Molecular cloning and characterization of ZFF29: a protein containing a unique Cys_2His_2 zinc-finger motif

Haruhiko ASANO*[1], Takashi MURATE†, Tomoki NAOE*, Hidehiko SAITO*[2] and George STAMATOYANNOPOULOS‡

*Department of Hematology, Nagoya University Graduate School of Medicine, Tsurumai-cho 65, Showa-ku, Nagoya, Japan 466-8550, †Nagoya University School of Health Science, Daiko-minami, 1-1-20, Higashi-ku, Nagoya, Japan 461-8673, and ‡Division of Medical Genetics, University of Washington, Seattle, WA 98195-7720, U.S.A.

We have cloned a gene, *ZFF29* (zinc-finger protein of human fetal liver erythroid cells 29), from human fetal liver erythroid cells. Two types of mature mRNA were identified and designated ZFF29a and ZFF29b. In human genome the *ZFF29* gene is on chromosome 9q, and the two forms are splice variants. There is a unique transcription start site, which predicts major mRNAs composed of 2485 bases for ZFF29a and 1801 bases for ZFF29b. The anticipated mRNAs were demonstrated in K562 cells, but not in any adult human tissues examined by Northern blotting. In the mouse, reverse transcription–PCR revealed that the ZFF29 mRNA is present in adult bone marrow and ovary at a higher level than in any other tissues examined. These findings suggest that ZFF29 proteins are expressed in embryonic/fetal erythroid tissues. The deduced polypeptide chains of ZFF29a and ZFF29b are composed of 306 and 350 amino acids respectively. A unique zinc-finger motif composed of two contiguous Cys_2His_2-type fingers is common to both forms of ZFF29. They are nuclear proteins and ZFF29b, but not ZFF29a, is an activator of erythroid gene promoters.

Key words: alternative splicing, fetal erythroid, transcriptional activator, zinc-finger protein.

INTRODUCTION

Expression of every gene is subject to individualized transcriptional regulation, which confers temporal and spatial specificity to the gene. This is achieved mainly at a transcriptional level by a number of gene-specific transcription factors. Recent advances in genome research have revealed that approximately 2000 hypothetical transcription factors are embedded in the whole human genome [1]. Among these factors, Cys_2His_2 zinc-finger proteins constitute the largest family. The Cys_2His_2 motif was first identified in *Xenopus* TFIIIA, and is usually composed of multiple repeats of the consensus sequence, C-X(2,4)-C-X(3)-[LIVMFYWC]-X(8)-H-X(3,5)-H, where X represents any amino acid. The zinc-finger domain enables proteins to bind to nucleic acids and to regulate expression of the target gene(s) (see review [2]).

We have searched for Cys_2His_2 zinc-finger proteins that potentially activate the transcription of fetal globin genes [3,4]. In screening of human fetal liver erythroid cells, we came across a zinc-finger motif. The motif, which is similar to that of yeast MIG1 [5], is unique in the human genome. In the present study we report that there are two mRNA species resulting from splicing variants, that they show non-ubiquitous expression patterns, and that at least one form functions as a transcriptional activator.

EXPERIMENTAL

Cell culture

K562 cells were cultured in RPMI 1640 medium supplemented with 10 % (v/v) fetal calf serum. COS-7 cells were cultured in Dulbecco's modified essential medium supplemented with 10 % (v/v) fetal calf serum.

cDNA cloning and plasmid construction

A cDNA fragment encoding the zinc-finger motif of ZFF29 (zinc-finger protein of human fetal liver erythroid cells 29) was obtained from human fetal liver erythroblasts (culture day 67) by PCR using a degenerate primer set, which has been described in detail elsewhere [3]. The 5′ and 3′ unknown sequences of the cDNA were obtained by RACE (rapid amplification of cDNA ends)–PCR from poly(A)$^+$ RNA of K562 cells using a Marathon cDNA amplification kit (Clontech) according to the manufacturer's instruction. ORFs (open reading frames) of ZFF29 cDNA were amplified by RT-PCR (reverse transcription–PCR). Subsequently they were cloned into a plasmid (T vector; Promega), and correct amplifications were verified by sequencing. The ORFs were cut out as a *Sma*I–*Spe*I fragment and subcloned into the eukaryotic expression vector pSG5DD [3] at *Bst*XI (blunted) and *Spe*I sites (to give pSG5/ZFF29a and pSG5/ZFF29b). pGFP/ZFF29 was constructed as follows: ZFF29 ORFs were cut as an *Eco*RI–*Bam*HI fragment from pSG5/ZFF29a and pSG5/ZFF29b, and were inserted into the same sites of vector pEGFPC3 (Clontech). Similarly, pFLAG-ZFF29 was constructed by inserting the ZFF29 ORFs of an *Eco*RI (blunted)–*Bam*HI fragment into pFLAG-CMV-2 (Sigma) cut with *Hin*dIII (blunted) and *Bam*HI. Information on primers and PCR conditions used for the amplification are available on request.

Abbreviations used: EKLF, erythroid Krüppel-like factor; FKLF, fetal Krüppel-like factor; GFP, green fluorescent protein; KLF, Krüppel-like factor; ORF, open reading frame; PI, propidium iodide; RACE, rapid amplification of cDNA ends; RT-PCR, reverse transcription–PCR; ZFF29, zinc-finger protein of human fetal liver erythroid cells 29.

[1] To whom correspondence should be addressed (email asanoh@med.nagoya-u.ac.jp).

[2] Present address: Nagoya National Hospital, Sannomaru 4-1-1, Naka-ku, Nagoya, Japan 460-0001.

The nucleotide sequence data reported will appear in DDBJ, EMBL, GenBank® and GSDB Nucleotide Sequence Databases under the accession numbers AY554164 (ZFF29a) and AY554165 (ZFF29b).

Transient transfection assays

K562 cells (approx. 5×10^5 in 1 ml of complete medium) were plated on a 12-well culture dish, and 0.5 μg of plasmid DNA (0.225 μg of activator, 0.225 μg of luciferase reporter and 0.05 μg of pSVβ-Gal) was transiently transfected into them using FuGENE™6 Transfection Reagent (Roche) according to the manufacturer's instructions (reagent/DNA, 3:1). After a 24 h incubation the cells were harvested, washed once with PBS, lysed in 200 μl of Reporter Lysis buffer (Promega), and subjected to one freeze/thaw cycle. Transfection assays were performed four times. Procedures of luciferase and β-galactosidase analyses have been described elsewhere [6].

Immunoblotting

Nuclear extracts were prepared as described previously [7]. Aliquots of 20 μg of the nuclear extracts were subjected to SDS/PAGE, transferred to a PVDF membrane and blotted overnight at 4 °C with 0.1 μg/ml anti-FLAG M2 antibody (Sigma) in PBS containing 0.2 % Tween 20 and 5 % skimmed milk. Signal detection was achieved with the Amersham Pharmacia ECL® (enhanced chemiluminescence) system according to the manufacturer's instructions.

Detection of the GFP (green fluorescent protein) signal

A 1 μg aliquot of pGFP/ZFF29s linearized with *Mlu*I was transfected using FuGENE™6 Transfection Reagent into 2 ml of COS-7 cells plated on a 6-well culture dish on the day before transfection at a concentration 1×10^5 cells/ml. In order to select stably transfected cells, G418 was added to the culture medium at a concentration of 800 μg/ml after 24 h of transfection. The green fluorescent signal was detected in G418-resistant cells cultured on a glass slide (Lab-Tek Chamber Slide; NUNC). Cells attached to the glass slide were washed twice with PBS and then treated with PI (propidium iodide) at a concentration of 0.5 μg/ml in PBS for 5 min at room temperature. Subsequently, after washing once with PBS, samples were examined by confocal microscopy (MRC-1024; Bio-Rad).

Northern blotting

Northern blotting was performed using a method described elsewhere [3]. Briefly, 4 μg of poly(A)$^+$ RNA extracted from the total RNA of K562 cells by passing it through an oligo(dT) column (Pharmacia) twice was run on a 1 % (w/v) agarose gel and immobilized on a nylon membrane (Hybond N$^+$; Amersham). The RNA was hybridized with ribo-probes. To make probes A and B, a *Sac*II–*Spe*I (blunted) fragment of ZFF29a and a *Sac*II–*Apa*I (blunted) fragment of ZFF29b respectively were subcloned into pGEM5-Zf(+) vector (Promega) digested by *Apa*I (blunted) and *Sac*II. To make probe C, a *Pst*I–*Sac*II fragment common to ZFF29a and ZFF29b was inserted into pGEM5-Zf(+) cut with *Pst*I and *Sac*II. Following digestion with *Bsm*FI, *Eco*O109I and *Eco*NI for probes A, B and C respectively, the antisense strand of the cDNA was transcribed with T7 RNA polymerase in the presence of [α-^{32}P]UTP.

Primer extension analysis

Primer extension analysis of RNA was carried out by a standard procedure [8] using a γ-^{33}P-labelled oligo DNA probe. The sequence was 5′-CAAAAATAACAACCTGCCGCCGCGGTGCCTGC-3′. A 4 μg portion of poly(A)$^+$ RNA of K562 cells was hybridized to the probe to determine the transcription start site of the *ZFF29* gene.

RT-PCR analysis

RT-PCR was performed as described previously [3]. Organs obtained from 10-month-old C57/B6 male mice were homogenized, and subsequently total RNA was extracted by a standard method [8]. Mouse total RNA from the testicle and the ovary was purchased (Ambion). Total RNA extracted from cell lines was treated with RQ1 DNase. Aliquots of 2 μg of total RNA were subjected to reverse transcription by MuMLV (Moloney murine leukaemia virus) reverse transcriptase (Invitogen) according to the manufacturer's recommendations in a 10 μl reaction volume. Aliquots of 0.5 μl of the reverse transcription samples, which had been diluted appropriately so as to give rise to uniform bands on amplification of the 28 S rRNA gene, were amplified by PCR in a 10 μl reaction mixture. The cycling conditions following initial denaturation at 95 °C for 3 min were as follows: for hZFF29a and hZFF29b, 95 °C for 35 s, 62 °C for 35 s and 72 °C for 40 s; for 28 S rRNA, 95 °C for 35 s, 56 °C for 35 s and 72 °C for 30 s; for murine ZFF29, 95 °C for 35 s, 58 °C for 35 s and 72 °C for 40 s; for β-actin, 95 °C for 35 s, 60 °C for 35 s and 72 °C for 45 s; for GATA-1, 95 °C for 35 s, 64 °C for 35 s and 72 °C for 35 s; for EKLF (erythroid Krüppel-like factor), 95 °C for 35 s, 53 °C for 35 s and 72 °C for 35 s. Following the PCR, 8.3 μl of the sample was run on an agarose gel and analysed. We performed each PCR multiple times with different numbers of cycles, and the results with unsaturated conditions are demonstrated. The primer sequences used are available on request.

RESULTS

ZFF29 is a two-zinc-finger-containing protein

We searched for an EKLF-like zinc-finger protein expressed in human fetal liver erythroid cells by the PCR-based approach that we had used to identify KLF11 {Krüppel-like factor 11; also known as FKLF (fetal Krüppel-like factor); [3]} and KLF13 (FKLF-2; [4]). In the course of the screening we came across a cDNA fragment encoding a novel zinc-finger motif. The same fragment was also found in a similar screening of cDNA prepared from the erythroid cell line K562 (results not shown). This gene was designated *ZFF29*. Unknown 5′ and 3′ regions of the cDNA were obtained by RACE/PCR. The 5′ RACE/PCR gave rise to a unique band, while 3′ RACE/PCR revealed two mRNA species containing a poly(A) tail. Sequencing analysis of the two reconstituted cDNAs revealed ORFs potentially encoding 306 and 350 amino acids. The respective subtypes of ZFF29 are designated as ZFF29a and ZFF29b, and the deduced polypeptide sequences are depicted in Figure 1(A). The molecular mass and pI are 32.9 kDa and 8.32 respectively for ZFF29a, and 38.4 kDa and 7.69 respectively for ZFF29b. They share two Krüppel-like Cys_2His_2 zinc fingers in tandem at the mid-portion. It looks as if there might be a 'third finger' contiguous to the second finger because of the presence of an apparently conserved inter-finger sequence, i.e. TGEKP (Thr-Gly-Glu-Lys-Pro), and cysteine and histidine residues (Figure 1A, broken underline). The 'third finger', however, is incomplete, as the number of amino acids between the second cysteine and the first histidine is nine (indicated by asterisks), while typically it should be 12. When this paper was in preparation, we found that an isologue of ZFF29b (ZNF367) had been predicted from the genome of the pufferfish, *Fugu rubripes* [9]. Although these authors reported ZNF367 as a zinc-finger

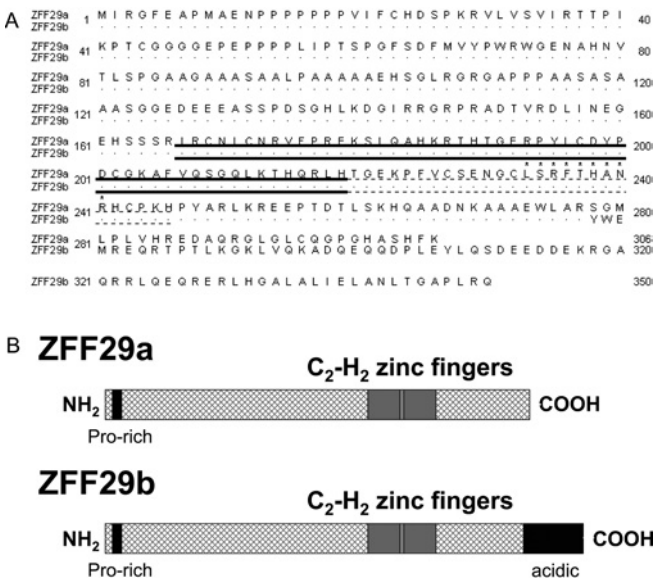

Figure 1 (A) Deduced polypeptide sequences of ZFF29a and ZFF29b, and (B) schematic representation of the deduced structures of ZFF29a and ZFF29b proteins

Residues that are common to the two polypeptide chains are indicated by dots in the ZFF29b sequence. The two zinc fingers are indicated by solid underlines. The broken underline shows an incomplete zinc finger. Note the short amino acid stretch indicated by asterisks between the conserved cysteine and histidine residues. (**B**) A common proline-rich domain, as well as zinc-finger domains and a C-terminal acidic domain of ZFF29b, are indicated.

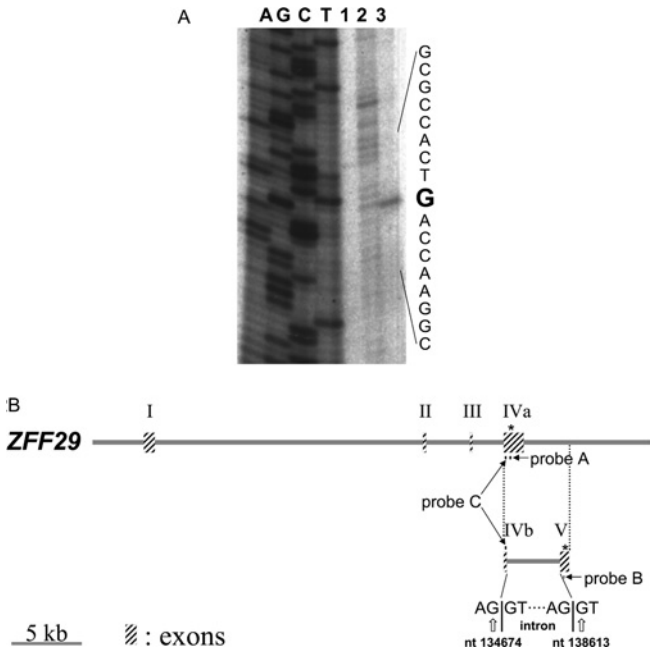

Figure 2 Transcription of the *ZFF29* gene

(**A**) Determination of the transcription start site of the *ZFF29* gene by primer extension analysis. Aliquots of 20 μg of total RNA (lane 2) and 4 μg of poly(A)$^+$ RNA (lane 3) extracted from K562 cells, as well as 20 μg of yeast tRNA (lane 1), were subjected to this assay. Reaction samples were run on the gel along with sequencing products (lanes A, G, C and T) of 5′ genomic DNA cloned into a plasmid vector using the same oligo DNA probe. Sequence reading, which represents the antisense strand, is depicted on the right. Note that poly(A)$^+$ RNA, but not total RNA, generated a distinct band at the guanine position indicated by a larger font size. (**B**) Transcription unit of the *ZFF29* gene. The number of each exon is shown above in Roman numerals. Asterisks indicate the stop codon, and the positions of probes used in the Northern analysis are indicated by solid rectangles. Note that an additional splicing in the fourth exon gives rise to another mRNA species (ZFF29b), the exon–intron junctions of which are also shown.

protein carrying three Cys_2His_2-type zinc fingers, the proposed 'third finger' fails to conform to the general structural rule of the Cys_2His_2 zinc-finger motif. Thus ZFF29 is most likely to be a two-zinc-finger-containing protein. There is a proline-rich region at the N-terminus, and the C-terminal domain, which is unique to the ZFF29b, is acidic (pI 4.9) (Figure 1B).

ZFF29 maps to chromosome 9q of the human genome

Sequence analysis and the subsequent database search mapped both *ZFF29* genes at human chromosome 9q22.1–31.1 (GenBank™ accession no. AL133477; hereafter nucleotide positions are expressed as the number in the genome sequence), indicating that the mRNAs of ZFF29a and ZFF29b are alternatively spliced products of a primary transcript. Of note is that ZFF29b is predicted as a novel protein by the Ensembl Human Genome Server (http://www.ensembl.org).

In order to identify the exact transcription units of ZFF29, we performed a primer extension analysis. As shown in Figure 2(A), we obtained a band from 4 μg of K562 cell poly(A)$^+$ RNA at nucleotide position 108 653, indicating a unique transcription start site. No distinct band, however, was observed with 20 μg of total RNA. The transcription units of the *ZFF29* gene are depicted in Figure 2(B). The transcripts of ZFF29a and ZFF29b are composed of four and five exons respectively; the first three exons are shared. Alternative splicing takes place in the fourth exon (see Figure 2B for the splice junctions). As a result, the shorter exon IVb and exon V, as well as exons I–III, comprise ZFF29b mRNA. Since the alternative splicing donor site precedes the stop codon, indicated by an asterisk in Figure 2(B), the C-terminal polypeptide sequences of ZFF29a and ZFF29b are different (Figure 1A). The theoretical length of the major mRNA without considering the poly(A) tail, is 2485 bases for ZFF29a and 1801 bases for ZFF29b.

Tissue-restricted and developmentally regulated expression of ZFF29 mRNA

The predicted sizes of ZFF29 mRNAs were confirmed by Northern blotting. Aliquots of 4 μg of poly(A)$^+$ RNA were hybridized with ZFF29a- and ZFF29b-specific RNA probes (probes A and B respectively; Figure 2B). These probes should detect different RNA species. The results are illustrated in Figure 3(A). Probe A generated a band at the 2.8 kb position (indicated by an arrow), as estimated by the positions of the 28 S (5025 bases) and 18 S (1868 bases) rRNAs. A size of 2.8 kb is reasonably close to that predicted for ZFF29a mRNA. Probe B generates multiple bands. The 2.0 kb band (indicated by an arrow in Figure 3A) was most intense, and its size is reasonably close to that predicted for ZFF29b mRNA. This 2.0 kb band, but not the 2.8 kb band generated by probe A, was also detected by probe C, suggesting that the mRNA expression of ZFF29a is much lower than that of ZFF29b. The uppermost bands (indicated by asterisks in Figure 3A) that were observed consistently with all probes appear at the position of the 28 S rRNA, suggesting that the bands are most probably the result of cross-hybridization of the probes with the residual 28 S rRNA. It is noteworthy that, according to the manufacturer's data, the purity of the poly(A)$^+$ RNA is $>90\,\%$, indicating that a significant amount of 28 S rRNA might have still

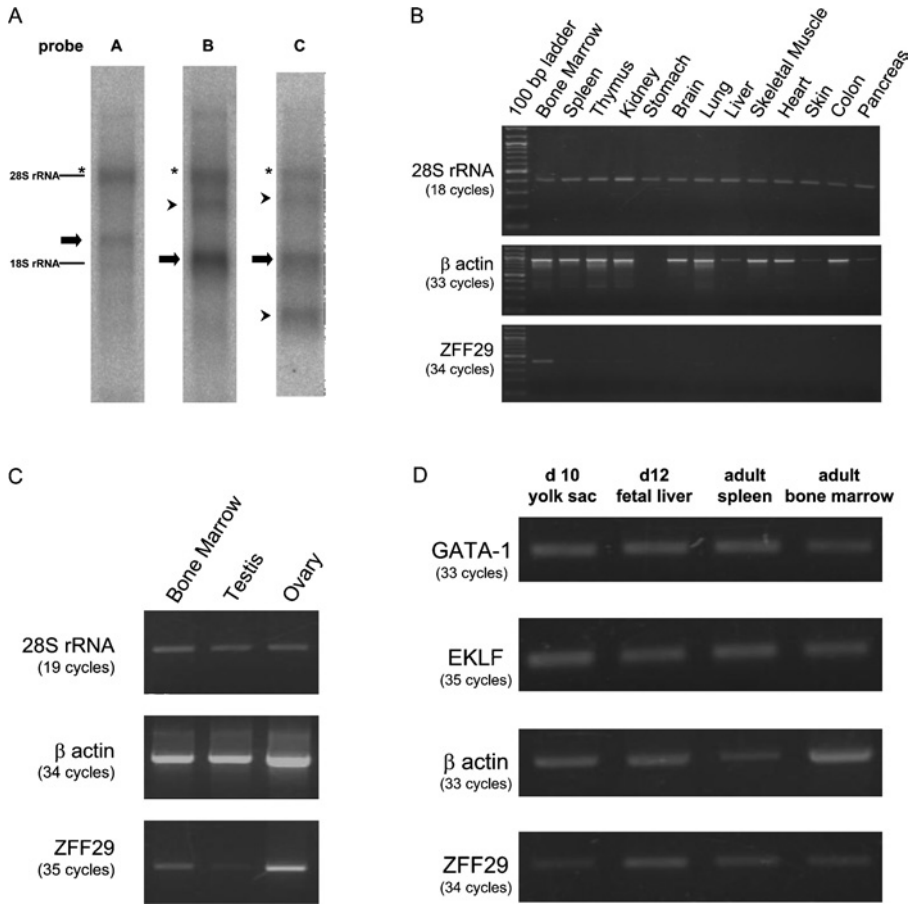

Figure 3 Expression of ZFF29 mRNA

(**A**) Transcripts of ZFF29a and ZFF29b were detected by Northern blotting. The riboprobes indicated above the lanes were hybridized to 4 μg of poly(A)$^+$ RNA extracted from K562 cells. The positions of 18 S and 28 S rRNA species are shown. Arrows indicate bands at the authentic positions of ZFF29a and ZFF29b. Bands marked with asterisks seem to be cross-hybridized to residual 28 S rRNA; the significance of bands marked by arrowheads is obscure. (**B**, **C**) ZFF29 expression in adult mouse tissues analysed by RT-PCR. Amplified cDNAs and numbers of PCR cycles are indicated on the left, and cDNA source tissues are shown above the lanes. Primers were derived from sequence common to ZFF29a and ZFF29b cDNAs. Note that intense bands were formed from the bone marrow and ovary cDNAs among those diluted to give rise to similar band intensities on amplification of the 28 S rRNA gene. (**D**) ZFF29 expression in mouse haematopoietic tissues analysed by RT-PCR. Amplified cDNAs and the numbers of PCR cycles are indicated on the left, and cDNA source tissues are shown above the lanes. The cDNA samples had been diluted to give rise to similar band intensities on amplification of the GATA-1 gene.

been present on the membrane. The significance of other bands (indicated by arrowheads in Figure 3A) is unknown. Thus the ZFF29 transcripts are present in K562 cells with the predicted sizes.

Subsequently, we examined the expression of ZFF29 in adult human tissues using commercially available membranes (MTN^{TM} Blots Human 12-Lane I and II; Clontech). Northern hybridizations were carried out multiple times. However, neither probe A nor probe B generated significant bands on these membranes (results not shown) on which portions of 2 μg of poly(A)$^+$ RNA from 24 different tissues had been immobilized, although the bands seemingly resulting from cross-hybridization with the residual 28 S rRNA were consistently produced with either probe. These results suggested that either (1) ZFF29 mRNA is absent from adult tissues; or (2) its expression is too low to be detected by the Northern analysis. To test these possibilities, we carried out RT-PCR analysis on adult mouse tissues. Prior to this experiment, the nucleotide sequence of mouse ZFF29b was determined (see below), and a mouse-specific primer set was designed. As shown in Figure 3(B), random-primed cDNAs from 13 different adult mouse tissues were diluted appropriately to generate uniform bands of amplified 28 S rRNA in unsaturated PCR conditions. These cDNAs, except for the stomach cDNA, were capable of generating bands by PCR for the β-actin gene, confirming the quality of the cDNAs. Then the samples were subjected to amplification of the ZFF29 cDNA. An obvious band was produced from the bone marrow cDNA (Figure 3B). In contrast with this, faint bands were detected in multiple tissues, e.g. the spleen, thymus, kidney and lung. Therefore low levels of ZFF29 are expressed in multiple tissues of the adult mouse, but the bone marrow displays the highest expression. Since the Fugu isologue of ZFF29b is expressed in the testis and the ovary [9], we further analysed murine ZFF29 expression among the sex glands and the bone marrow by RT-PCR. As shown in Figure 3(C), the ovary showed higher ZFF29 expression than the bone marrow or the testis. These results demonstrate that ZFF29 is expressed predominantly in the bone marrow and the ovary among adult mouse tissues.

In order to test whether a temporal expression pattern exists in erythroid cells, in which ZFF29 was originally identified, expression of ZFF29 mRNA in four different haematopoietic tissues, i.e. the yolk sac of embryos at 10 days' gestation, the liver of fetuses at 12 days' gestation, the adult spleen and the adult bone

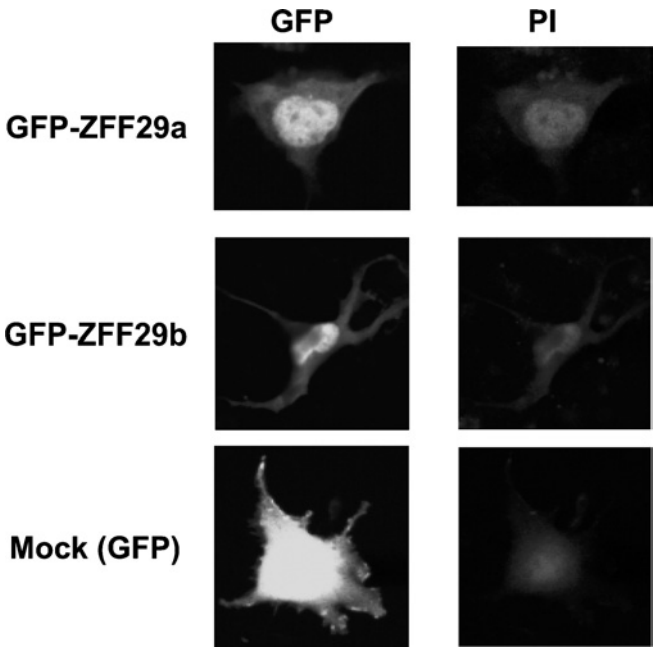

Figure 4 Nuclear localization of ZFF29 proteins

ZFF29 proteins were stably expressed as fusion proteins with GFP in COS-7 cells. After staining of nuclei with PI, cells were observed by confocal microscopy. Note that the fluorescence signals of GFP and PI in ZFF29-transfected cells obviously show the same localization. This is not the case in mock-transfected cells, in which the GFP signal is present diffusely in the cytoplasm.

marrow, was analysed by RT-PCR. The bands of ZFF29, along with those of two erythroid-specific genes, GATA-1 and EKLF, are illustrated in Figure 3(C). The cDNA samples were first diluted to give rise to uniform band patterns with amplification of GATA-1 gene; amplicons of EKLF were also uniform, but this was not the case with the amplification of the β-actin gene. In contrast with the patterns with GATA-1 and EKLF, amplicons of ZFF29 from fetal liver apparently gave rise to the most intense band among the four tissues tested, indicating that the ZFF29 mRNA is more abundant in the fetal liver compared with the yolk sac or adult haematopoietic tissues.

Transcriptional activities of ZFF29 proteins on erythroid gene promoters

Since both ZFF29a and ZFF29b contain a zinc-finger motif, which is generally considered to be a DNA-binding domain, it was reasonable to speculate that ZFF29 is localized in the nucleus and functions as a transcription factor. To test for nuclear localization, full-length cDNAs of ZFF29a and ZFF29b were inserted into a GFP (green fluorescent protein) expression vector, and the chimaeric proteins were expressed in COS-7 cells. In GFP–ZFF29a- and GFP–ZFF29b-transfected cells, the GFP signal was located in the nucleus, and the localization pattern was the same as that of the fluorescence signal generated by PI (Figure 4). In contrast, the GFP signal of mock-transfected cells was present diffusely in the cytoplasm (Figure 4). Thus both forms of ZFF29 are nuclear proteins.

Next we explored effects of the ZFF29 proteins on gene transcription. Since ZFF29s were cloned from erythroid cells, we tested their transcriptional activity on the promoters of five genes expressed in erythroid cells. Luciferase reporter constructs carrying the promoter of the 5-aminolaevulinate synthase, glycophorin B, porphobilinogen deaminase, ferrochelatase or GATA-1 genes [4] were transiently transfected into K562 cells with and without the ZFF29 expression construct. Luciferase counts were corrected by β-galactosidase activities. The luciferase activities in the absence of ZFF29 were taken to be 100%. The results are shown in Figure 5(A). Mean luciferase activities in the presence of ZFF29a and ZFF29b were: 85% and 290% respectively for 5-aminolaevulinate synthase; 108% and 177% for glycophorin B; 117% and 226% respectively for porphobilinogen deaminase; 92% and 511% respectively for ferrochelatase; and 80% and 410% respectively for GATA-1. Expression of the ZFF29 proteins was tested by using FLAG–ZFF29 constructs. It should be noted that the presence of the FLAG peptide did not affect the activities of ZFF29s on the ferrochelatase promoter (results not shown). Western analysis of FLAG–ZFF29a- and FLAG–ZFF29b-transfected K562 cells clearly demonstrated the expression of both types of ZFF29 protein (Figure 5B). ZFF29b, but not ZFF29a, is thus a potential transcriptional activator of erythroid genes.

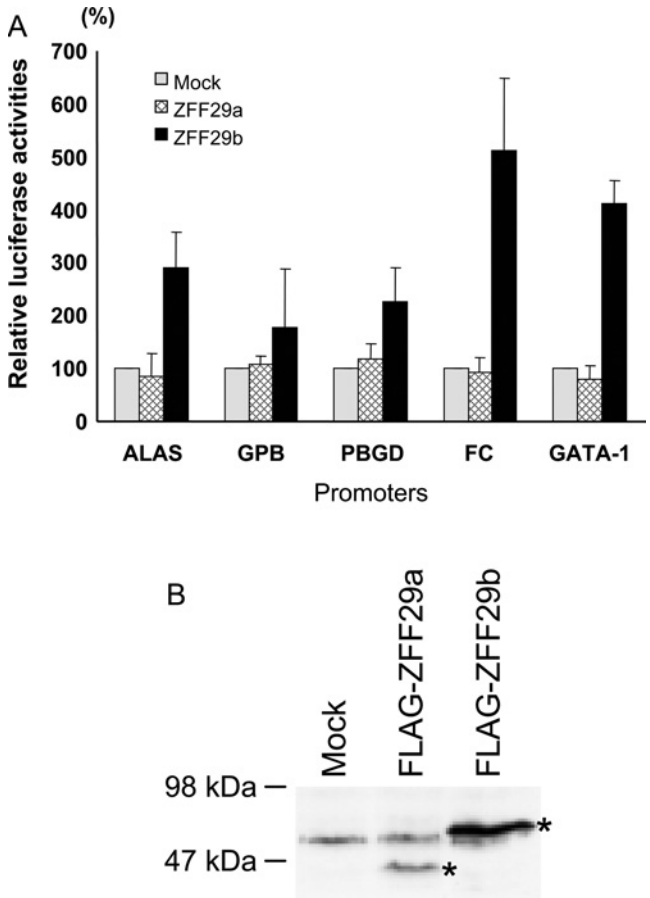

Figure 5 (A) *Trans*-activation by ZFF29 of promoters of various genes expressed in erythroid cells, and (B) expression of ZFF29 proteins in transiently transfected K562 cells

(**A**) Reporter constructs of 5-aminolaevulinate synthase (ALAS), glycophorin B (GPB), porphobilinogen deaminase (PBGD), ferrochelatase (FC) or GATA-1 gene promoters were transiently transfected into K562 cells with pSG5/ZFF29a or pSG5/ZFF29b. Luciferase activities were corrected by β-galactosidase activities, and expressed as a percentage of luciferase activity obtained by mock co-transfection (100%). Data are expressed as means + S.D. derived from multiple transfections. Note that ZFF29b apparently activates these promoters. (**B**) FLAG-tagged ZFF29 constructs were transfected using FuGENE™6 Transfection Reagent and the products were detected by Western analysis using an anti-FLAG antibody. Specific bands are indicated by asterisks.

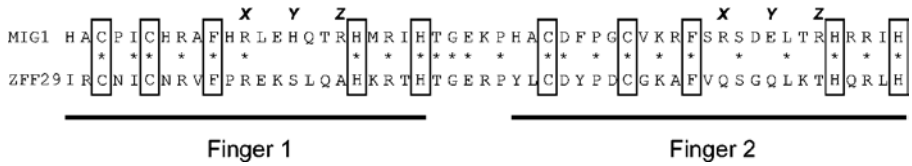

Figure 6 Comparison of zinc fingers between yeast MIG1 and ZFF29

Identical amino acid residues are indicated by asterisks. Structurally conserved residues, e.g. cysteines and histidines, are highlighted by rectangles. The letters ***X***, ***Y*** and ***Z*** indicate the positions directly involved in the recognition of nucleotides. Note that amino acid identity, excluding the structurally conserved residues and the inter-finger residues, is 27 % (10/37), and only one residue is identical among the six ***X***, ***Y*** and ***Z*** positions.

DISCUSSION

In our search for new EKLF-like or Sp1-like zinc-finger structures in early human fetal liver erythroblasts, we found a new zinc-finger motif; we cloned the full-length cDNA, and the deduced protein was designated ZFF29. The two tandem Krüppel-like zinc-finger motifs are the characteristic feature of ZFF29. The same motif is present in the yeast proteins MIG1, MIG2 and Yer028 [10]. These proteins are considered to be closely related, as their zinc fingers are very similar. In particular, key amino acid residues that are involved in nucleotide recognition are completely identical. However, ZFF29 is not related to the 'MIG1 family'. As illustrated in Figure 6, only one residue among the six key residues (indicated by X, Y and Z in Figure 6) is identical between ZFF29 and MIG1. Therefore the zinc-finger motif of ZFF29 exhibits a unique design. In humans no other factors, including hypothetical ones, have been identified with the same zinc-finger structure (see protein family ENSF00000034058; Ensembl human genome database: http://www.ensembl.org).

The expression pattern of ZFF29 reported in the present study confirms the non-ubiquitous expression pattern of the Fugu (*Fugu rubripes*) isologue reported by Gilligan et al. [9]: the ovary is the major tissue in which this gene is expressed. Besides the ovary, the bone marrow and especially the fetal erythroid tissue are other major sites of ZFF29 expression. On the basis of these expression patterns, ZFF29 thus appears to be a factor characterized by ovarian and fetal erythroid expression. Whether its action is limited to these tissues, or it is involved in the regulation of genes in other tissues, is not possible to surmise on the basis of the results of these expression studies. The following facts are relevant to this point. GATA-1 [11] and NF-E2 [12], although established erythroid specific factors, are also expressed and carry out substantial functions in non-erythroid tissues [13–15]. Of all the erythroid lineage-specific factors described so far, only EKLF appears to be characterized by true erythroid lineage specificity, in the sense that it has not been shown to affect, *in vivo*, the expression of any non-erythroid genes. By analogy with these erythroid factors, ZFF29 might play a role in the non-erythroid and non-ovarian tissues in which it is expressed. With regard to fetal stage, two other factors that we have cloned, KLF11 and KLF13, are capable of activating ε and γ globin gene promoters *in vitro* [3,4], but they do not show a unique expression pattern; although KLF11 is not detectable by Northern blotting in the adult bone marrow, the pancreas is the major tissue in which it is expressed in the adult [16]. KLF13 is expressed in the fetal liver, but also in other lineages [17,18]. Compared with these factors, ZFF29 shows relatively higher expression in haematopoietic tissues, implying that ZFF29 is involved in the regulation of erythroid lineage genes. In support of this notion, we have shown that ZFF29b, but not ZFF29a, is capable of activating erythroid gene promoters (Figure 5A). When considering the difference between ZFF29a and ZFF29b, the C-terminal acidic domain of ZFF29b (Figure 1B) is likely to be responsible for the function of ZFF29b as a transcriptional activator. Since the ZFF29 proteins carry a zinc-finger motif, the *trans*-activation of erythroid promoters by ZFF29b is possibly accomplished by a direct protein–DNA interaction on the promoter. To test this, we carried out a gel retardation experiment using the FLAG–ZFF29b protein expressed in COS-7 cells and a ^{32}P-labelled ferrochelatase gene promoter as a probe. No retarded bands, however, were formed by purified (i.e. immunoprecipitated) FLAG–ZFF29b (results not shown). Thus the mechanisms by which ZFF29b activates erythroid gene promoters, and the DNA sequence recognized by the ZFF29 proteins, remain to be elucidated.

We thank Ms Izumi Ojika and Sonoko Hatano for their excellent technical assistance. We are also grateful to Dr Tatsuya Ito and Dr Kazuyoshi Imaizumi for providing mouse and PCR primers for the mouse β-actin gene respectively. This work was supported by the Hibino Foundation and NIH grant HL20899 (subcontract 225053).

REFERENCES

1 Tupler, R., Perini, G. and Green, M. R. (2001) Expressing the human genome. Nature (London) **409**, 832–833
2 Wolfe, S. A., Nekludova, L. and Pabo, C. O. (1999) DNA recognition by Cys2His2 zinc finger proteins. Annu. Rev. Biophys. Biomol. Struct. **3**, 183–212
3 Asano, H., Li, X. S. and Stamatoyannopoulos, G. (1999) FKLF, a novel Krüppel-like factor that activates human embryonic and fetal β-like globin genes. Mol. Cell. Biol. **19**, 3571–3579
4 Asano, H., Li, X. S. and Stamatoyannopoulos, G. (2000) FKLF-2: a novel Krüppel-like transcriptional factor that activates globin and other erythroid lineage genes. Blood **95**, 3578–3584
5 Nehlin, J. O. and Ronne, H. (1990) Yeast MIG1 repressor is related to the mammalian early growth response and Wilms' tumour finger proteins. EMBO J. **9**, 2891–2898
6 Asano, H. and Stamatoyannopoulos, G. (1998) Activation of β-globin promoter by erythroid Krüppel-like factor. Mol. Cell. Biol. **18**, 102–109
7 Andrews, N. C. and Faller, D. V. (1991) A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. Nucleic Acids Res. **19**, 2499
8 Sambrook, J. and Russell, D. W. (2001) Molecular Cloning: A Laboratory Manual, 3rd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
9 Gilligan, P., Brenner, S. and Venkatesh, B. (2002) Fugu and human sequence comparison identifies novel human genes and conserved non-coding sequences. Gene **294**, 35–44
10 Lutfiyya, L. L., Iyer, V. R., DeRisi, J., DeVit, M. J., Brown, P. O. and Johnston, M. (1998) Characterization of three related glucose repressors and genes they regulate in Saccharomyces cerevisiae. Genetics **150**, 1377–1391
11 Evans, T. and Felsenfeld, G. (1989) The erythroid-specific transcription factor Eryf1: a new finger protein. Cell **58**, 877–885
12 Andrews, N. C., Erdjument-Bromage, H., Davidson, M. B., Tempst, P. and Orkin, S. H. (1993) Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. Nature (London) **362**, 722–728
13 Ito, E., Toki, T., Ishihara, H., Ohtani, H., Gu, L., Yokoyama, M., Engel, J. D. and Yamamoto, M. (1993) Erythroid transcription factor GATA-1 is abundantly transcribed in mouse testis. Nature (London) **362**, 466–468
14 Hirasawa, R., Shimizu, R., Takahashi, S., Osawa, M., Takayanagi, S., Kato, Y., Onodera, M., Minegishi, N., Yamamoto, M., Fukao, K. et al. (2002) Essential and instructive roles of GATA factors in eosinophil development. J. Exp. Med. **195**, 1473–1480

15 Peters, L. L., Andrews, N. C., Eicher, E. M., Davidson, M. B., Orkin, S. H. and Lux, S. E. (1993) Mouse microcytic anaemia caused by a defect in the gene encoding the globin enhancer-binding protein NF-E2. Nature (London) **362**, 768–770
16 Cook, T., Gebelein, B., Mesa, K., Mladek, A. and Urrutia, R. (1998) Molecular cloning and characterization of TIEG2 reveals a new subfamily of transforming growth factor–inducible Sp1-like zinc finger-encoding genes involved in the regulation of cell growth. J. Biol. Chem. **273**, 25929–25936
17 Song, A., Chen, Y. F., Thamatrakoln, K., Storm, T. A. and Krensky, A. M. (1999) RFLAT-1: a new zinc finger transcription factor that activates RANTES gene expression in T lymphocytes. Immunity **10**, 93–103
18 Kaczynski, J. A., Conley, A. A., Fernandez Zapico, M., Delgado, S. M., Zhang, J. S. and Urrutia, R. (2002) Functional analysis of basic transcription element (BTE)-binding protein (BTEB) 3 and BTEB4, a novel Sp1-like protein, reveals a subfamily of transcriptional repressors for the BTE site of the cytochrome P4501A1 gene promoter. Biochem. J. **366**, 873–882

Received 11 March 2004/20 August 2004; accepted 2 September 2004
Published as BJ Immediate Publication 2 September 2004, DOI 10.1042/BJ20040394