Urvalek AM, Nguyen KH, Monzur F, et al. A unique sequence in the N-terminal regulatory region controls the nuclear localization of KLF8 by cooperating with the C-terminal zinc-fingers. Cell Res. 2009; 19(9):1098–109. https://doi.org/10.1038/cr.2009.64 PMID: 19488069.

51. Peisach E, Pabo CO. Constraints for zinc finger linker design as inferred from X-ray crystal structure of tandem Zif268-DNA complexes. J Mol Biol. 2003; 330(1):1–7. PMID: 12818197.

52. UniProt C. The Universal Protein Resource (UniProt). Nucleic Acids Res. 2007; 35(Database issue): D193–7. https://doi.org/10.1093/nar/gkl929 PMID: 17142230.

53. Austen M, Luscher B, Luscher-Firzlaff JM. Characterization of the transcriptional regulator YY1. The bipartite transactivation domain is independent of interaction with the TATA box-binding protein, transcription factor IIB, TAFII55, or cAMP-responsive element-binding protein (CPB)-binding protein. J Biol Chem. 1997; 272(3):1709–17. PMID: 8999850.

54. Oka S, Shiraishi Y, Yoshida T, Ohkubo T, Sugiura Y, Kobayashi Y. NMR structure of transcription factor Sp1 DNA binding domain. Biochemistry. 2004; 43(51):16027–35. https://doi.org/10.1021/bi048438p PMID: 15609997.

55. Pavletich NP, Pabo CO. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science. 1991; 252(5007):809–17. PMID: 2028256.

56. Elrod-Erickson M, Benson TE, Pabo CO. High-resolution structures of variant Zif268-DNA complexes: implications for understanding zinc finger-DNA recognition. Structure. 1998; 6(4):451–64. PMID: 9562555.

57. Gersbach CA, Gaj T, Barbas CF 3rd. Synthetic zinc finger proteins: the advent of targeted gene regulation and genome modification technologies. Acc Chem Res. 2014; 47(8):2309–18. Epub 2014/06/01. https://doi.org/10.1021/ar500039w PMID: 24877793.

58. Kim MS, Kini AG. Engineering and Application of Zinc Finger Proteins and TALEs for Biomedical Research. Mol Cells. 2017; 40(8):533–41. Epub 2017/08/25. https://doi.org/10.14348/molcells.2017.0139 PMID: 28835021.

59. Gaj T, Guo J, Kato Y, Sirk SJ, Barbas CF 3rd. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. Nat Methods. 2012; 9(8):805–7. Epub 2012/07/04. https://doi.org/10.1038/nmeth.2030 PMID: 22751204.

60. Milletti F. Cell-penetrating peptides: classes, origin, and current landscape. Drug Discov Today. 2012; 17(15–16):850–60. Epub 2012/04/03. https://doi.org/10.1016/j.drudis.2012.03.002 PMID: 22465171.

61. Gaj T, Liu J, Anderson KE, Sirk SJ, Barbas CF 3rd. Protein delivery using Cys2-His2 zinc-finger domains. ACS Chem Biol. 2014; 9(8):1662–7. Epub 2014/06/18. https://doi.org/10.1021/cb500282g PMID: 24936957.