Human Molecular Genetics, 2008, Vol. 17, No. 22 ***3459–3473***
doi:10.1093/hmg/ddn239
Advance Access published on August 20, 2008

Functional and structural basis of the nuclear localization signal in the ZIC3 zinc finger domain

Minoru Hatayama[1], Tadashi Tomizawa[3], Kumiko Sakai-Kato[4], Patrice Bouvagnet[5], Shingo Kose[6], Naoko Imamoto[6], Shigeyuki Yokoyama[3,7], Naoko Utsunomiya-Tate[4], Katsuhiko Mikoshiba[2], Takanori Kigawa[3,8] and Jun Aruga[1,*]

[1]Laboratory for Behavioral and Developmental Disorders and [2]Laboratory for Developmental Neurobiology, RIKEN Brain Science Institute, Wako-shi, Saitama 351-0198, Japan, [3]Systems and Structural Biology Center (SSBC), Yokohama Institute, RIKEN, Yokohama-shi, Kanagawa 230-0045, Japan, [4]Research Institute of Pharmaceutical Sciences, Musashino University Nishitokyo-shi, Tokyo 202-8585, Japan, [5]INSERM, Equipe de Recherche Methodologique (ERM) 0107, Bron 69677, France, [6]Cellular Dynamics Laboratory, Discovery Research Institute, RIKEN, Wako-shi, Saitama 351-0198, Japan, [7]Graduate School of Science, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan and [8]Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, Yokohama 226-8502, Japan

Received May 26, 2008; Revised and Accepted August 7, 2008

Disruptions in ZIC3 cause heterotaxy, a congenital anomaly of the left–right axis. ZIC3 encodes a nuclear protein with a zinc finger (ZF) domain that contains five tandem C2H2 ZF motifs. Missense mutations in the first ZF motif (ZF1) result in defective nuclear localization, which may underlie the pathogenesis of heterotaxy. Here we revealed the structural and functional basis of the nuclear localization signal (NLS) of ZIC3 and investigated its relationship to the defect caused by ZF1 mutation. The ZIC3 NLS was located in the ZF2 and ZF3 regions, rather than ZF1. Several basic residues interspersed throughout these regions were responsible for the nuclear localization, but R320, K337 and R350 were particularly important. NMR structure analysis revealed that ZF1–4 had a similar structure to GLI ZF, and the basic side chains of the NLS clustered together in two regions on the protein surface, similar to classical bipartite NLSs. Among the residues for the ZF1 mutations, C253 and H286 were positioned for the metal chelation, whereas W255 was positioned in the hydrophobic core formed by ZF1 and ZF2. Tryptophan 255 was a highly conserved inter-finger connector and formed part of a structural motif (tandem CXW-C-H-H) that is shared with GLI, Glis and some fungal ZF proteins. Furthermore, we found that knockdown of Karyopherin α1/α6 impaired ZIC3 nuclear localization, and physical interactions between the NLS and the nuclear import adapter proteins were disturbed by mutations in the NLS but not by W255G. These results indicate that ZIC3 is imported into the cell nucleus by the Karyopherin (Importin) system and that the impaired nuclear localization by the ZF1 mutation is not due to a direct influence on the NLS.

INTRODUCTION

The zinc finger domains (ZFDs) of the Zic family proteins have been strongly conserved over the evolution of a broad range of eumetazoan animals (1) and may provide the structural basis for the essential roles of Zic family proteins in animal development (reviewed in 2,3, and references therein). In humans, there are five members of the Zic family that are involved in human congenital anomalies (reviewed in 4). One of them, ZIC3, causes X-linked heterotaxy (HTX1), which is a left–right axis disturbance that manifests as variable combinations of heart malformation, altered lung lobation, splenic abnormality and gastrointestinal malrotation (5,6).

*To whom correspondence should be addressed. Tel: +81 484679791; Fax: +81 484679792; Email: jaruga@brain.riken.jp

Zic3 was originally identified as a gene preferentially expressed in mouse cerebellum (7) and is considered to be a transcriptional regulator, based on its ability to bind DNA and activate transcription (8). Zic3-deficient mice have a variety of abnormalities that include not only heterotaxy, but also neural tube defects, skeletal patterning defects, cerebellar hypoplasia and abnormal behavior (9–12). *Xenopus* Zic3 has also been demonstrated to function in neural and left–right axis development (13,14). Hemilateral expression of the *Xenopus* Zic3 ZFD caused a left–right axis disturbance comparable to the misexpression of full-length ZIC3 (14), suggesting that the ZFD is critical for left–right axis determination.

Zic ZFDs are generally composed of five tandemly repeated C2H2 zinc finger (ZF) motifs (reviewed in 2). Comparison of amino acid sequences indicates that the Zic ZFD shows a substantial similarity to the ZF proteins of the Gli and Glis families, whose biological importance in vertebrate and ecdysozoan animals has been demonstrated. Zic, Gli and Glis—the three distinct ZF protein families are considered to be derived from a common ancestral gene (1,2). Although the ZFDs of Gli and Glis are also composed of five ZFs, the N-terminal (Nt) ZF (ZF1) of the Zic ZFD diverges from those of the Gli and Glis proteins (2). Zic ZF1 is unique in that it possesses more amino acid residues (6–38 amino acids) between the two cysteine residues of the C2H2 motif than the Gli and Glis ZF1s or any of the other ZFs (ZF2–5) in the Gli/Glis/Zic superfamily of proteins, which usually have 2–4 amino acids between the cysteine residues. Although the amino acid sequences between the two cysteine resides of Zic ZF1 are not strongly conserved within Zic family proteins as a whole, the second amino acid residue from the first cysteine is always tryptophan.

In human ZIC3, a missense mutation in this tryptophan residue (W255G) results in transposition of great artery (TGA), a congenital heart defect that might be an expression of a left–right laterality disturbance (15). ZIC3 W255G shows abnormal extranuclear localization and impaired transcriptional activation ability (15), indicating that W255 has an essential role in the functional integrity of the Zic3 protein. Mutations in cysteine 253 (C253S) or histidine 286 (H286R) in ZIC3 ZF1, which are found in heterotaxy patients, also result in extranuclear localization of the mutant ZIC3 protein (6). Together, these results raise the possibility that mutations in the evolutionarily conserved amino acid residues of ZF1 generally impair nuclear localization. However, the mechanism of nuclear localization, including the identity of the nuclear localization signal (NLS) in Zic family proteins, has not been investigated. This situation makes it difficult to clarify the role of ZF1 in the regulation of their subcellular distribution.

Some molecular components required for the nuclear transport of proteins have been identified (16,17). Macromolecules cannot disperse through the nuclear pore complexes and are actively transported into the nucleus by the importin system. Cargo proteins containing a classical NLS are bound by importin α, which is the adapter protein between the cargo proteins and importin β. The cargo protein–importin α–importin β heterotrimeric complex formed in the cytoplasm is targeted to the nuclear pore complexes and then is translocated into the nucleus. Importin-α proteins are categorized into three groups based on amino acid sequence identity. The first group includes human IMPORTIN α1 and mouse Karyopherin α2 (Kpna2); the second group includes human IMPORTIN α3 and α4, which are homologous to mouse Kpna4 and Kpna3, respectively; and the last group contains human IMPORTIN α5, α6 and α7 and mouse Kpna1 and Kpna6 (18) [Although it is confusing that the human and mouse family numbers do not match, we used the name 'Kpna' and the corresponding number designations because the Mouse Genome Informatics (MGI; http://www.informatics.jax.org/) and National Center for Biotechnology Information (NCBI; http://www.ncbi.nib/gov/) databases use this name]. There are five or six Kpna proteins that show overlapping but differential modes of NLS recognition without recognizing a clear consensus sequence. It would be interesting to see which Kpnas are used for ZIC3 nuclear transport since subtype switching of Kpna may be involved in neural differentiation (19).

Here we investigated how Zic family proteins are localized in the nucleus and the relationship of this mechanism to ZF1 in the ZIC3 ZFD. We identify the NLS in ZF2 and ZF3 and describe its spatial relationship to ZF1 and its interaction with nuclear import receptor proteins Kpna1 and Kpna6. In the course of this study, we determined the three-dimensional (3D) structure of the ZF1 to ZF4 region, providing the first result of an NMR-based protein structure for a Zic family protein.

RESULTS

Mutations of the conserved residues in ZIC3 ZF1 disturb its subcellular localization

Mutations of the conserved residues W255, C253 and H286 in ZF1 affect the subcellular distribution of ZIC3 (6,15). We first tested whether mutations in the other conserved residues in ZF1 also affect its subcellular distribution. For this purpose, we prepared novel mutant ZIC3 proteins containing C268S and H281R (mutations in the second cysteine and the first histidine residues in the C2H2 motif, respectively; Fig. 1) and compared their subcellular distribution to that of the wild-type ZIC3 protein and W255G, which causes extranuclear localization of ZIC3. Immunofluorescence staining of the NIH3T3 transfectants indicated that H281R strongly and C268S weakly increased the proportion of cells with extranuclear ZIC3 protein. These results indicate that mutations in the conserved C2H2 or tryptophan residues in ZF1 of ZIC3 generally inhibit the nuclear localization of ZIC3.

Influence of the ZF1 mutations on ZIC3 ZFD secondary structure

Circular dichroism (CD) spectroscopy is a widely used method for studying protein structures in solution (20,21). To investigate the possible structural alterations caused by the ZF1 mutations, we measured the CD spectrum of ZIC3 ZF1–5. The ZIC3 protein fragments were expressed in *Escherichia coli* and purified for the analysis (Fig. 2A). The CD spectra in the far UV range (200–250 nm) were measured in the presence of Zn^{2+} (see Materials and Methods). C253S and H286R showed a lower mean residue ellipticity (θ) around 200 nm than wild-type ZFD (Fig. 2B). This reduction indicates that the C253S and H286R ZFD contained more random coil

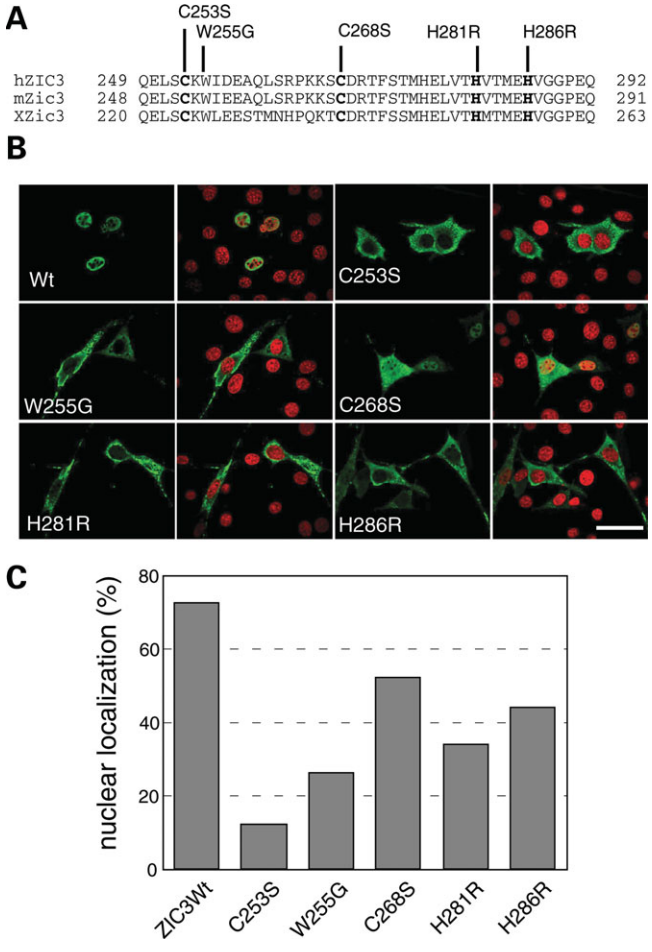

Figure 1. Intracellular localization of ZIC3 with missense mutations in evolutionarily conserved residues of ZF1. (**A**) ZF1 amino acid sequence of vertebrate Zic3 ZF1. Human (hZIC3), mouse (mZic3) and frog (xZic3) sequences are aligned. Cysteine and histidine residues in the C2H2 ZF motif are shown in bold. Known missense mutations (C253S, W255G and H286R) and newly generated ZF1 mutations (C268S and H281R) are indicated. (**B**) Representative images of the intracellular localization of wild-type and mutant ZIC3. Expression vectors for HA-ZIC3 or the HA-ZIC3 mutants were transfected into NIH3T3 cells. ZIC3 was stained with an anti-HA antibody. Wild-type ZIC3 was localized in nuclei, but the mutations in ZIC3 ZF1 increased the percentage of cells with cytoplasmic localization. Scale bar, 50 μm. (**C**) Nuclear localization percentages of each mutant. Photographs were taken of randomly selected fields under the same conditions, and all cells within each field were counted. All of the ZF1 mutants showed decreased nuclear localization percentages.

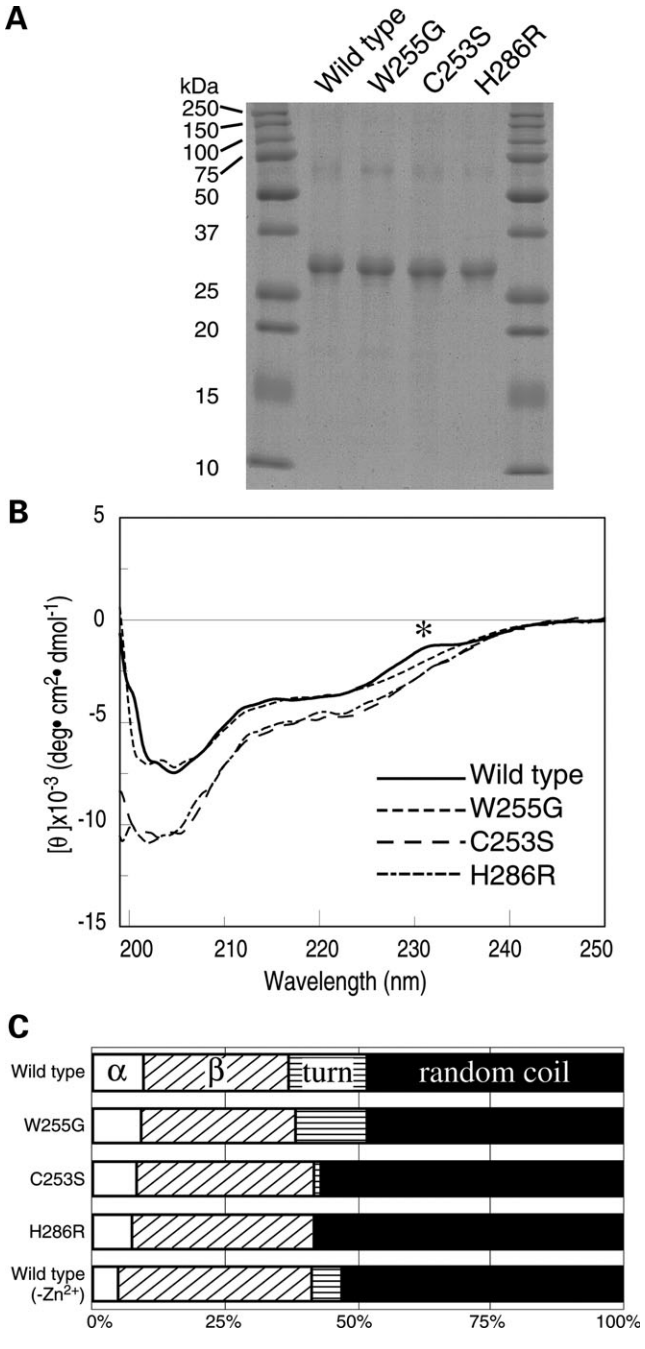

Figure 2. Circular dichroism (CD) spectroscopy analysis of ZIC3 ZF and its ZF1 mutants. (**A**) Coomassie brilliant blue-stained SDS–PAGE gel indicates the purity of the ZIC3 mutant proteins. (**B**) CD spectrum of ZF1 mutants. The solid line indicates wild-type spectrum and broken lines indicate mutants. The signal of C253S and H286R around 200 nm decreased, indicating that the amount of random coil had increased in the ZF1 mutants. The asterisk indicates a change at 230 nm that is probably due to the absence of the tryptophan indole side chain. (**C**) The relative proportions of the secondary structures were estimated as previously described (44).

than wild-type ZFD. In fact, the estimated random coil content of the wild type was 48.2% (Fig. 2C), whereas C253S and H286R had much higher percentages (57.0 and 57.8%, respectively). Random coil also increased in wild-type protein at zinc ion-free condition (53.0%). Reflecting this, ZFD C253S and H286R were easily aggregated during protein refolding after denaturation by urea (data not shown). The CD spectrum of W255G ZFD was similar to that of wild-type ZFD. However, a positive peak at 230 nm in the wild-type spectrum was not observed in the W255G spectrum. It is likely that this difference reflects the absence of the tryptophan residue, which has a peak absorbance around 230 nm. The similarity of the spectrum between the wild-type and W255G suggested that the secondary structure of the ZFD was not strongly affected by the W255G mutation. The abnormal conformation of the C253S and H286R proteins may have impaired their import into the nucleus after translation; however, we needed to further investigate the molecular pathogenesis of the W255G mutation.

A nuclear export inhibitor does not affect the subcellular localization of a ZIC3 ZF1 protein

We hypothesized that the extranuclear localization of the ZIC3 W255G was due to either acceleration of nuclear export or inhibition of nuclear import. To test the former possibility, we treated the ZIC3 transfectants with leptomycin B (LMB), an inhibitor of the CRM-1-dependent nuclear export pathway. LMB also causes the nuclear accumulation of RanBP1, a cytoplasmic factor required for Ran-dependent export pathways. Prolonged incubation of cells with LMB is therefore expected to inhibit most Ran-dependent export pathways (22). Consistent with this, LMB causes nucleocytoplasmic shuttling proteins such as Gli1 to accumulate in the nucleus (23). GLI1 effectively accumulated in nuclei in our experiments, but the subcellular localization of the W255G mutant was not affected even after prolonged incubation with LMB (Fig. 3). Therefore, nuclear export is probably not markedly accelerated in the ZIC3 ZF1 mutant; rather, impaired nuclear import is a more likely explanation for the extranuclear localization of the ZIC3 ZF1 mutant.

Mapping of the NLS in the ZIC3 protein

We next postulated that W255G may inhibit an NLS in ZF1 itself or nearby. To test this hypothesis, we first mapped the NLS in ZIC3 using a glutathione-*S*-transferase (GST)–enhanced green fluorescence protein (EGFP) fusion protein system. In this assay, the peptide fragments to be tested were inserted between GST and EGFP. GST was added to increase the total molecular mass of the fusion protein because small-molecular-weight proteins passively diffuse through the nuclear pores. Constructs containing various parts of ZIC3 were prepared (Fig. 4A) and transfected into NIH3T3 cells, and the subcellular localization of the EGFP signal was evaluated. The GST-EGFP fusion protein without any inserts (negative control) was localized in the nucleus in $<2\%$ of cells, whereas more than 90% of the cells showed nuclear localization when SV40-NLS or GLI1-NLS was inserted into the fusion proteins between GST and EGFP (positive controls, Fig. 4A and B).

In this assay system, neither the ZIC3 Nt nor the C-terminal (Ct) fragment showed any NLS activity. However, the GST-EGFP fusion protein containing the whole ZFD was located in cell nuclei in 73% of the cells (Fig. 4B). This percentage was less than that of SV40 or GLI1-NLS, but was comparable to that of full-length ZIC3 (70%, Fig. 1C). Starting with this construct, we generated serial deletion mutants to map the NLS in ZIC3 ZFD. ZF2 + 3 was the minimal unit that had nuclear localization activity (74%) comparable to the whole ZFD. Further division of ZF2 + 3 into single ZF units (ZF2 and ZF3) severely lowered the nuclear localization frequencies (Fig. 4B). ZF4 +5 showed weak nuclear localization (21%), whereas the other deletion mutants did not show substantial nuclear localization (Fig. 4B). We also confirmed that ZF2 + 3 has major NLS activity in HeLa cells and in the multipotent neural stem cell line MNS-70 by using the same NLS mapping analysis (Supplementary Material, Fig. S1). These results suggest that the ZF2 + 3 domain is widely utilized as an NLS in mammalian cells. GLI1-NLS is composed of bipartite basic residue clusters located in ZF5 and its Ct flanking region (23,24). This region also showed a strong NLS activity in our assay system (GLI1-NLS, Fig. 4B, Supplementary Material, Fig. S1). However, the corresponding Ct region in ZIC3 did not show clear NLS activity (ZF5 + Ct, Fig. 4B, Supplementary Material, Fig. S1). Conversely, when we assayed the NLS activity of GLI1 ZF2 + 3, only 4% of cells showed nuclear localization, suggesting that the location of the NLS is different between the ZIC3 and GLI1 ZFDs.

The basic amino acid residues lysine and arginine are important for NLS activity. Although basic amino acid residues were abundantly contained in the ZF2 and ZF3 region (13 of 65 residues), there were no basic amino acid residue clusters. We therefore sequentially replaced each lysine and arginine in ZF2 and ZF3 with alanine to investigate which amino acid residues are essential for the ZIC3 NLS. GST–ZF123–EGFP was chosen as the backbone construct for the substitution analysis (Fig. 5A), because a ZF1 mutation affects the nuclear localization. When the substitution mutants were produced in NIH3T3 cells, most showed greater or lesser decrements in their nuclear localization frequencies. However, only the R320A, K337A and R350A mutants of ZF123 showed a reduction to less than half of the wild-type nuclear localization frequency (Fig. 5B). This result was reproduced in MNS-70 and HeLa cells (Supplementary Material, Fig. S2).

To investigate the necessity of these basic residues for the nuclear localization of intact ZIC3 protein, we generated ZIC3 mutants containing R320A, K337A, R341A, K346A, K349A and R350A mutations (ΔNLS23), or R320A, K337A and R350A mutations (ΔNLS23′) and compared their NLS activity with those of the wild-type ZIC3 and the ZIC3 'ΔNLS45' mutant (identical to Δ*NLS* in ref. 25) in which four lysine/arginine residues in the ZF4 +5 region mutant are mutated. ΔNLS23, ΔNLS23′ and ΔNLS45 decreased the ratio of nuclear localization to 13, 19 and 53%, respectively, whereas 70% of wild-type ZIC3 was localized in the nucleus (Fig. 5C and D). LMB treatment did not significantly increase the nuclear localization of ΔNLS23 in NIH3T3 cells (Fig. 3), in agreement with the predicted impairment in the nuclear import process. Furthermore, the CD spectrum of ΔNLS23 ZFD was very similar to that of the wild-type ZFD (Supplementary Material, Fig. S3), indicating that the secondary structure was not strongly affected by the substitution mutations. These results collectively indicate that ZF2 + 3 is a major NLS for ZIC3.

Solution structure of ZIC3 ZF1–4

Having mapped the major NLS in ZIC3 to the ZF2 + 3 region, we investigated the relative positioning of ZF1 and the NLS by determining the solution structure of the ZIC3 ZF1–4 region (Fig. 6). The corresponding peptide was prepared using a cell-free protein synthesis system, and the structure was determined using NMR. The statistics on experimental constraints and structure quality as determined by PROCHECK-NMR (26) are summarized in Table 1. A total of 2108 NOE distance restraints and 96 user restraints for zinc fixation were used as the CYANA calculation input. Based on the values of the

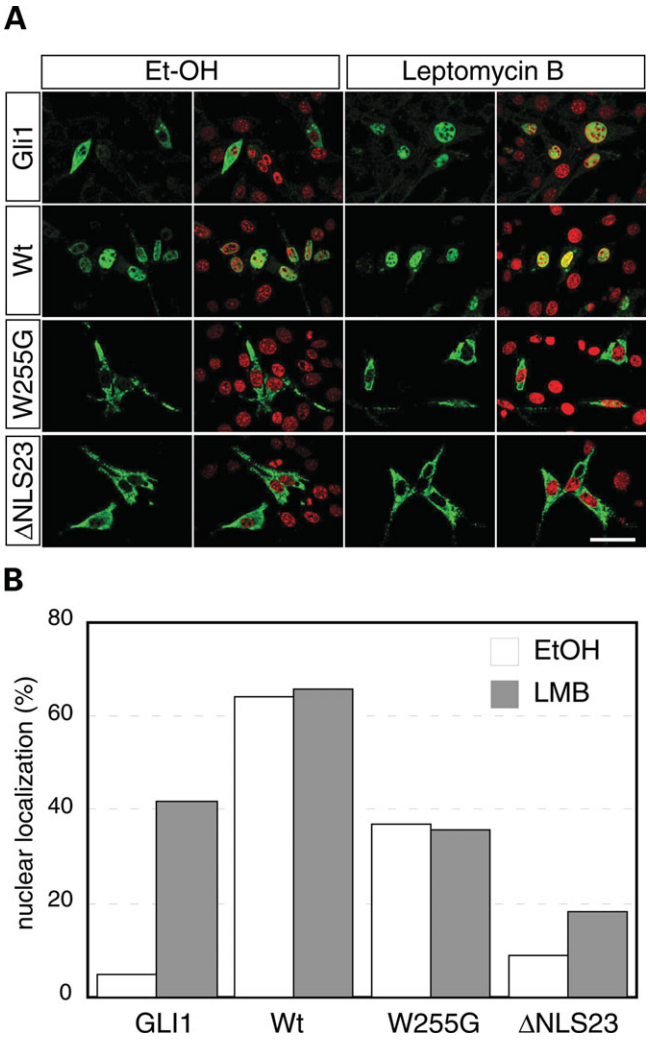

Figure 3. ZIC3 localization in LMB-treated cells. (**A**) Representative images of the intracellular localization of each protein. GLI1 was used as a control protein whose nuclear accumulation is enhanced by the LMB treatment. Scale bar, 50 μm. Et-OH, vehicle control. (**B**) Nuclear localization percentages. W255G did not accumulate in nuclei even with LMB treatment. The NLS-deficient mutant ΔNLS23 (described later) accumulated slightly in nuclei.

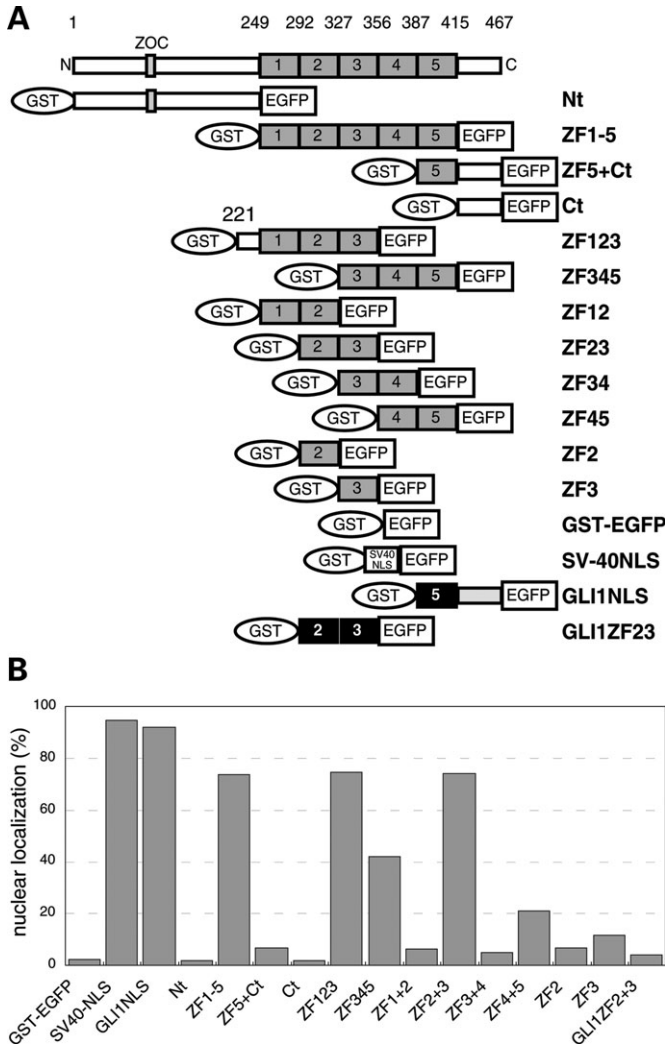

Figure 4. Mapping of ZIC3 NLS. (**A**) Scheme of the ZIC3 domain structure and the GST-EGFP fusion proteins used for the NLS mapping. The SV40-NLS and GLI1-NLS were used as positive controls. ZF123 contains 28 amino acids flanking the N-terminal side of ZF1 because this region was required to maintain the fusion protein stability (H.M. and J.A., unpublished data). (**B**) Nuclear localization percentages. ZIC3 ZF5 + C, a region analogous to the GLI1-NLS, did not show NLS activity. The whole ZF, ZF123 and ZF2 + 3 had strong NLS activity, and ZF345 and ZF4 + 5 showed weak NLS activity. ZF2 and ZF3 each showed low NLS activities when separated; ZF2 + 3 was therefore considered to be the minimal functional domain for NLS activity. ZOC, Zic Opa Conserved motif.

target function, and the Ramachandran plot (Table 1), we concluded that the structures are well-defined.

A global view of the structure (Fig. 6A) indicated that each ZF has a typical C2H2 ZF structure in which the two cysteines are in a β-sheet region and the two histidines are in an α-helical region, and together they are tetrahedrally coordinated to the zinc ion (27,28). ZF1 and ZF2 were folded into a single structural unit, whereas ZF3 and ZF4 existed as independent units. No domain–domain interaction was detected between ZF12 and ZF3, but a weak interaction of hydrophobic residues occurred between ZF3 and ZF4, as indicated by several interresidual NOEs between F332 in ZF3 and P357 in ZF4 (Supplementary Material, Fig. S4A). The soluble structure of the ZF12 single unit, ZF3 and ZF4 had fixed structures (Fig. 6B–E).

In the structural unit composed of ZF1 and ZF2, the secondary structure elements in the polypeptide were two helices (helix 1, amino acids 276–286; helix 2, amino acids 313–322) and a two-stranded antiparallel β-sheet (amino acids 251–255 and 268–272). These secondary structure elements were packed against each other by zinc coordination and by hydrophobic interactions among a group of hydrophobic residues: L278, V279 and V282 in helix 1; L315, V316 and I319 in helix 2; and W255, Y298 and W299 (Fig. 7A). W255 was situated in the center of the hydrophobic core formed by these residues. The NOEs between W255 and hydrophobic core-forming residues are shown in Supplementary Material, Figure S4B. This structural unit is similar to GLI1 (PDBID: 2GLI) ZF1 and ZF2 (Fig. 6F and G).

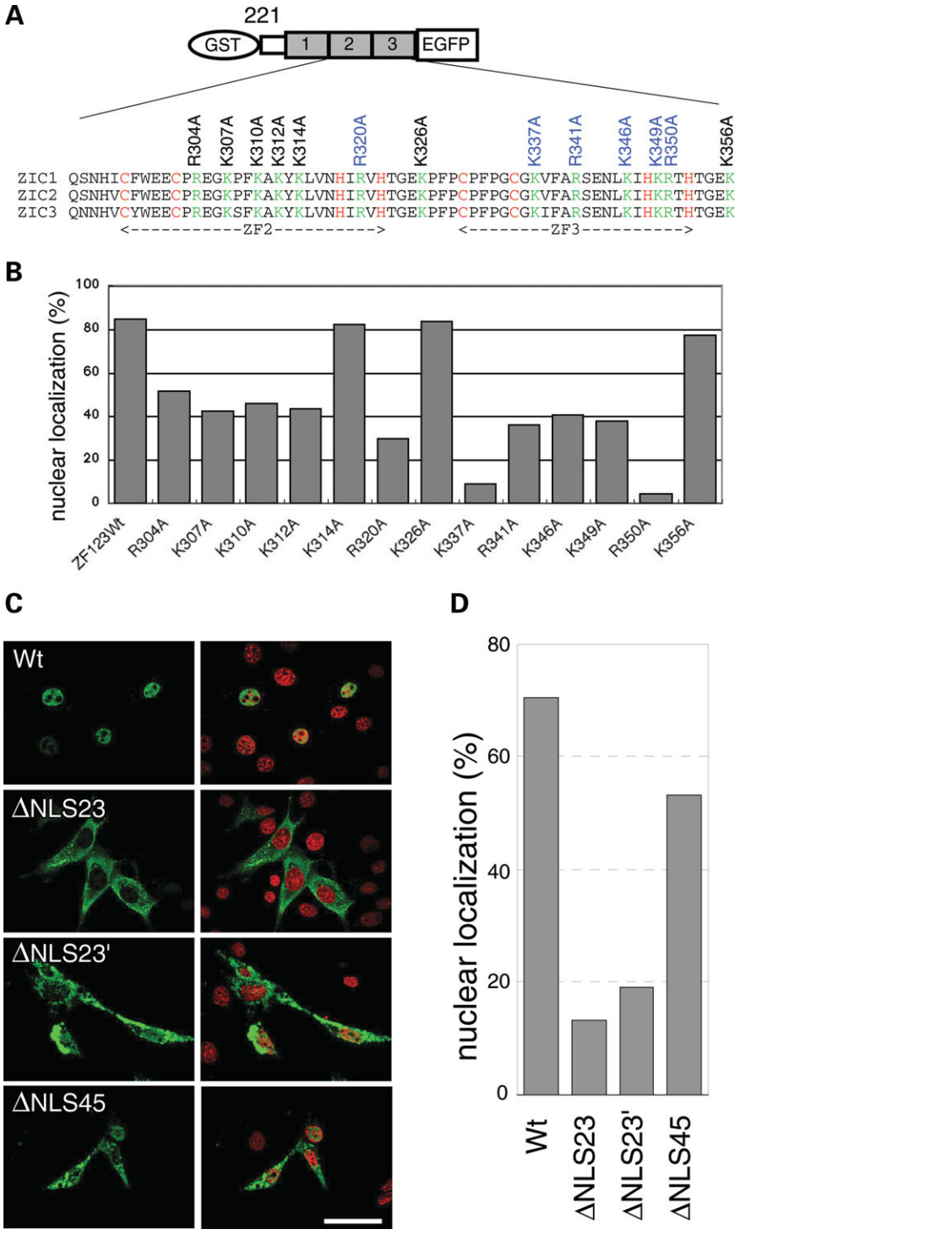

Figure 5. Mapping of the amino acid residues critical for the NLS activity in ZF123. (**A**) Schematic drawing of backbone constructs of ZF123 and amino acid sequences of human ZIC1, ZIC2 and ZIC3 ZF2 to ZF3 domain. Lysine and arginine residues (green text) are highly conserved in the ZIC proteins. Each mutant contains a substitution of alanine for the residue indicated in the sequence. Cysteine and histidine residues in the C2H2 ZF motif are shown in red. Positions of mutations in the 'ΔNLS23', which contains six amino acid substitutions, are indicated in blue. (**B**) Nuclear localization percentages. Most of the mutants had lower percentages of nuclear localization. The linker region between the ZFs (K326A and K356A) did not affect nuclear localization. (**C**) Representative images of the intracellular localization of the wild-type and the ΔNLS23, ΔNLS23′ (R320A, K337A, R350A) and ΔNLS45. Scale bar, 50 μm. (**D**) Nuclear localization percentages of HA-ZIC3 containing amino acid substitutions. Both the ZF4 +5 NLS mutant (ΔNLS45) and the ZF2 + 3 NLS mutant (ΔNLS23, ΔNLS23′) decreased nuclear localization.

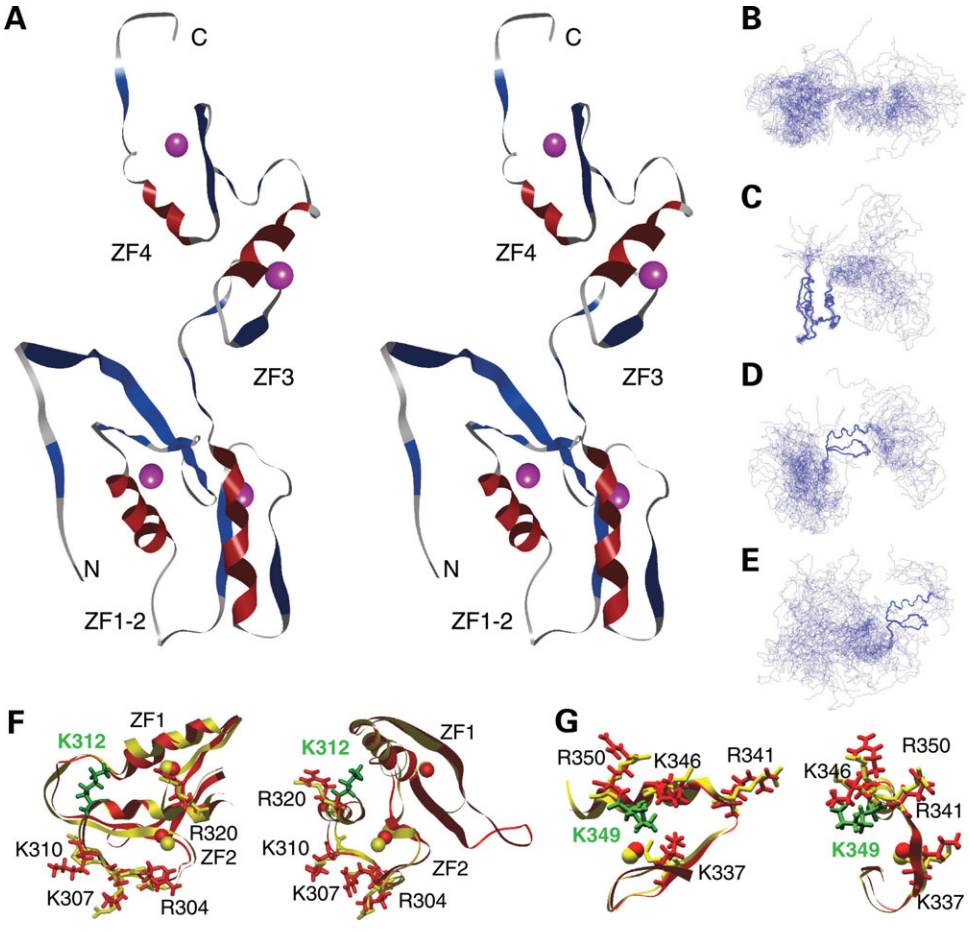

Figure 6. Three-dimensional structure of ZIC3 ZF1 to ZF4 (amino acids 238–386). (**A**) A stereo view of ZF1 to ZF4 (ribbon model). Magenta balls are zinc atoms. The top is the C-terminus and the bottom is the N-terminus. (**B**–**E**) Superimposition of the 20 models. (B), (C), (D) and (E) are aligned to maximize the overlap of the overall, ZF12, ZF3 and ZF4 structures, respectively. (**F** and **G**) Superimposition of GLI1 and ZIC3 ZFs. The red ribbon model is ZIC3, and the yellow ribbon represents GLI1 ZF structures from PDBID 2GLI. The red and yellow balls are zinc atoms of ZIC3 and GLI1. ZF1-2 (F) and ZF3 (G) structures are shown. Right models indicate horizontal 90° rotation of left models. The side chain of ZIC3 NLS forming lysine and arginine are shown in red. Green indicates the ZIC3 NLS-forming residues that are not conserved in GLI1 ZF.

We next examined the spatial relationship of the basic residues essential for the NLS. When the basic residues were mapped on a surface model (Fig. 7B), the NLS-forming lysine and arginine residues were located close to each other in each ZF, and many of their side chains were exposed on the molecular surface facing the same side.

ZIC3 NLS was bound by Kpna1 and Kpna6

Some NLSs are bound by the nuclear import adaptor protein Kpna (Importin α). To test whether the ZIC3 NLS is bound by Kpna, we assessed the physical interaction between *in vitro*-translated ZIC3 and the GST-Kpna fusion protein with a GST pull-down assay. Mouse Kpna2, Kpna4 and Kpna6 were chosen as representatives of each of the three groups of the Kpna family. Purified GST-Kpna proteins produced in *E. coli* (Fig. 8A) or cell lysates of 293T-expressed GST-Kpna proteins were used in these experiments. Among the three Kpna proteins, Kpna6 efficiently co-precipitated ZIC3 (Fig. 8B). Next we examined the physical interactions between Kpna6 and the ZIC3 NLS mutants ΔNLS23 (ZF2 + 3) and ΔNLS45 (ZF4 + 5) (Fig. 8C). In a binding assay using GST-Kpna expressed in 293T, Kpna6 precipitated wild-type ZIC3, but did not precipitate either ΔNLS23 or ΔNLS45 mutant. Kpna1, which belongs to the same group as Kpna6, showed a similar binding preference except that it weakly bound ΔNLS45.

To reveal the functional significance of the Kpna/ZIC3 interaction, we performed Kpna knockdown experiments by means of RNA interference (RNAi) for Kpna1 and Kpna6. Kpna1 and Kpna6 are strongly expressed in the NIH3T3 cells and in developing neural tissue, where mouse Zic3 is also expressed (19, M.H. and J.A., data not shown). The amount of Kpna1/6 proteins in siRNA-transfected NIH3T3 cells was ∼35% of that in cells transfected with the control pSilencer vector (Fig. 8D), as quantified from immunofluorescence staining. The results indicate that the knockdown was effective in the NIH3T3 cells.

Kpna knockdown using this experimental system impaired the nuclear localization of ZIC3 (Fig. 8E). The reduction in the ZIC3 nuclear localization ratio was not markedly reversed by the cotransfection of a Kpna1 or Kpna6 expression vector that contains the siRNA target sequence but was reversed by

Table 1. Structural statistics of ZIC3 ZF1 to ZF4 region (amino acids 245–286) based on 20 structures

NOE upper distance restraints		
Short-range ($\|i-j\|=0$)		668
Medium-range ($1 \le \|i-j\| \le 4$)		1070
Long-range ($\|i-j\|>4$)		370
Total		2108
CYANA target function value (Å^2)		0.40 ± 0.035
Number of restraint violations		
Distance restraint violations (>0.20 Å)		0
Dihedral angle restraint violations (>5.0°)		0
Rmsd from the averaged coodinates for each regions (Å)[a]		
	Backbone atoms	Heavy atoms
ZF12	0.66 ± 0.14	1.44 ± 0.15
ZF3	0.50 ± 0.17	1.17 ± 0.18
ZF4	0.28 ± 0.08	1.35 ± 0.19
Ramachandran plot for ordered region (%)		
Residues in most favored regions		84.1
Residues in additional allowed regions		15.8
Residues in generously allowed regions		1.1
Residues in disallowed regions		0.0

Data were analyzed by PROCHECK-NMR (26).
[a]Rmsd values were calculated by best fitted to the corresponding region.

a modified Kpna expression vector that lacks the siRNA target sequence (ΔsiRNAseq). These results eliminate the possibility that the RNAi had off-target effects and indicate that Kpna1 and Kpna6 are necessary for the ZIC3 nuclear transport.

Finally, we tested whether Kpna1/6 bind the ZIC3 ZF1 mutant W255G in the GST pull-down assay. W255G bound Kpna1 and Kpna6 as efficiently as wild-type ZIC3 (Fig. 8F), suggesting that the ZIC3 ZF1 mutation does not directly inhibit the interaction between the NLS and Kpna proteins.

DISCUSSION

NLS in ZIC3

In the course of this study, a paper dealing with the ZIC3 NLS was published (25). It seems beneficial to consider the NLS in ZIC3 by discussing their results with ours. We mapped significant NLS activities in ZF2 + 3 region and ZF4 +5 region using a GST–EGFP mapping system. The NLS activity of the ZF2 + 3 region was greater than that of ZF4 +5. Bedard *et al*. (25) used an EGFP-(test fragment)-LacZ construct and mapped the minimal NLS-containing region within amino acids 290–420, which contains the ZF2 + 3 region (amino acids292–356). In their subsequent analysis, Bedard *et al*. (25) focused on two subregions within the ZF4 + 5 region, 367–382 and 403–412, based on the presence of the multiple positively charged amino acid residues. Together with our comparative analysis using ΔNLS23 and ΔNLS45 mutants, we may conclude that these regions are cooperatively acting as the NLS for ZIC3 nuclear localization. This idea is also supported by our result that the disruption of the NLS in either ZF2 + 3 or ZF4 +5 was sufficient to disturb the interaction between ZIC3 and Kpna1/6.

Recently, an NLS was identified in Glis3 that, like the ZIC3 NLS, has a five-C2H2-type ZFD (29). The Glis3 NLS was determined to be located in its ZF4 because a mutant lacking the tetrahedral configuration of ZF4 also lacked nuclear localization. An NLS is also located in the ZF3 of PacC, a fungus transcription factor that contains a three-C2H2-type ZFD similar to ZF1–3 of the Gli/Glis/Zic ZFDs. When we aligned the NLSs of ZIC3 (ZF2 + ZF3), Glis3 (ZF3 + ZF4) and PacC (ZF2 + ZF3), we could see that the three basic residues that are critical for the ZIC3 nuclear localization (R320, K337 and R350) are conserved among these sequences (Fig. 9). In addition, F339 and T351 are conserved in ZIC3, Glis3 and PacC in addition to the ZF core-forming Cys/His and the sequence generally conserved in Gli-Krüppel superfamily proteins (TGEKP). The results suggest that the NLSs in these three proteins are conserved to some extent.

On the other hand, in Cubitus interruptus, a fly homolog of Gli, the NLS was mapped to ZF5 and the region flanking the C-terminus of the ZFD (23,24) that is not conserved in ZIC3. In our assay system, GLI1 (ZF2 + ZF3) does not possess NLS activity despite the apparent sequence similarity to ZIC3 (Fig. 4). Because the NLS in ZFD does not solely depend on the sequence similarity, we compared 3D structure models of ZIC3 and GLI1 ZF2–ZF3, focusing on the basic residues in this region. By the superimposition of their 3D structures (Fig. 6F and G), it is clear that some residues that contribute to ZIC3 nuclear localization are not conserved in the GLI ZFD. These residues may partly account for the differences in the NLSs of ZIC3 and GLI1 proteins. Furthermore, some residues adjacent to the critical residues (R320, K337 and R350) are conserved in the ZIC1/ZIC2/ZIC3/Glis3/PacC group, but not in GLI1 (closed circles in Fig. 9). They can also modulate the binding to the nuclear import adaptor proteins.

The ZIC3 NLS can be categorized as an interspersed-type NLS, different from the classic NLS that contains a short stretch or bipartite cluster(s) of basic residues (reviewed in 30,31). The interspersed-type NLS has also been found in other ZF proteins (32). However, so far as we know, there have been little detailed analyses of the interspersed NLSs in ZFD. This is probably due to technical difficulties, because many amino acids residues cooperatively and redundantly participate in the nuclear localization activity of the interspersed NLS (33–35). The results of this study may help to deepen our understanding of the structural basis of the NLS despite such difficulty.

In previous studies, the structure of the NLS–Karyopherin α complexes have been described for NLSs of the classic single-basic-cluster type [e.g. SV40 T antigen (36); androgen receptor (37)] and bipartite-cluster type NLS [e.g. Nucleoplasmin (38,39)]. There are two NLS binding sites in Kpna. The N-terminus of a bipartite NLS binds to the minor site and the C-terminus binds to the major site (36,38,40,41). When we compare the solution structure of ZIC3 and that of the *Xenopus* nucleoplasmin NLS (PDBID: 1EJY; 38,39), five out of the twenty ZIC3 solution structure models largely fit the nucleoplasmin structure in terms of its basic residue positioning (Fig. 10A). In our fitting, ZIC3 K310, K312, K337, K349 and R350 were positioned close to the nucleoplasmin K155, R156, K168, K169 and K170, respectively (root mean square between the backbone atoms of these residues was

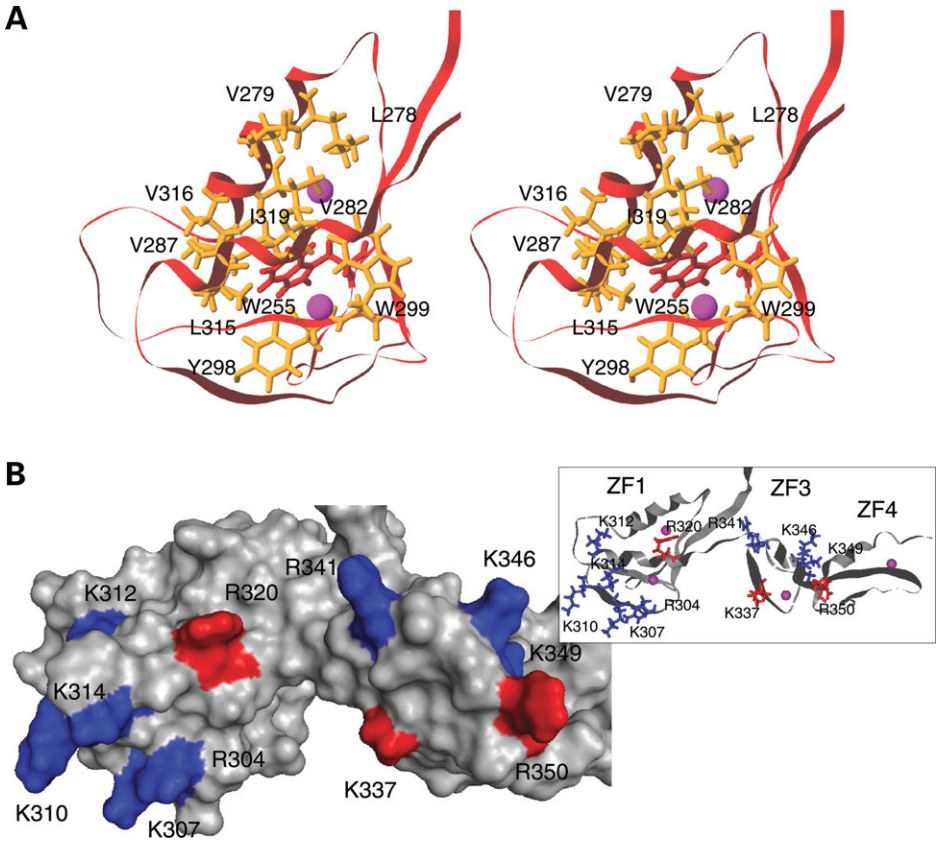

Figure 7. (**A**) Hydrophobic core of ZF1 and ZF2 (stereo view). Side chains of hydrophobic core-forming residues are shown in orange, and W255 is shown in red. (**B**) Surface model showing ZIC3 NLS. A corresponding ribbon model is drawn at the right side. NLS core residues (R320, K337 and R350) are indicated by red, and the other basic residues in ZF2 and ZF3 are in blue.

3.66 ± 0.32 Å). Based on this result, we further overlaid ZIC3 ZF1–3 structures with Kpna1 that has been co-crystallized with nucleoplasmin NLS (Fig. 10B). In this superimposition, ZIC3 NLS was assumed to be in the same position as that of the nucleoplasmin NLS in its binding to Kpna. The result suggested that the ZIC3 ZF1–3 structures fit the Kpna1 surface groove that forms the NLS binding sites in this positioning. There was no acute steric exclusion in the presumptive ZIC3 NLS–Kpna1 interaction. It was surprising that an interspersed-type NLS as in ZIC3 ZF can be fitted into the classical bipartite NLS. Although this fitting needs to be examined further, it might be a clue to consider nuclear import of many nuclear proteins that do not have classical NLS with one or two clusters of basic residues.

ZIC3 W255 as a novel inter-finger connector

This study provided the first solution structure of a ZFD in any Zic family protein. The four Nt ZFs possess typical globular ZF structures with a zinc ion at the core. Among the four ZFs analyzed, ZF1 and ZF2 are juxtaposed and together form a hydrophobic core. The structure predicts that not only W255 but also W299 and Y298 are constituents of this hydrophobic core. In a recent study (1), W255 and W299 were absolutely conserved among 45 Zic proteins covering a broad range of eumetazoan animals. The structural motif can be summarized as CXW–C–H–H–CXW–C–H–H. Hereafter, we call this structure a 'tandem CWCH2'.

The tandem CWCH2 motif can be found in the ZF1 and ZF2 of both GLI and GLIS protein families despite the apparent sequence divergence of the amino acids between the two cysteine residues. Pavletich and Pabo (27) determined the crystal structure of the GLI–DNA complex. They found that GLI ZF1 makes extensive protein–protein contacts with ZF2. In their structure (PDBID: 2GLI), the two tryptophan residues in ZF1 and ZF2 are wedged between the two ZF segments, and GLI ZF1 and ZF2 also form a united fold structure, similar to ZIC3. Although there are no reports concerning the GLIS ZF structure, Gli/Glis/Zic superfamily proteins could share the common feature that the ZF1 and ZF2 domains form a unified fold.

In fact, the tandem CWCH2 structure is not limited to Gli/Glis/Zic superfamily proteins. *Aspergillus* PacC, possesses a tandem CWCH2 structure in the ZFD that contains three C2H2 ZFs. Missense mutations in either tryptophan residue in PacC result in a partial loss-of-function phenotype (42). In addition, Zap1, a yeast zinc-sensing transcription factor, also possesses a tandem CWCH2 structure in its zinc-responsive domain, where the interaction between the two ZF is required for the high affinity binding of the zinc ion (43). The solution structure of Zap1 indicated that the two tryptophan residues provides numerous non-helical inter-finger contacts (PDBID: 1ZW8). Thus, the functional importance of the tandem

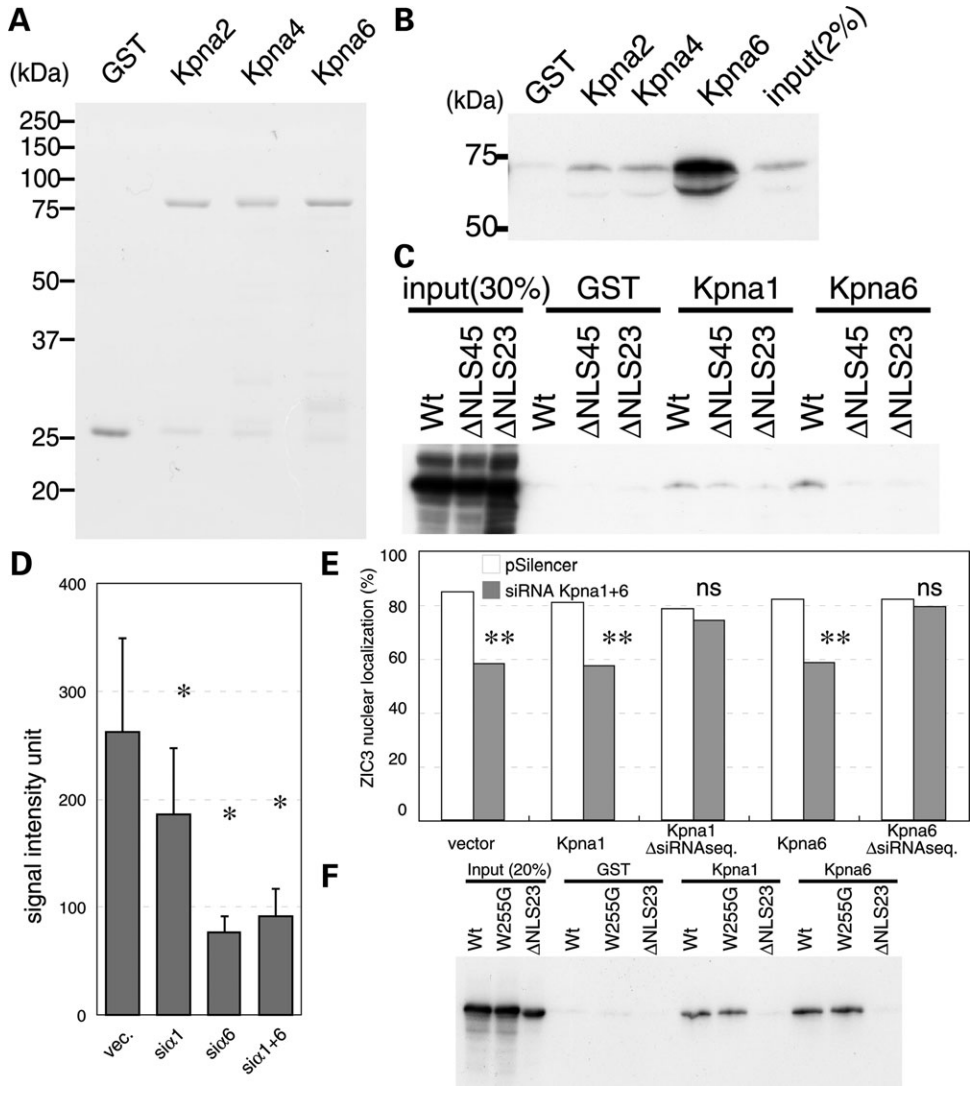

Figure 8. Physical interaction between Kpna and ZIC3. (**A**) Coomassie brilliant blue-stained SDS–PAGE gel indicates the purity of the GST-Kpna recombinant proteins. (**B**) GST pull-down analysis using recombinant GST-Kpna and *in vitro*-translated Myc-ZIC3. The pulled-down ZIC3 was detected with an anti-Myc-tag antibody (9E10). ZIC3 was co-precipitated with Kpna6. (**C**) GST pull-down assay using GST-Kpna expressed in 293T cells. *In vitro*-translated wild-type ZIC3 (Wt), a ZF4 +5 NLS mutant (ΔNLS45), and a ZF 2 + 3 NLS mutant (ΔNLS23) were used in this assay. Wild-type ZIC3 co-precipitated with Kpna1 and Kpna6. ΔNLS23 had little affinity for Kpna1 or Kpna6. ΔNLS45 showed slight binding to Kpna1, but little to Kpna6. (**D**) Kpna1 and Kpna6 protein amount in NIH3T3 cells. Kpna1 and/or Kpna6 siRNA expression vector or its empty vector was transfected into NIH3T3 and quantified by immunofluorescence staining as in Materials and Methods. Kpna1 siRNA significantly reduced the protein amount. (UNIT = signal intensity/background signal) $^{*}P < 0.05$ (*t*-test) (**E**) Effects of Kpna1 and Kpna6 double knockdown on ZIC3 nuclear localization. HA-ZIC3 was co-transfected into NIH3T3 with Kpna1 siRNA, FLAG-Kpna1, siRNA-resistant Kpna1 mutant (Kpna1ΔsiRNAseq), FLAG-Kpna6, siRNA-resistant Kpna6 (Kpna6ΔsiRNAseq) or parental empty vectors. Gray bar, Kpna1 siRNA vector; open bar, empty vector. (**F**) GST pull-down assay using GST-Kpna expressed in 293T cells. *In vitro*-translated ZIC3 and ZIC3 W255G were used for this assay. W255G bound to both Kpna1 and Kpna6 with affinity comparable to that of wild-type ZIC3. ΔNLS23 was used as negative control. $^{**}P < 5 \times 10^{-9}$; ns, not significant ($P = 0.31$) ($\chi^2$ test).

CWCH2 structure is supported by the analysis of these two fungus proteins as well as the ZIC3 W255G mutation. Further investigation of its distribution and biological significance may be beneficial for understanding protein structure–function relationships.

Nuclear import of ZIC3 and its impairment in the ZF1 mutant

This study revealed that all of the evolutionarily conserved residues in ZF1—C253, W255, C268, H281 and H286 (1)—are required for nuclear localization, but ZF1 is not itself the NLS. Zinc ion has an important role in ZF protein folding. Addition of zinc ion increases α helix and decreases random coil content in ZIC3 ZF (Fig. 2C), and these changes were consistent with other C2H2 ZFs (44–47). In our result, C253S and H286R similarly increased the random coil and decreased the turn content, suggesting that they are actually required to preserve structural integrity by chelating zinc ion. Bedard *et al.* (25) reported that H286R was not sensitive to LMB. Together, these findings may suggest that the ZF1 mutants missing the zinc-chelating residue are recognized as misfolded proteins

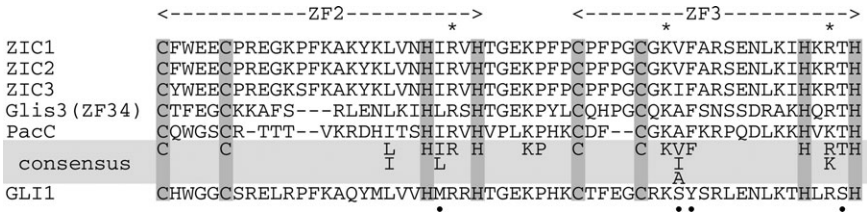

Figure 9. Amino acid sequence alignment of the NLS-containing regions of the selected ZF proteins. The ZF2 and ZF3 amino acid sequences of human ZIC1, ZIC2, ZIC3 and PacC and the mouse Glis3 ZF3 and ZF4 sequence are aligned. Core C2H2 residues are indicated with gray boxes. NLS core residues in ZIC3 (R320, K337 and R350) are indicated with asterisks over the top line. Consensus residues are indicated at the bottom.

and trapped by a molecular chaperone or degradation system (48,49).

In the ZIC3 W255G mutant, derangement of the CD spectrum was not obvious in comparison to the C253S and H286R mutants, indicating that the W255G mutation causes little change in the secondary structure. Furthermore, the interactions between the NLS and Kpna1/Kpna6 were not affected by the W255G substitution, excluding the possibility of a direct impairment of the NLS–Kpna1/Kpna6 interaction. However, the possibility remains that the mutation may cause a structural change that cannot be uncovered by a CD analysis. Such a structural alteration may result in the trapping of the protein by molecular chaperones or degradation systems, or the impairment of processes after the Kpna-binding such as interactions with Karyopherin β (Importin β) and the nucleopore complex. This idea is supported by the facts that LMB treatment did not cause significant nuclear accumulation of W255G (Fig. 3) and W255G lowers the stability of the ZIC3 protein (15). Although the consequences of the W255G mutation are not known in detail at this point, it is clear that it is a pathogenic mutation that occurs in the newly identified inter-finger connector residue.

MATERIALS AND METHODS

Plasmid construction

The human ZIC3 cDNA and C253S, W255G and H268R mutants were previously described (15). ZIC3 C268S and H281R were generated by using an LA PCR *in vitro* mutagenesis kit (TakaraBIO) with appropriate primers. cDNAs encoding the wild-type and mutant ZIC3 were inserted into the pcDNA3.1 vector (Invitrogen), which contains an HA-epitope tag (15), or the pCS2+Myc tag (MT) vector (50). For NLS mapping, a pCS2+ vector (50) expressing GST– EGFP was constructed by inserting GST- and EGFP-encoding DNA fragments from pGEX4T1 (GE Healthcare) and pEGFP (Invitrogen), respectively, into pCS2+. The GST and EGFP coding regions were connected by a linker sequence (5′-CAT GGA TAT CGC A-3′). The SV40 NLS from pCS2+NLS MT, the GLI1 open reading frame (ORF) (51) and ZIC3 ORF fragments were also inserted into this vector. For the CD spectrum analysis, we generated cDNA fragments containing a wild-type or mutant ZF domain by PCR amplifying the wild-type and mutant ZIC3 F239–S427 region using the primers 5′-CCGGATCCTTCTTCCGTTATATG-3′ and 5′-CCGTCG ACGATTCATAGCCTGAAC-3′. The PCR fragments were digested with restriction enzymes and inserted into the *Bam*HI- and *Sal*I-cut pET29a+ vector (Novagen). To adjust the reading frame to the Ct His-tag, the *Not*I site in the vector was filled with the Klenow enzyme.

Mouse Kpna cDNA clones were obtained from Riken FANTOM clones (http://www.gsc.riken.go.jp/e/FANTOM/; 52). We used the mouse cDNAs because human and mouse Kpna amino acid sequences are highly conserved (Supplementary Material, Fig. S5). The clone IDs of Kpna1, -2, -3, -4 and -6 are 3010020J22, I920184H24, C330037N10, A130074P15 and D230009G09, respectively. The Kpna ORFs were cloned into the *Bam*HI site of the pCMV-Tag 2B expression vector (Stratagene), pCS2+ expression vector or pGEX 4T1 using following PCR primers:

Kpna1

5′-GGA TCC ATG TCC ACA CCA GG-3′
5′-GGA TCC TCA AAG CTG GAA AC-3′

Kpna2

5′-GGA TCC ATG TCC ACG AAC GAG-3′
5′-GGA TCC TTA GAA GTT AAA GG-3′

Kpna3

5′-GGA TCC ATG GCC GAG AAC CCC-3′
5′-GGA TCC TTA GAA ATT AAA TTC-3′

Kpna4

5′-GGA TCC ATG GCG GAC AAC GAG-3′
5′-GGA TCC CTA AAA CTG GAA CCC C-3′

Kpna6

5′-GGA TCC ATG GAG ACC ATG GC-3′
5′-GGA TCC TTA TAG CTG GAA GCC C-3′

A control vector that generates an siRNA-resistant Kpna1 and Kpna6 mRNA through additional site-directed mutagenesis of Kpna1 (pCMV-Kpna1mut) and Kpna6 (pCMV-Kpna6mut) were prepared by introducing five synonymous mutations into the siRNA-targeted sequence. The following primers were used for site-directed mutagenesis with an LA PCR *in vitro* mutagenesis kit (Takara-BIO).

Kpna1mut

5′-TGAGCGCGGTTACCAGCTG-3′
5′-CCACCGCTAGTAGCATTTGTG-3′

Kpna6mut

5′-AAAGATCGTACAGGTAGCCCTCAATG-3′

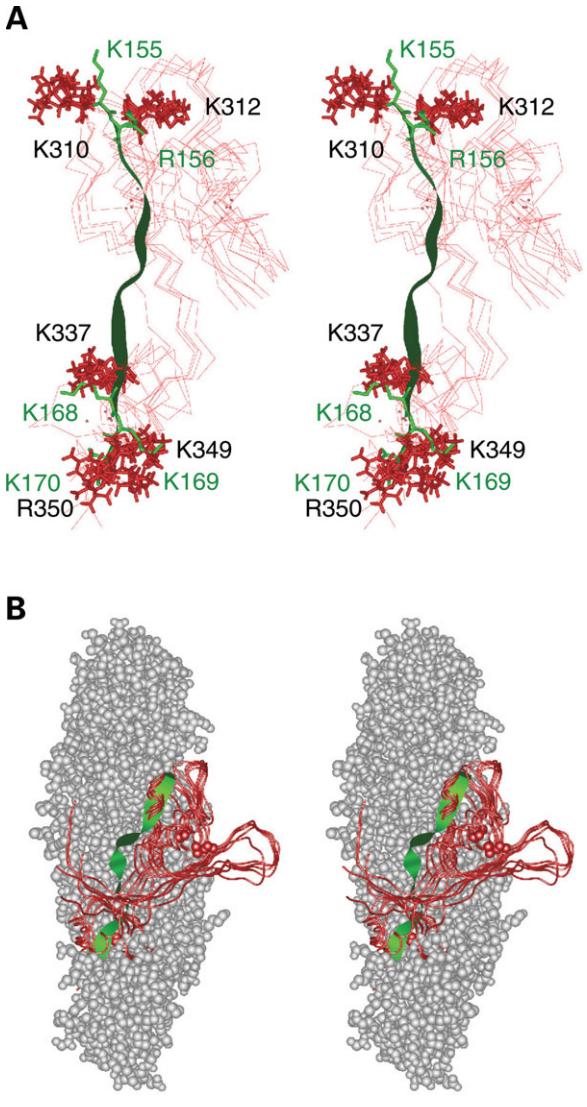

Figure 10. Structural comparison of ZIC3 NLS with reference proteins. (**A**) Comparison with a bipartite NLS from nucleoplasmin. Five of the 20 ZIC3 NMR structures fitted the nucleoplasmin structure (stereo view). Side chains of the some ZIC3 NLS-forming basic residues overlapped with those of nucleoplasmin. Corresponding side chains are drawn as red sticks (ZIC3) or green sticks (nucleoplasmin). Green ribbon model, nucleoplasmin; red wire model, ZIC3. (**B**) Presumptive superimposition of Kpna–nucleoplasmin crystal structure and ZIC3 ZF1–3 (stereo view). Gray surface model, Kpna; green ribbon model, nucleoplasmin; red tube model, ZIC3 ZF1–3.

Cell culture and transfection

NIH3T3, HeLa and MNS-70 cells were maintained in Dulbecco's modified Eagle's medium with 10% fetal bovine serum and a 1% antibiotics–antimycotics cocktail (Invitrogen). Lipofectamine 2000 (Invitrogen) was used for transfection according to the manufacturer's protocols. MNS-70 is the cell line established from rat neural stem cell (53).

Analysis of subcellular localization

NIH3T3 cells were seeded into 3.5-cm-diameter glass-bottom dishes (1×10^5 cells/dish). At 24–30 h after transfection, the cells were washed twice with phosphate-buffered saline (PBS), fixed with 4% paraformaldehyde in 0.1 M sodium phosphate buffer (pH 7.5) for 30 min, and immersed in a blocking buffer (1% bovine serum albumin, 0.1% Triton X-100 in PBS) for 1 h at room temperature. For immunofluorescence staining, the cells were incubated with an anti-HA antibody (3F10, Roche) at a dilution of 1:1000 in the blocking buffer for 1 h at room temperature. The cells were washed with the blocking buffer three times and incubated for 1 h with an Alexa 488-conjugated anti-rat IgG polyclonal antibody (Molecular Probes, Eugene, OR, USA). After three washes with PBS, the cells were mounted with Vectashield with DAPI (Vector Laboratories). The cells transfected with the GST–EGFP construct were observed after fixation and mounting with DAPI without immunostaining. The fluorescence images were obtained by confocal microscope and CCD camera (Olympus FV1000). The images were analyzed by using PhotoshopCS software (Adobe) to calculate the translocation score (54). More than 100 cells were counted in at least three independent experiments.

CD spectrum analysis

BL21 *E. coli* cells were transformed with the pET29a+ ZIC3-ZF vectors. His-tagged proteins were prepared according to the method of Sakai-Kato *et al*. (44). Purified proteins were dialyzed with 150 mM NaCl, 20 mM phosphate buffer (pH 7.0), 5 mM DTT and 0.03 mM $ZnCl_2$. Measurement of the CD spectrum was described previously (44).

LMB treatment

At 20–24 h after transfection, 1 ng/ml LMB (Alexis biochemicals, San Diego, CA, USA) or vehicle (0.1 μl/ml ethanol) was added to the medium. Cells were fixed 6 h after treatment and immunostained. The human GLI1 expression vector (55) was used as a positive control in this experiment.

RNA interference

The siRNA expression vectors against mouse Kpna1 [NPI-1 1104 and NPI-1 1221 (19)] were gifts from Dr Y. Yoneda (Osaka University). The Kpna6 siRNA expression vector was designed using the Ambion web site (http://www.ambion.com/techlib/misc/siRNA_finder.html). We chose 5′-AAATTG TGCAAGTGGCCCTCA-3′ as the target sequence for Kpna6 (Kpna6 1307). The vectors and the control parental vector (pSilencer, Ambion) were co-transfected with ZIC3 and the pCMV-Tag2B, pCMV-Kpna1, pCMV-Kpna1mut, pCMV-Kpna6 or pCMV-Kpna6mut vector (siRNA:ZIC3:Kpna = 3:6:1) using Lipofectamine 2000 (Invitrogen). The siRNA expression vectors NPI-1 1104, NPI-1 1221 and Kpna6 1307 were mixed 1:1:1. The cells were fixed 48 h later and subjected to the ZIC3 localization analysis. Endogenous Kpna1/6 expression was detected with immunofluorescence staining with the monoclonal rat antibody 2D9 (MBL, Nagoya; 1:100) and an anti-rat IgG Alexa 594-conjugated secondary antibody (1:1000). The 2D9 monoclonal antibody recognizes Kpna1/6 specifically, with low cross-reactivity against Kpna2 and 4 (data not shown). The fluorescence images were taken with a

confocal microscope (FV-1000, Olympus) and were analyzed with Image-J software. The quantified fluorescence signals were normalized to negative controls without the primary antibody.

Sample preparation for NMR analysis

A ZIC3 cDNA encoding ZF1, ZF2, ZF3 and ZF4 and their flanking regions (amino acids 245–386) was subcloned into the expression vector pCR2.1 (Invitrogen) as a fusion with an Nt His6 affinity tag and a tobacco etch virus (TEV) protease cleavage site. The actual construct contains seven extra residues (GSSGSSG) after the TEV cleavage site and six extra residues at the C-terminus (SGPSSG) that are derived from the expression vector. The $^{13}C/^{15}N$-labeled protein was expressed using a large-scale, cell-free system (56,57) and was purified using a chelating column, as described elsewhere (58). The purified protein was concentrated to 0.57 mM in 20 mM *d*-Tris/HCl (pH 7.0), 100 mM NaCl, 1 mM 1,4-DL-dithiothreitol-d_{10}, 0.02% NaN_3, 50 μM $ZnCl_2$, 1 mM IDA and 10% D_2O.

NMR measurements, resonance assignment and structural calculations

For chemical shift assignment, ^{15}N-HSQC, ^{13}C-HSQC, HNCO, HNCACO, HNCA, HNCOCA, HNCACB, CBCACONH, HHBHACONH, CCCONNH, HCCH-COSY, HCCH-TOCSY, CCH-TOCSY, HNHB and HNCOHB spectra were recorded on a Brucker AVANCE 600 spectrometer with a CryoProbe at 296 K. For structure determination, ^{15}N-edited Nuclear Overhauser Enhancement Spectroscopy (NOESY) and ^{13}C-edited NOESY spectra with 150 ms mixing times were recorded on a Brucker AVANCE 900 spectrometer at 296 K. Sequence-specific backbone chemical shifts were assigned using standard triple-resonance experiments (59). Side-chain chemical shift assignments were obtained from HBHACONH, CCCONNH and HCCH-TOCSY spectra (60–63). The distance restraints were obtained from ^{15}N-edited NOESY and ^{13}C-edited NOESY spectra.

All spectra were processed using NMRPipe (64), and the programs KUJIRA (65) and NMRView (66,67) were used to visualize the NMR spectra and chemical shift assignments. Automated NOE assignments and structure calculations with torsion angle dynamics were performed with the program CYANA 2.1 (68–71). A total of 100 structures were independently calculated and 20 conformers with the lowest target-function values were finally selected. The structure quality was evaluated with PROCHECK-NMR (26). The structural coordinates have been deposited in the RCSB Protein Data Bank under the accession code 2RPC.

GST pull-down assay

A GST pull-down assay was performed according to the methods of Kelly *et al*. (72), with modification. The GST-Kpna recombinant protein was purified from *E. coli* BL21 that were transformed by pGEX-4T1 derivatives or 293T cells that were transfected with a pCS2+ expression vector by using Lipofectamine 2000. The ZIC3 protein was prepared with an *in vitro* translation kit (Retic Lysate IVT™, Ambion). The capped mRNA used as a template was described in a previous study (15). GST fusion proteins were immobilized on glutathione-Sepharose 4B beads (GE Healthcare) and applied to *in vitro*-translated ZIC3 that had been pre-absorbed by free glutathione-Sepharose 4B beads in 1× transport buffer (20 mM HEPES–KOH, pH 7.3, 11 mM KAc, 5 mM NaAc, 2 mM MgAc, 1 mM EDTA, 1 mM DTT, 0.01% BSA and 0.1% NP-40). After overnight incubation at 4°C, the beads were washed three times with 1× transport buffer and resuspended in 1× SDS sample-loading buffer. The dissociated proteins were separated by SDS–PAGE and immunoblotted with an anti-Myc-tag antibody (9E10, Sigma).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* Online.

FUNDING

This work was supported by RIKEN BSI Funds and by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

ACKNOWLEDGEMENTS

We thank Yoshihiro Yoneda and Noriko Yasuhara for the Kpna1 knockdown plasmids, Masato Nakafuku for MNS-70 cells and RRC RIKEN BSI for technical assistance.

Conflict of Interest statement. None declared.

REFERENCES

1. Aruga, J., Kamiya, A., Takahashi, H., Fujimi, T.J., Shimizu, Y., Ohkawa, K., Yazawa, S., Umesono, Y., Noguchi, H., Shimizu, T. *et al.* (2006) A wide-range phylogenetic analysis of Zic proteins: implications for correlations between protein structure conservation and body plan complexity. *Genomics*, **87**, 783–792.
2. Aruga, J. (2004) The role of Zic genes in neural development. *Mol. Cell. Neurosci.*, **26**, 205–221.
3. Herman, G.E. and El-Hodiri, H.M. (2002) The role of ZIC3 in vertebrate development. *Cytogenet. Genome Res.*, **99**, 229–235.
4. Grinberg, I. and Millen, K.J. (2005) The ZIC gene family in development and disease. *Clin. Genet.*, **67**, 290–296.
5. Gebbia, M., Ferrero, G.B., Pilia, G., Bassi, M.T., Aylsworth, A., Penman-Splitt, M., Bird, L.M., Bamforth, J.S., Burn, J., Schlessinger, D. *et al.* (1997) X-linked situs abnormalities result from mutations in ZIC3. *Nat. Genet.*, **17**, 305–308.
6. Ware, S.M., Peng, J., Zhu, L., Fernbach, S., Colicos, S., Casey, B., Towbin, J. and Belmont, J.W. (2004) Identification and functional analysis of ZIC3 mutations in heterotaxy and related congenital heart defects. *Am. J. Hum. Genet.*, **74**, 93–105.
7. Aruga, J., Nagai, T., Tokuyama, T., Hayashizaki, Y., Okazaki, Y., Chapman, V.M. and Mikoshiba, K. (1996) The mouse zic gene family. Homologues of the *Drosophila* pair-rule gene odd-paired. *J. Biol. Chem.*, **271**, 1043–1047.
8. Mizugishi, K., Aruga, J., Nakata, K. and Mikoshiba, K. (2001) Molecular properties of Zic proteins as transcriptional regulators and their relationship to GLI proteins. *J. Biol. Chem.*, **276**, 2180–2188.
9. Carrel, T., Purandare, S.M., Harrison, W., Elder, F., Fox, T., Casey, B. and Herman, G.E. (2000) The X-linked mouse mutation Bent tail is associated with a deletion of the Zic3 locus. *Hum. Mol. Genet.*, **9**, 1937–1942.

10. Klootwijk, R., Franke, B., van der Zee, C.E., de Boer, R.T., Wilms, W., Hol, F.A. and Mariman, E.C. (2000) A deletion encompassing Zic3 in bent tail, a mouse model for X-linked neural tube defects. *Hum. Mol. Genet.*, **9**, 1615–1622.
11. Purandare, S.M., Ware, S.M., Kwan, K.M., Gebbia, M., Bassi, M.T., Deng, J.M., Vogel, H., Behringer, R.R., Belmont, J.W. and Casey, B. (2002) A complex syndrome of left–right axis, central nervous system and axial skeleton defects in Zic3 mutant mice. *Development*, **129**, 2293–2302.
12. Aruga, J., Ogura, H., Shutoh, F., Ogawa, M., Franke, B., Nagao, S. and Mikoshiba, K. (2004) Locomotor and oculomotor impairment associated with cerebellar dysgenesis in Zic3-deficient (Bent tail) mutant mice. *Eur. J. Neurosci.*, **20**, 2159–2167.
13. Nakata, K., Nagai, T., Aruga, J. and Mikoshiba, K. (1997) *Xenopus* Zic3, a primary regulator both in neural and neural crest development. *Proc. Natl. Acad. Sci. USA*, **94**, 11980–11985.
14. Kitaguchi, T., Nagai, T., Nakata, K., Aruga, J. and Mikoshiba, K. (2000) Zic3 is involved in the left–right specification of the *Xenopus* embryo. *Development*, **127**, 4787–4795.
15. Chhin, B., Hatayama, M., Bozon, D., Ogawa, M., Schon, P., Tohmonda, T., Sassolas, F., Aruga, J., Valard, A.G., Chen, S.C. *et al.* (2007) Elucidation of penetrance variability of a ZIC3 mutation in a family with complex heart defects and functional analysis of ZIC3 mutations in the first zinc finger domain. *Hum. Mutat.*, **28**, 563–570.
16. Gorlich, D. and Mattaj, I.W. (1996) Nucleocytoplasmic transport. *Science*, **271**, 1513–1518.
17. Goldfarb, D.S., Corbett, A.H., Mason, D.A., Harreman, M.T. and Adam, S.A. (2004) Importin alpha: a multipurpose nuclear-transport receptor. *Trends Cell Biol.*, **14**, 505–514.
18. Hogarth, C.A., Calanni, S., Jans, D.A. and Loveland, K.L. (2006) Importin alpha mRNAs have distinct expression profiles during spermatogenesis. *Dev. Dyn.*, **235**, 253–262.
19. Yasuhara, N., Shibazaki, N., Tanaka, S., Nagai, M., Kamikawa, Y., Oe, S., Asally, M., Kamachi, Y., Kondoh, H. and Yoneda, Y. (2007) Triggering neural differentiation of ES cells by subtype switching of importin-alpha. *Nat. Cell Biol.*, **9**, 72–79.
20. Pelton, J.T. and McLean, L.R. (2000) Spectroscopic methods for analysis of protein secondary structure. *Anal. Biochem.*, **277**, 167–176.
21. Kelly, S.M. and Price, N.C. (2000) The use of circular dichroism in the investigation of protein structure and function. *Curr. Protein Pept. Sci.*, **1**, 349–384.
22. Plafker, K. and Macara, I.G. (2000) Facilitated nucleocytoplasmic shuttling of the Ran binding protein RanBP1. *Mol. Cell. Biol.*, **20**, 3510–3521.
23. Sheng, T., Chi, S., Zhang, X. and Xie, J. (2006) Regulation of Gli1 localization by the cAMP/protein kinase A signaling axis through a site near the nuclear localization signal. *J. Biol. Chem.*, **281**, 9–12.
24. Wang, Q.T. and Holmgren, R.A. (1999) The subcellular localization and activity of *Drosophila* cubitus interruptus are regulated at multiple levels. *Development*, **126**, 5097–5106.
25. Bedard, J.E., Purnell, J.D. and Ware, S.M. (2007) Nuclear import and export signals are essential for proper cellular trafficking and function of ZIC3. *Hum. Mol. Genet.*, **16**, 187–198.
26. Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*, **8**, 477–486.
27. Pavletich, N.P. and Pabo, C.O. (1993) Crystal structure of a five-finger GLI-DNA complex: new perspectives on zinc fingers. *Science*, **261**, 1701–1707.
28. Pavletich, N.P. and Pabo, C.O. (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science*, **252**, 809–817.
29. Beak, J.Y., Kang, H.S., Kim, Y.S. and Jetten, A.M. (2008) Functional analysis of the zinc finger and activation domains of Glis3 and mutant Glis3(NDH1). *Nucleic Acids Res.*, **36**, 1690–1702.
30. Dingwall, C. and Laskey, R.A. (1991) Nuclear targeting sequences–a consensus? *Trends Biochem. Sci.*, **16**, 478–481.
31. Jans, D.A., Xiao, C.Y. and Lam, M.H. (2000) Nuclear targeting signal recognition: a key control point in nuclear transport? *Bioessays*, **22**, 532–544.
32. Pandya, K. and Townes, T.M. (2002) Basic residues within the Kruppel zinc finger DNA binding domains are the critical nuclear localization determinants of EKLF/KLF-1. *J. Biol. Chem.*, **277**, 16304–16312.
33. Yang, M., May, W.S. and Ito, T. (1999) JAZ requires the double-stranded RNA-binding zinc finger motifs for nuclear localization. *J. Biol. Chem.*, **274**, 27399–27406.
34. Huang, S.M., Huang, S.P., Wang, S.L. and Liu, P.Y. (2007) Importin alpha1 is involved in the nuclear localization of Zac1 and the induction of p21WAF1/CIP1 by Zac1. *Biochem. J.*, **402**, 359–366.
35. Quadrini, K.J. and Bieker, J.J. (2002) Krüppel-like zinc fingers bind to nuclear import proteins and are required for efficient nuclear localization of erythroid Krüppel-like factor. *J. Biol. Chem.*, **277**, 32243–32252.
36. Conti, E., Uy, M., Leighton, L., Blobel, G. and Kuriyan, J. (1998) Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell*, **94**, 193–204.
37. Cutress, M.L., Whitaker, H.C., Mills, I.G., Stewart, M. and Neal, D.E. (2008) Structural basis for the nuclear import of the human androgen receptor. *J. Cell Sci.*, **121**, 957–968.
38. Fontes, M.R., Teh, T. and Kobe, B. (2000) Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha. *J. Mol. Biol.*, **297**, 1183–1194.
39. Conti, E. and Kuriyan, J. (2000) Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha. *Structure*, **8**, 329–338.
40. Tarendeau, F., Boudet, J., Guilligay, D., Mas, P.J., Bougault, C.M., Boulo, S., Baudin, F., Ruigrok, R.W., Daigle, N., Ellenberg, J. *et al.* (2007) Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. *Nat. Struct. Mol. Biol.*, **14**, 229–233.
41. Fontes, M.R., Teh, T., Jans, D., Brinkworth, R.I. and Kobe, B. (2003) Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-alpha. *J. Biol. Chem.*, **278**, 27981–27987.
42. Espeso, E.A., Tilburn, J., Sanchez-Pulido, L., Brown, C.V., Valencia, A., Arst, H.N. Jr and Penalva, M.A. (1997) Specific DNA recognition by the Aspergillus nidulans three zinc finger transcription factor PacC. *J. Mol. Biol.*, **274**, 466–480.
43. Wang, Z., Feng, L.S., Matskevich, V., Venkataraman, K., Parasuram, P. and Laity, J.H. (2006) Solution structure of a Zap1 zinc-responsive domain provides insights into metalloregulatory transcriptional repression in Saccharomyces cerevisiae. *J. Mol. Biol.*, **357**, 1167–1183.
44. Sakai-Kato, K., Ishiguro, A., Mikoshiba, K., Aruga, J. and Utsunomiya-Tate, N. (2008) CD spectra show the relational style between Zic-, Gli-, Glis-zinc finger protein and DNA. *Biochim. Biophys. Acta*, **1784**, 1011–1019.
45. Parraga, G., Horvath, S.J., Eisen, A., Taylor, W.E., Hood, L., Young, E.T. and Klevit, R.E. (1988) Zinc-dependent structure of a single-finger domain of yeast ADR1. *Science*, **241**, 1489–1492.
46. Nomura, A. and Sugiura, Y. (2002) Contribution of individual zinc ligands to metal binding and peptide folding of zinc finger peptides. *Inorg. Chem.*, **41**, 3693–3698.
47. Frankel, A.D., Berg, J.M. and Pabo, C.O. (1987) Metal-dependent folding of a single zinc finger from transcription factor IIIA. *Proc. Natl. Acad. Sci. USA*, **84**, 4841–4845.
48. Hartl, F.U. and Hayer-Hartl, M. (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science*, **295**, 1852–1858.
49. Spiess, C., Meyer, A.S., Reissmann, S. and Frydman, J. (2004) Mechanism of the eukaryotic chaperonin: protein folding in the chamber of secrets. *Trends Cell Biol.*, **14**, 598–604.
50. Turner, D.L. and Weintraub, H. (1994) Expression of achaete-scute homolog 3 in *Xenopus* embryos converts ectodermal cells to a neural fate. *Genes Dev.*, **8**, 1434–1447.
51. Kinzler, K.W. and Vogelstein, B. (1990) The GLI gene encodes a nuclear protein which binds specific sequences in the human genome. *Mol. Cell. Biol.*, **10**, 634–642.
52. Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C. *et al.* (2005) The transcriptional landscape of the mammalian genome. *Science*, **309**, 1559–1563.
53. Kamitori, K., Machide, M., Tomita, K., Nakafuku, M. and Kohsaka, S. (2002) Cell-type-specific expression of protein tyrosine kinase-related receptor RYK in the central nervous system of the rat. *Brain Res. Mol. Brain Res.*, **104**, 255–266.

54. Galigniana, M.D., Scruggs, J.L., Herrington, J., Welsh, M.J., Carter-Su, C., Housley, P.R. and Pratt, W.B. (1998) Heat shock protein 90-dependent (geldanamycin-inhibited) movement of the glucocorticoid receptor through the cytoplasm to the nucleus requires intact cytoskeleton. *Mol. Endocrinol.*, **12**, 1903–1913.
55. Koyabu, Y., Nakata, K., Mizugishi, K., Aruga, J. and Mikoshiba, K. (2001) Physical and functional interactions between Zic and Gli proteins. *J. Biol. Chem.*, **276**, 6889–6892.
56. Kigawa, T., Yabuki, T., Yoshida, Y., Tsutsui, M., Ito, Y., Shibata, T. and Yokoyama, S. (1999) Cell-free production and stable-isotope labeling of milligram quantities of proteins. *FEBS Lett.*, **442**, 15–19.
57. Matsuda, T., Koshiba, S., Tochio, N., Seki, E., Iwasaki, N., Yabuki, T., Inoue, M., Yokoyama, S. and Kigawa, T. (2007) Improving cell-free protein synthesis for stable-isotope labeling. *J. Biomol. NMR*, **37**, 225–229.
58. Tochio, N., Umehara, T., Koshiba, S., Inoue, M., Yabuki, T., Aoki, M., Seki, E., Watanabe, S., Tomo, Y., Hanada, M. *et al.* (2006) Solution structure of the SWIRM domain of human histone demethylase LSD1. *Structure*, **14**, 457–468.
59. Cavanagh, J., Fairbrother, J.W., Palmer, G.A. and Skelton, J.N. (1996) *Protein NMR Spectroscopy, Principles and Practice*. Academic Press, Inc, San Diego.
60. Ikura, M., Kay, L.E. and Bax, A. (1990) A novel approach for sequential assignment of 1H, 13C, and 15N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry*, **29**, 4659–4667.
61. Bax, A., Vuister, G.W., Grzesiek, S., Delaglio, F., Wang, A.C., Tschudin, R. and Zhu, G. (1994) Measurement of homo- and heteronuclear J couplings from quantitative J correlation. *Methods Enzymol.*, **239**, 79–105.
62. Kay, L.E. (1997) NMR methods for the study of protein structure and dynamics. *Biochem. Cell Biol.*, **75**, 1–15.
63. Sattler, M., Schleucher, J. and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. NMR Spectrosc.*, **24**, 93–158.
64. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
65. Kobayashi, N., Iwahara, J., Koshiba, S., Tomizawa, T., Tochio, N., Güntert, P., Kigawa, T. and Yokoyama, S. (2007) KUJIRA, a package of integrated modules for systematic and interactive analysis of NMR data directed to high-throughput NMR structure studies. *J. Biomol. NMR*, **39**, 31–52.
66. Johnson, A.B. and Blevins, R. (1994) NMRView: a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR*, **4**, 603–614.
67. Johnson, B.A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol.*, **278**, 313–352.
68. Güntert, P., Mumenthaler, C. and Wuthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.*, **273**, 283–298.
69. Herrmann, T., Güntert, P. and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.*, **319**, 209–227.
70. Jee, J. and Güntert, P. (2003) Influence of the completeness of chemical shift assignments on NMR structures obtained with automated NOE assignment. *J. Struct. Funct. Genomics*, **4**, 179–189.
71. Güntert, P. (2004) Automated NMR structure calculation with CYANA. *Methods Mol. Biol.*, **278**, 353–378.
72. Kelly, K.F., Otchere, A.A., Graham, M. and Daniel, J.M. (2004) Nuclear import of the BTB/POZ transcriptional regulator Kaiso. *J. Cell Sci.*, **117**, 6143–6152.