Available online at www.sciencedirect.com
SCIENCE DIRECT
Biochimica et Biophysica Acta 1727 (2005) 125–133
BIOCHIMICA ET BIOPHYSICA ACTA BBA
http://www.elsevier.com/locate/bba

Transcriptional regulation of FKLF-2 (KLF13) gene in erythroid cells

Ayako Mitsuma[a], Haruhiko Asano[a,*], Tomohiro Kinoshita[a], Takashi Murate[b], Hidehiko Saito[a,1], George Stamatoyannopoulos[c], Tomoki Naoe[a]

[a]*Department of Hematology, Nagoya University Graduate School of Medicine, Tsurumai-cho 65, Showa-ku, Nagoya, 466-8550, Japan*
[b]*Nagoya University School of Health Sciences, Daiko-minami, 1-1-20, Higashi-ku, Nagoya, 461-8673, Japan*
[c]*Division of Medical Genetics, University of Washington, Seattle, WA, 98195-7720, USA*

Received 5 May 2004; received in revised form 15 December 2004; accepted 20 December 2004
Available online 6 January 2005

Abstract

FKLF-2 (KLF13) was cloned from fetal globin-expressing tissues and has been shown to be abundantly expressed in erythroid cells. In this study we examined the transcriptional regulation of the *KLF13* gene. A 5.5 kb 5′ flanking region cloned from mouse erythroleukemia (MEL) cell genomic DNA showed that major *cis* regulatory activities exist in the 550 bp sequence to the unique transcription start site, and that the promoter is more active in K562 cells than in COS-7 cells. The promoter was *trans*-activated by co-expressed GATA-1 through the sequence containing two CCAAT motifs, suggesting that GATA-1 is involved in the abundant expression of KLF13 mRNA in the erythroid tissue. Dual action, i.e. activating effect in COS-7 and repressive effect in K562 cell, was observed on its own promoter, suggesting a feedback mechanism for the transcriptional control of the KLF13 gene in the erythroid environment. These findings provide an insight on the mechanism of inducible mRNA expression of the *KLF13* gene in erythroid cells.

Keywords: Erythroid differentiation; KLF family; Promoter; Transcriptional regulation

1. Introduction

Proteins containing three contiguous Cys_2–His_2 zinc fingers similar to Sp1 or EKLF constitute a KLF family [1–3]. By using the zinc finger motif, these proteins bind to G-rich (GC or GT/CACCC) motifs that are widely distributed in *cis*-regulatory DNA sequences for gene expression. Some KLFs are ubiquitously expressed, while others show a tissue-restricted expression pattern. It may be reasonable to speculate that the non-ubiquitous KLFs are involved in cell type specific gene expression. Confirming this notion, EKLF (KLF1), which is specifically expressed in erythroid cells [4], plays a critical role for the expression of the β-globin gene [5,6] by interacting with the proximal CACCC motif of the β gene promoter [7,8]. LKLF (KLF2), which is predominantly expressed in the lung [9], is essential for the normal lung development since *LKLF*$^{-/-}$ ES cells do not contribute to the lung formation in chimeric mice [10]. KLFs thus play a substantial role in their major expression tissues.

Fetal *Krüppel*-like factor-2 (FKLF-2: KLF13) was originally cloned from mouse yolk sac and human fetal liver erythroid cells, i.e. fetal globin-expressing tissues [11]. It is predominantly expressed in the bone marrow, striated muscles and a subset of T cells. Other groups, in fact, cloned the same gene from these tissues [12,13]. Of note is that the mRNA expression of KLF13 gene is up-regulated upon the induction of differentiation with chemicals in cell lines with erythroid features [11]. Similarly, the expression of KLF13 protein is up-regulated upon the maturation of T cells [12]. These facts strongly suggest that KLF13 may be involved in the differentiation of the cells in which it is expressed. In addition KLF13 is a powerful activator of a broad range of promoters of erythroid genes in vitro [11]. These results suggest that

* Corresponding author. Tel.: +81 52 744 2158; fax: +81 52 744 2141.
E-mail address: asanoh@med.nagoya-u.ac.jp (H. Asano).
[1] Present address: Nagoya National Hospital, Sannomaru 4-1-1, Naka-ku, Nagoya 460-0001, Japan.

0167-4781/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbaexp.2004.12.007

KLF13 plays a role in the development of erythroid phenotype. The transcriptional regulation for KLF13 gene expression may therefore be a part of molecular events of erythroid cell differentiation.

In this study we have cloned the 5′ flanking region of the mouse KLF13 gene. The promoter activity of the DNA fragment was tested in both erythroid (K562) and non-erythroid (COS-7) cells. Our data show that: the KLF13 gene promoter is more active in the erythroid environment than in the non-erythroid environment; and GATA-1 and KLF13 itself may have substantial effect in the transcriptional regulation of the KLF13 gene expression.

2. Materials and methods

2.1. Isolation of 5′ flanking DNA of the KLF13 gene

The 5′ flanking region of *KLF13* gene was obtained by inverse PCR (Fig. 1A). Four hundred nanograms of genomic

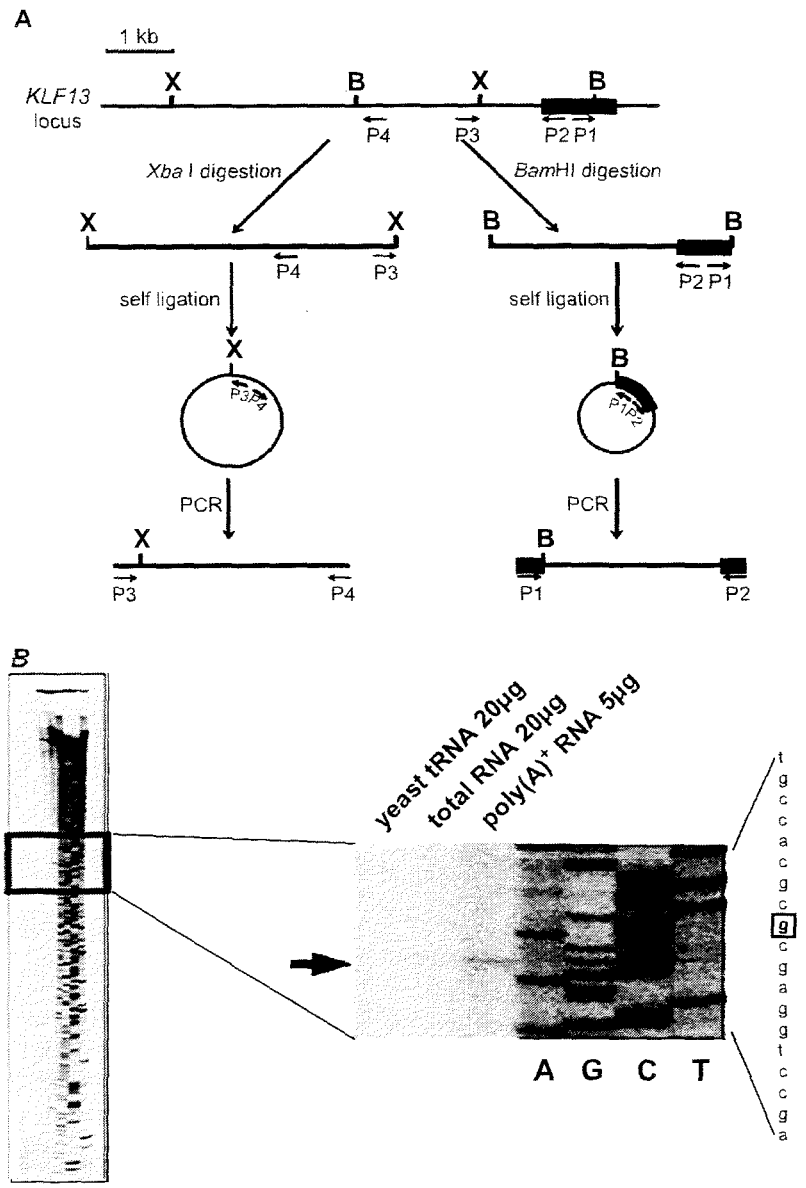

Fig. 1. Promoter of the *KLF13* gene. (A) Cloning of the 5′ flanking region of *KLF13* gene by inverse PCRs. The first exon is shown by a solid rectangle. B and X represent the *Bam*HI and *Xba*I sites, respectively. Primers used in inverse PCRs are indicated by arrows (P1–P4). Note that the positions of upstream *Bam*HI and *Xba*I sites are unknown prior to the experiments. (B) Primer extension analysis of KLF13 mRNA. RNAs used in the analysis are indicated above. Products of the sequencing reaction using the same probe were run on gel together, of which reading is shown at right. An arrow indicates the position of the extension product. The nucleotide in the sequence is highlighted by a square.

DNA extracted from MEL cells was digested with *Bam*HI or *Xba*I. Subsequently, the DNA was phenol-extracted, ethanol-precipitated and then self-ligated in 100 μL of reaction mixture at 16 °C for 30 min. The self-ligated DNA was dissolved in 8 μL low TE (5 mM Tris–HCl pH 8.0 and 0.1 mM EDTA) and 50 ng of the DNA was subjected to PCR by Expand Long PCR System (Roche) according to manufacture's instructions. The primer sets used in the PCR are 5′-GCCGCGACGGCAAGGACAGCC-3′ (P1) and 5′-GCATTGTGGCAGGCGGGACAGCC-3′ (P2), and 5′-GGTCTCCCTGACTTGATGGCTG-3′ (P3) and 5′-CAATGCTGAGGGAGTGAGCTAC-3′ (P4) for the *Bam*HI- and *Xba*I-cut DNA, respectively. The PCR products were cloned into T-vector (Promega) and were fully sequenced. The respective constructs were referred to pBam35 and pXba29.

2.2. Plasmid constructions

The 5.5 kb 5′ flanking DNA of the *KLF13* gene was reconstituted in pBluescript SK⁻ vector (Stratagene). The *Not*I–*Bam*HI fragment from pBam35 was subcloned into the same sites of the vector (referred to pBS28). Subsequently, the *Bam*HI–*Xba*I (blunted) fragment of pXba29 was inserted at the *Bam*HI and *Eco*RV sites of pBS28 (referred to pBS55). The 5.5 kb DNA was cut as a *Not*I (blunted)–*Kpn*I fragment and inserted into pGL2-Basic vector (Promega) at the *Hin*dIII (blunted) and *Kpn*I sites (referred to pGL55). The truncation of the DNA was performed by either (1) blunting and self-ligation after double digestion with *Kpn*I and a desired enzyme or (2) cutting with *Xba*I and a desired enzyme (blunted), and subsequent re-insertion into the pGL2-Basic vector cut with *Sma*I/*Xba*I.

GATA-1 expression vector was constructed using the coding region of GATA-1 amplified by reverse transcription (RT)-PCR using random-hexamer-primed cDNA from MEL cell total RNA. The PCR primers are 5′-CAGGTTCAACCCCAGTGTTCCCA-3′ and 5′-CCTTCAAGAACTGAGTGGGGCG-3′. The cDNA fragment was cloned into the T vector, and correct amplification was verified by sequencing. Subsequently, the open reading frame was cut out as a *Sph*I (blunted)–*Sac*I fragment and subcloned into a eukaryotic expression vector pSG5DD [14] at the *Eco*RI (blunted) and *Sac*I sites.

Mutations were introduced using a kit (Altered Sites II in vitro Mutagenesis Systems, Promega) following the manufacture's instructions. The nucleotide substitution of individual motifs is listed in Table 1.

2.3. RNA extractions

Total RNA was extracted by a standard method [15]. Poly(A)⁺ RNA was separated using an oligo (dT)-cellulose spun column (Pharmacia Biotech).

Table 1
List of nucleotide substitutions of *cis* elements

Motif	Sequence[a]	
	Wild type	Mutant
−0.37 kb CACCC	GGGTG	CTAGT
−0.23 kb CACCC	CACCC	ACTAG
Distal CCAAT	ATTGG	CAGAT
Proximal CCAAT	CCAAT	ATCTC

[a] Sense strand.

2.4. Primer extension and sequencing

Primer extension analysis was performed by a standard method [15]. An oligo DNA (5′-GCATTGTGGCAGGCGGGACAGCC-3′) was end-labeled with [γ^{33}P]ATP and 50,000 cpm was put into the extension reaction. Sequencing of the pBam35 by the same probe was performed using a kit (Cyclist, Stratagene).

2.5. Semi-quantitative RT-PCR

Semi-quantitative RT-PCR [11] was carried out using RQ1 DNase (Promega)-treated total RNA of COS-7 and K562 cells. The primers used for the amplification of KLF13 cDNA were 5′-TTCGCCTGCAGCTGGCAGGA-3′ and 5′-TGGCCGGGCTGATGGTGGG-3′. The condition of PCR was 95 °C (3 min) followed by variable number of cycles of 95 °C (30 s), 62 °C (30 s) and 72 °C (30 s).

2.6. Cell culture and transfections

K562 and COS-7 cells were cultured in RPMI1640 and DMEM, respectively, supplemented with 10% fetal calf serum. Transient transfections were performed using FuGENE 6 Reagent (Roche) according to the manufacture's instructions (reagent: DNA=3:1). Briefly, 1 μg plasmid DNA mixture (0.45 μg activator, 0.45 μg reporter and 0.1 μg pSVβ-Gal, or 0.9 μg reporter and 0.1 μg pSVβ-Gal) was transfected into 1×10^5 COS-7 and 5×10^5 K562 cells in 1 mL culture medium. COS-7 cells had been plated on a 12-well culture dish 24 h prior to the transfection. K562 cells were plated on the 12-well dish just before the transfection. After, the transfection cells were incubated at 37 °C for 24 h. Then, the cells were harvested, washed once with PBS and lysed in 200 μL Reporter Lysis Buffer (Promega). The cell lysate was subjected to luciferase and β-Gal assays as described elsewhere [8]. All transfections were performed three to five times in triplicate.

2.7. Data analysis and statistics

Data expressed as relative values to the average of the standard group were stored, analyzed and reported with the packages STATISTICA (StatSoft). Tukey's HSD (honestly significant difference) procedure (two-way ANOVA) was

used to evaluate the probability of significant differences. P values less than 0.05 were considered statistically significant.

3. Results

3.1. Promoter of the KLF13 gene

To examine the transcriptional regulation of *KLF13* gene, it was required to clone the 5′-flanking DNA. When we started this project the DNA sequence was still uncovered. To determine the unknown genomic DNA sequence we carried out two sequential inverse PCRs using MEL cell genomic DNA (Fig. 1A). Each reaction gave rise to a single band (data not shown). Subsequently, these DNA fragments were fully sequenced, which revealed 6183 nucleotides. (GenBank accession no. AY601638). Using the DNA we first determined the transcription start site of the *KLF13* gene by the primer extension analysis. As illustrated in Fig. 1B the poly(A)$^+$ RNA generated a unique band that is at the nucleotide position 5471. A band at the same position, although faint, is formed from the total RNA while no bands were observed in the yeast tRNA. These results indicate that: the transcription of the KLF13 gene starts at a unique site, and the DNA obtained by inverse PCRs contains the promoter region of the KLF13 gene. The presence of the promoter activity was confirmed by the luciferase assay using K562 cells. The DNA fragment clearly raised the luciferase activity (pGL55 in Table 2). The average luciferase count driven by the KLF13 genomic DNA was about 314 times higher compared with that obtained from the control vector (pGL2-Basic). The same experiment was performed using COS-7, i.e. non-erythroid, cells. The pGL55 construct generated approximately 164-fold more luciferase counts compared with the pGL2-Basic vector (Table 2). The fold increase, i.e. 314 in K562 cells vs. 164 in COS-7 cells, was significantly different (P<0.001). These results show that: there exists a promoter activity in the DNA fragment that we cloned, and the promoter is more active in erythroid cells than in non-erythroid cells.

To test whether the different promoter activities were reflected to the gene expression, the expression of KLF13 mRNA was analyzed by semi-quantitative RT-PCR. Since the nucleotide sequence of the KLF13 gene of COS-7 cells was unknown, primers were set in the sequence common between mouse and human KLF13 cDNAs. The primer set amplified the COS-7 and K562 genomic DNA at a comparable efficiency (Fig. 2A). Subsequently, cDNAs of the two cell lines that were diluted to generate similar band patterns on amplification of 28S ribosomal RNA (rRNA) were subjected to PCR of the KLF13 gene. As shown in Fig. 2B cDNA of K562 cells yielded more intense bands in all cycles tested than that of COS-7 cells, demonstrating that higher expression of KLF13 mRNA in K562 cells than in COS-7 cells.

Table 2
Promoter activity of the 5.5 kb DNA fragment

	COS-7	K562
pGL2-Basic	1±0.17	1±0.23
pGL55	163.5±27.7	313.8±54.2

Luciferase counts were corrected by the β-Gal activities. The average luciferase activities obtained from pGL2-Basic vector was considered as 1. Data are expressed as mean±1 S.D. of three independent transfections in triplicate.

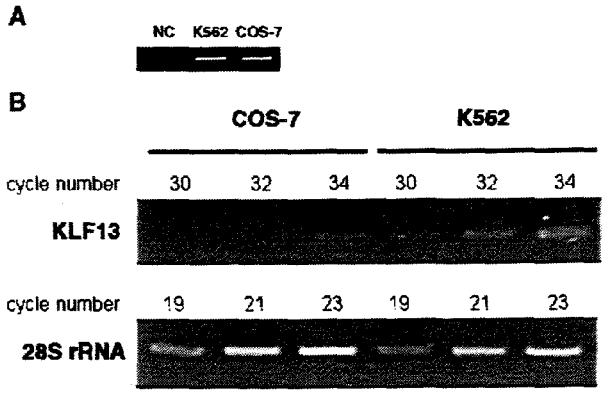

Fig. 2. Expression of KLF13 mRNAs in COS-7 and K562 cells. (A) PCR using the same amount of genomic DNA. One hundred nanograms of genomic DNA was amplified by 33 cycles of PCR in a 50 μL reaction volume. Subsequently, 10 μL was run on gel. Note that similar bands were generated in COS-7 and K562 cells, indicating comparable efficiencies of amplification between the two cell lines. NC means negative control in which TE buffer instead of DNA was put in the PCR reaction. (B) Semi-quantitative RT-PCR using COS-7 and K562 cell cDNAs. The numbers shown above indicate PCR cycles. Note the similar amplification of 28S rRNA gene between the two cells, and obviously more intense bands on the amplification of KLF13 gene in K562 cells than in COS-7 cells.

3.2. Erythroid cell specific regulation of the KLF13 gene promoter

To elucidate the *cis* regulatory elements of the KLF13 gene promoter, a series of truncated DNA fragments were prepared (Fig. 3A) and their promoter activities were analyzed in COS-7 and K562 cells. Results are expressed as promoter activities relative to those obtained from the 0.13 kb DNA fragment (considered as 1). As shown in Fig. 3B, in COS-7 cells promoter activities were significantly different between 5.5 and 2.9 kb (P<0.001), 2.0 and 0.73 kb (P<0.001), 0.55 and 0.37 kb (P<0.001), 0.37 and 0.29 kb (P<0.001), 0.29 and 0.22 kb (P<0.001), and 0.22 and 0.13 kb (P<0.001), suggesting that two DNA regions, i.e. −0.13 to −0.37 kb and −0.73 to −2.0 kb, function as positive regulatory elements while −0.37 to −0.55 kb and −2.9 to −5.5 kb DNA regions function as negative regulatory elements in non-erythroid cells (Fig. 3C). Since the 0.37 kb fragment exhibited the most powerful promoter activity, which is 12.4-fold higher relative to that of 0.13 kb fragment, among the 9 truncated promoters, major

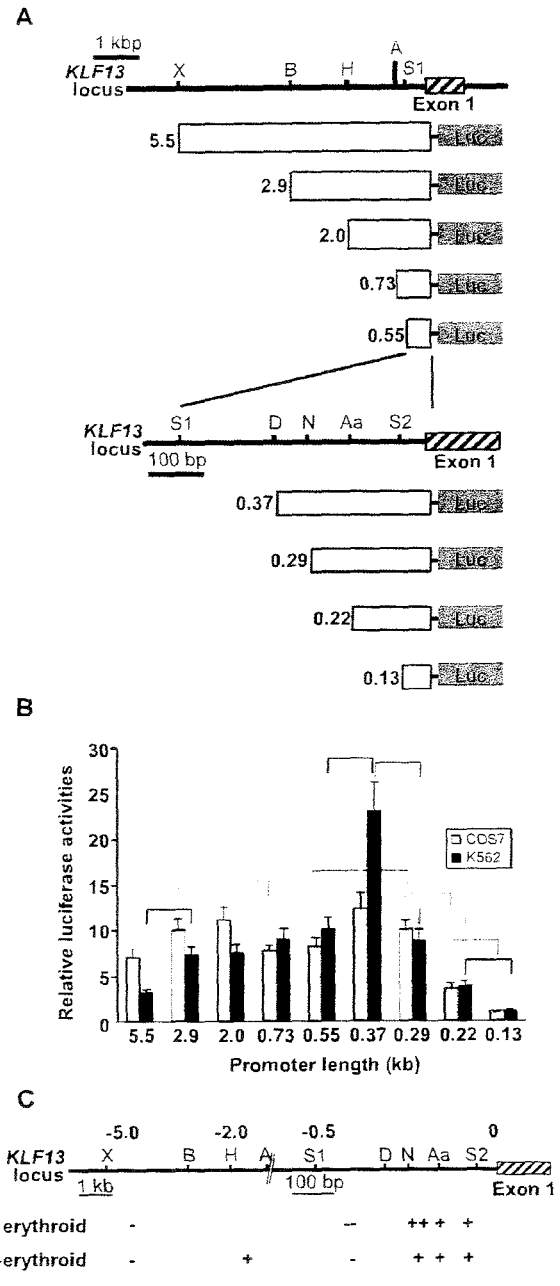

Fig. 3. Activities of KLF13 gene promoter in erythroid and non-erythroid cells. (A) Reporter constructs and *KLF13* locus. Luciferase gene is driven by truncated promoters indicated by light gray rectangles. Restriction sites used for the truncation are indicated above. Abbreviations for restriction sites are X, *Xba*I; B, *Bam*HI; H, *Hin*dIII; A, *Apa*I; S1, *Sac*I; D, *Dra*III; N, *Nae*I; Aa, *Aat*II; and S2, *Sac*II. The length of the promoters is shown at left in kb. (B) Relative luciferase activities obtained from COS-7 (gray bars) and K562 (black bars) cells are shown. Luciferase counts were corrected by the β-Gal activity, and the average count obtained from the 0.13 kb promoter was considered as 1. Error bars indicate one S.D. Results were obtained from three independent transfections in triplicate. Promoter pairs showing significantly different activities are indicated by broken (COS-7) and solid (K562) lines. (C) Positive (+) and negative (−) regulatory regions of KLF13 gene promoter. *KLF13* locus is shown above and numbers are the distance to the transcription start site in kb. Note that the scale is altered at −0.7 kb position. Regions having a potent effect are indicated by ++ and −−.

promoter activities exist in this DNA sequence. The highest promoter activity of the 0.37 kb fragment was also observed in K562 cells (Fig. 3B). However, it is of note that the activity is 22.9-fold higher than that of the 0.13 kb fragment, which should be contrasted to the 12.4-fold increase observed in COS-7 cells. The relative promoter activities of DNA fragments containing the upstream sequence (5.5, 2.9 and 2.0 kb) were not as high as those of COS-7 cells (Fig. 3B). The relative activities of 0.73 and 2.0 kb promoters were 9.0 and 7.5, respectively, which was not statistically significant (P=0.46). Therefore the positive regulatory effect of the −0.73 to −2.0 kb region observed in COS-7 cells was absent in K562 cells (Fig. 3C). Together, it is suggested that: the 80 bp sequence between −0.29 and −0.37 kb is more powerful in an erythroid environment than in a non-erythroid environment (expressed as ++ in Fig. 3C), and among the negative regulatory elements the −0.37 to −0.55 kb sequence has a strong effect (expressed as - in Fig. 3C) since the powerful activity of the −0.29 to −0.37 kb region disappeared in the presence of the 180 bp, i.e. −0.37 to −0.55 kb, DNA.

3.3. The erythroid factor GATA-1 and KLF13 are potential transcriptional regulators of the KLF13 gene promoter

The results shown above indicate that: the regulation of KLF13 gene promoter is distinct in erythroid cells, and major regulatory elements of the promoter are located in the proximal 550 bp region. Inspection and computer database search (TRANSFAC, http://www.motif.genome.jp/) of the DNA sequence revealed a number of transcription factor binding sites. Of remark is the presence of multiple GC, CACCC and GATA-1-binding motifs (Fig. 4), suggesting that Sp1/KLF and GATA factors participate in the regulation of KLF13 gene promoter. To test how the transcription of *KLF13* gene is regulated in erythroid cells we analyzed the effects of GATA-1 and KLF13, both of which are expressed in erythroid cells, on the KLF13 gene promoter. The expression vectors of these factors and luciferase reporter constructs were co-transfected into COS-7 and K562 cells. To focus the *cis* sequence corresponding to the effects, if any, of these factors we utilized three promoters with different length, i.e. 0.55, 0.29 and 0.13 kb fragments. As shown in Fig. 5A GATA-1 activated the 0.55 and 0.29, but not 0.13, kb promoters in COS-7 cells. The respective fold increases compared with

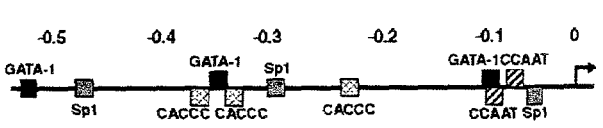

Fig. 4. *cis* elements of the 0.55 kb promoter that are potentially recognized by Sp1 KLFs and GATA-1. Numbers shown above are the distance to the transcription start site in kb. Motifs are indicated by black (GATA-1), gray (Sp1), hatched (CACCC) and striped (CCAAT) squares. The transcription start site is indicated by an arrow.

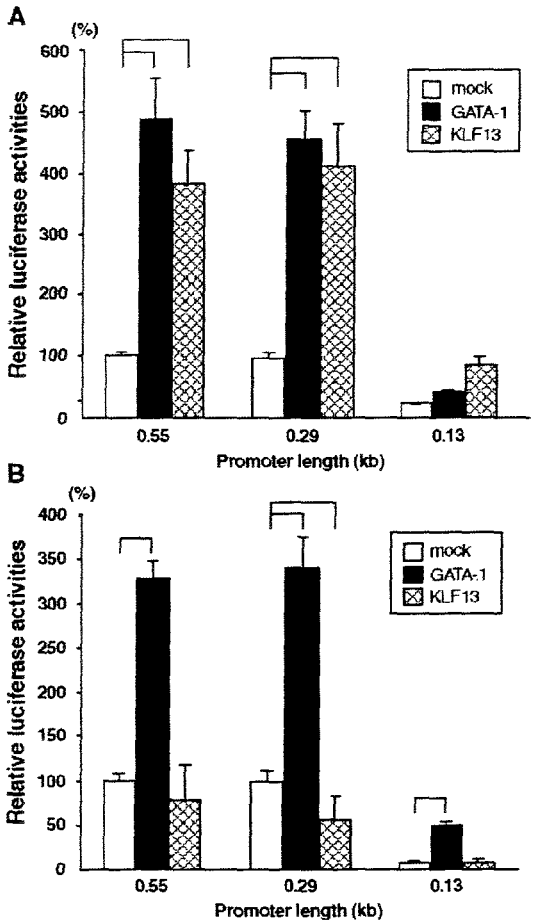

Fig. 5. (A and B) Effects of GATA-1 and KLF13 itself on the KLF13 gene promoter. Relative luciferase activities derived from 0.55, 0.29 and 0.13 kb promoters are shown. Reporter constructs were co-transfected with either mock, GATA-1 or KLF13 into COS-7 (A) and K562 (B) cells. Luciferase counts were corrected by the β-Gal activity, and the average count of the 0.55 kb promoter obtained from mock-transfected cells was considered as 100%. Error bars indicate one S.D. Solid lines shown above indicate that the promoter activities were significantly different. Results were obtained from three (COS-7) and four (K562) independent transfections in triplicate.

its absence were 4.9 ($P<0.001$), 4.7 ($P<0.01$) and 1.9 ($P=0.99$). The activation of the KLF13 gene promoter by GATA-1 was also observed in K562 cells (Fig. 5B). In the presence of GATA-1 the respective promoter activities were 3.3, 3.5 and 6.5 times higher compared with those without GATA-1 ($P<0.001$ in all promoters). These results show that GATA-1 is a potential *trans*-activator of the KLF13 gene promoter, and the proximal 130 bp DNA sequence is capable of mediating the GATA-1 effects specifically in the erythroid environment. In addition it is noteworthy that the 0.55 and 0.29 kb promoters showed essentially the same activities, which is similar than the result shown in Fig. 3B, in any experimental condition, suggesting that neither GATA-1 nor KLF13 is responsible for the erythroid specificity in the sequence between −0.55 and −0.29 kb (Figs. 3B and C).

KLF13 activates its own promoter in COS-7 cells (Fig. 5A). The activities of the 0.55, 0.29 and 0.13 kb promoters in the presence of KLF13 were 3.8 ($P<0.01$), 4.2 ($P<0.01$) and 4.0 ($P=0.05$) times higher compared with its absence. In K562 cells, however, this was not the case i.e. the respective promoter activities relative to those mock-transfected were 0.78 ($P=0.27$), 0.56 ($P<0.001$) and 1.04 ($P=1.0$) (Fig. 5B). KLF13 thus failed to activate its own promoter in K562 cells. Rather, KLF13 tended to repress the 0.55 and 0.29 kb promoters.

3.4. GATA-1 activates KLF13 gene promoter through the sequence containing CCAAT motifs

The data shown above indicate that erythroid specific GATA-1-responsive element(s) of the KLF13 promoter are present in the very proximal 0.13 kb sequence. There is a non-canonical GATA-1-binding site within the sequence (Fig. 4). In addition two CCAAT boxes could be a potential GATA-1-binding site [16,17]. We, however, failed to demonstrate GATA-1-binding to these sequences by gel shift experiments, in which there was a distinct GATA-1-binding to the canonical GATA sequence (data not shown). Therefore, we focused on the role of activation by GATA-1 of the major *cis* elements in the 0.13 kb sequence, i.e. two CCAAT boxes. Mutations were introduced into individual CCAAT motifs of the 0.55 kb KLF13 gene promoter. We analyzed how these mutations affect the activity of GATA-1 on the promoter in K562 cells by transient transfection experiments. Results are shown in percentage luciferase counts relative to that obtained from mock-transfected wild type (WT) promoter (considered as 100%). As illustrated in Fig. 6 mutation of the single CCAAT motifs failed to affect the basal promoter activity: 83% in $CCAAT^{mut1}$ ($P=0.73$), and 105% in $CCAAT^{mut2}$ ($P=1.0$). In contrast to these single mutations, double mutation ($CCAAT^{mut3}$) brought about remarkable decline of the promoter activity (15%, $P<0.001$). GATA-1 activated these CCAAT-mutated promoters, however the relative luciferase activities were significantly lower in $CCAAT^{mut1}$ (185%, $P<0.001$), $CCAAT^{mut2}$ (303%, $P<0.001$) and $CCAAT^{mut3}$ (56%, $P<0.001$) compared with that of WT (406%). Thus GATA-1 less effectively activated CCAAT-mutated KLF13 gene promoters.

3.5. KLF13 represses its own promoter in K562 cells

As shown in Fig. 5B KLF13 has a tendency to repress its own promoter in K562 cells. Since this was observed in 0.55 and 0.29 kb promoters, and not in 0.13 kb promoter, we focused on two CACCC motifs located at the −0.37 and −0.23 kb positions, potential target sequences of the KLF13 protein. Mutations were introduced into individual motifs of the 0.55 kb promoter ($CACCC^{mut1}$, $CACCC^{mut2}$ and $CACCC^{mut3}$ in Fig. 7). We tested whether the mutations affect the suppressive activity of KLF13 on its

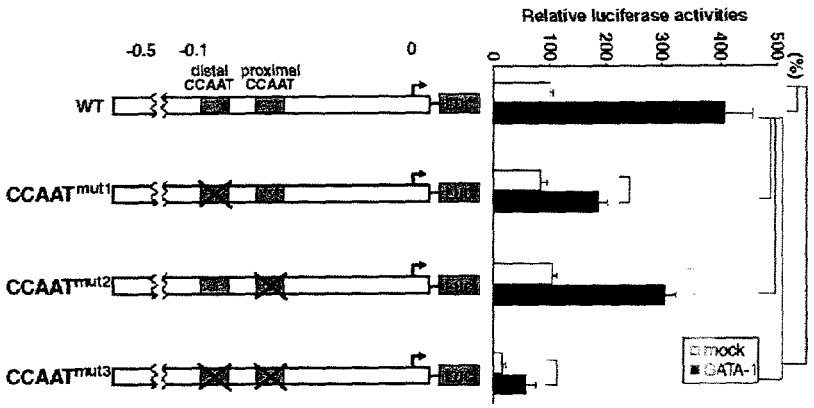

Fig. 6. Role of CCAAT motifs for the activity of GATA-1 on the KLF13 gene promoter. They were disrupted by point mutations listed in Table 1. Promoters lacking the distal, proximal and both CCAAT motifs are referred to $CCAAT^{mut1}$, $CCAAT^{mut2}$ and $CCAAT^{mut3}$, respectively, which are shown in the left panel. The numbers shown above are the distance to the transcription start site in kb. The right panel shows the relative luciferase activities tested in K562 cells. Luciferase counts were corrected by the β-Gal activity, and the average count of wild type (WT) promoter without co-expressed GATA-1 was considered as 100%. Error bars indicate one S.D. Solid lines shown at the right indicate that the promoter activities were significantly different. Results were obtained from three independent transfections in triplicate.

own promoter in K562 cells by the transient transfection assay. Results are shown in percentage luciferase counts of the basal WT promoter (Fig. 7). Mutations did not reduce the promoter strength at all: 119% for $CACCC^{mut1}$ (P=0.61), 111% for $CACCC^{mut2}$ (P=0.96) and 102% for $CACCC^{mut3}$ (P=1.0). In the presence of KLF13 the average luciferase counts were low: 78% (WT), 85% ($CACCC^{mut1}$), 82% ($CACCC^{mut2}$) and 57% ($CACCC^{mut3}$). The respective fold increases from the basal promoter activity were 0.78 (P=0.46), 0.71 (P=0.03), 0.74 (P=0.11) and 0.56 (P=0.001). It was thus consistently observed that KLF13, to some extent, represses its own promoter. The repressive effect was strengthened by the mutations of the two CACCC boxes.

4. Discussion

KLF13 gene is highly expressed in erythroid cells, and its mRNA expression is up-regulated upon the induction of differentiation in erythroid lines [11], which raises the possibility that KLF13 participates in molecular events of erythroid cell differentiation. In view of this it is of interest to investigate how the transcription of the KLF13 gene is regulated in erythroid cells. In this study we cloned and characterized the 5′ flanking regulatory region of the KLF13 gene.

We identified a unique transcription start point of KLF13 gene. To reveal how the transcription is regulated in erythroid cells we analyzed the promoter activity of the

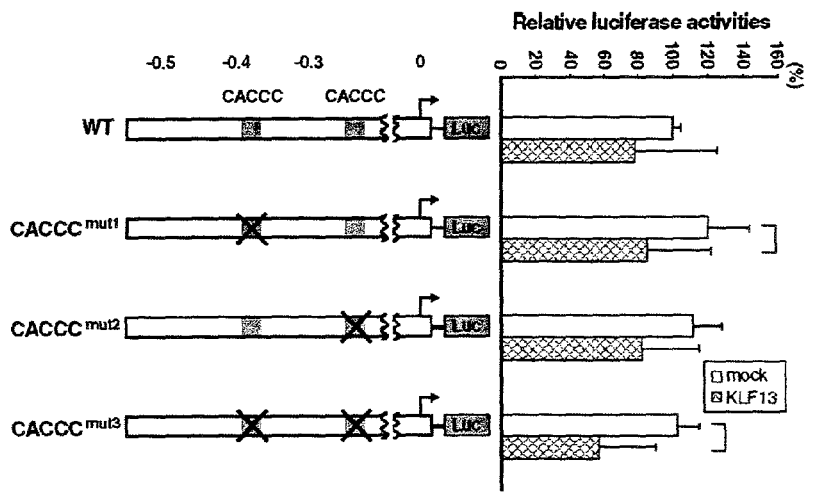

Fig. 7. Role of CACCC motifs for the activity of KLF13 on its own promoter. They were disrupted by the point mutations listed in Table 1. Promoters lacking the −0.37 kb, −0.23 kb and both CACCC motifs are referred to $CACCC^{mut1}$, $CACCCT^{mut2}$ and $CACCC^{mut3}$, respectively, which are shown in the left panel. The numbers shown above are the distance to the transcription start site in kb. Right panel shows relative luciferase activities tested in K562 cells. Luciferase counts were corrected by the β-Gal activity, and the average count of wild type (WT) promoter without co-expressed KLF13 was considered as 100%. Error bars indicate one S.D. The solid lines shown at the right indicate that the promoter activities were significantly different. Results were obtained from five independent transfections in triplicate.

5.5 kb flanking DNA sequence in K562 and COS-7 cells. Our data show that the KLF13 gene promoter in the erythroid environment is more powerful than in the non-erythroid environment, which may be related to the higher KLF13 gene expression in K562 cells than in COS-7 cells (Fig. 2B). In addition positive and negative regulatory regions were at least in part different between the two environments. Of particular interest is the DNA sequence with a potent positive activity between −0.29 and −0.37 kb, which was distinctive in erythroid cells. However, the potent promoter activity disappeared when the −0.55 kb DNA was used, which in turn suggests the presence of a potent negative regulatory activity in the DNA sequence between −0.37 and −0.55 kb. The identification of *cis* elements responsible for the positive and negative regulation of these DNA sequences may be the next step to uncover the mechanism of erythroid cell specific transcriptional control of the KLF13 gene.

It is interesting that the KLF13 gene promoter was activated by KLF13 itself in COS-7 cells, whereas this was not the case in K562 cells. Rather, the promoter tended to be repressed in the presence of KLF13, suggesting that there exists a feedback mechanism in the control of KLF13 gene transcription in erythroid cells. *cis* Element(s) responsible for the repressive action of KLF13 remain to be elucidated. Regarding this issue we explored the role of two CACCC boxes of the KLF13 gene promoter. Mutations of the two CACCC motifs, however, made the repressive effect of KLF13 more evident. This might be relevant to the fact that these CACCC motifs were likely involved in the *trans*-activation of the promoter by KLF13 itself in COS-7 cells since KLF13 failed to sufficiently activate the promoter (data not shown). Therefore, it is suggested that the CACCC boxes may be key *cis* elements corresponding to the *trans*-activating effect of KLF13 on its own promoter.

We have shown that the KLF13 promoter is *trans*-activated by GATA-1, indicating that KLF13 is potentially a downstream gene of GATA-1. The mechanism by which GATA-1 activates the KLF13 gene promoter needs to be determined. The fact that the promoter with distal CCAAT box mutation was less effectively activated by GATA-1 ($CCAAT^{mut1}$ in Fig. 6) should be taken into account, which suggests that the distal CCAAT box or the overlapping non-canonical GATA-1-binding site may play a role in the KLF13 promoter activation by GATA-1. This may be supported by the observation that the reduction of the activation of the $CCAAT^{mut1}$ by GATA-1 was significant, $P<0.01$, when the same experiment was performed using COS-7 (data not shown). The overlapping CCAAT box and GATA-1-binding site of the KLF13 promoter is structurally similar to the Aγ globin gene promoter [16]. In contrast to the Aγ promoter where GATA-1 binding could be demonstrated by gel shift assays, we have failed to demonstrate GATA-1 binding to the KLF13 promoter containing CCAAT motifs. How does GATA-1 activate the KLF13 promoter in such a situation? There are two possibilities: either GATA-1 binds to the promoter region but the binding is too weak to be detected; or the activation of the promoter by GATA-1 is indirect, a case such as transcription factor(s) regulated by GATA-1 activate the KLF13 gene promoter through the CCAAT or neighboring sequences. The former possibility may be supported by the evidence that the binding to CCAAT element is very unstable [17]. In addition, the non-canonical GATA-1 binding sequence, GATT that overlaps with the distal CCAAT motif of the KLF13 gene promoter (Fig. 4), is supposedly recognized by the GATA-1 carboxyl finger [18]. Therefore, the binding of GATA-1 to the promoter region in vivo may not necessarily be ruled out, even though we failed to demonstrate it in vitro. It should also be mentioned that a number of factors including CP1/NF-Y and CP2 which are expressed in a wide variety of tissues interact with the CCAAT motif [19]. It is therefore possible that GATA-1 activates KLF13 through enhancements of expression of such ubiquitously expressed factors. GATA-1 is also known to influence the expression of other erythroid transcription factors, e.g. EKLF [20], KLFD [21] and GATA-1 itself [22,23]. It has been shown that NF-E4 is a co-factor of CP2 [24], a transcription factor abundantly expressed in erythroid cells and involved in γ globin gene regulation. It is conceivable that GATA-1 activates NF-E4 which may bind to the KLF13 CCAAT box through CP2 and enhance KLF13 expression. It remains to be, however, tested whether the expression of NF-E4 is under the control of GATA-1. The mechanism of action of GATA-1 on the KLF13 promoter remains therefore unknown.

Acknowledgments

This work was supported by NIH grant HL20899 (subcontract 225053).

References

[1] S. Philipsen, G. Suske, A tale of three fingers: the family of mammalian Sp/XKLF transcription factors, Nucleic Acids Res. 27 (1999) 2991–3000.
[2] J.J. Bieker, Krüppel-like factors: three fingers in many pies, J. Biol. Chem. 276 (2001) 34355–34358.
[3] J. Kaczynski, T. Cook, R. Urrutia, Sp1- and Krüppel-like transcription factors, Genome Biol. 4 (2003) 206.
[4] I.J. Miller, J.J. Bieker, A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Krüppel family of nuclear proteins, Mol. Cell. Biol. 13 (1993) 2776–2786.
[5] B. Nuez, D. Michalovich, A. Bygrave, R. Ploemacher, F. Grosveld, Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene, Nature 375 (1995) 316–318.
[6] A.C. Perkins, A.H. Sharpe, S.H. Orkin, Lethal β-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF, Nature 375 (1995) 318–322.
[7] D. Donze, T.M. Townes, J.J. Bieker, Role of erythroid Kruppel-like factor in human γ- to β-globin gene switching, J. Biol. Chem. 270 (1995) 1955–1959.

[8] H. Asano, G. Stamatoyannopoulos, Activation of β-globin promoter by erythroid Krüppel-like factor, Mol. Cell. Biol. 18 (1998) 102–109.
[9] K.P. Anderson, C.B. Kern, S.C. Crable, J.B. Lingrel, Isolation of a gene encoding a functional zinc finger protein homologous to erythroid Krüppel-like factor: identification of a new multigene family, Mol. Cell. Biol. 15 (1995) 5957–5965.
[10] M.A. Wani, S.E. Wert, J.B. Lingrel, Lung Krüppel-like factor, a zinc finger transcription factor, is essential for normal lung development, J. Biol. Chem. 274 (1999) 21180–21185.
[11] H. Asano, X.S. Li, G. Stamatoyannopoulos, FKLF-2: a novel Krüppel-like transcriptional factor that activates globin and other erythroid lineage genes, Blood 95 (2000) 3578–3584.
[12] A. Song, Y.F. Chen, K. Thamatrakoln, T.A. Storm, A.M. Krensky, Transcriptional regulation of RANTES expression in T lymphocytes, Immunity 10 (1999) 93–103.
[13] K.M. Martin, W.N. Cooper, J.C. Metcalfe, P.R. Kemp, Mouse BTEB3, a new member of the basic transcription element binding protein (BTEB) family, activates expression from GC-rich minimal promoter regions, Biochem. J. 345 (2000) 529–533.
[14] H. Asano, X.S. Li, G. Stamatoyannopoulos, FKLF, a novel Krüppel-like factor that activates human embryonic and fetal β-like globin genes, Mol. Cell. Biol. 19 (1999) 3571–3579.
[15] J. Sambrook, D.W. Russell, Molecular Cloning: A Laboratory Manual, 3rd ed., Cold Spring Harbor Laboratory Press, New York, 2001.
[16] M. Berry, F. Grosveld, N. Dillon, A single point mutation is the cause of the Greek form of hereditary persistence of fetal haemoglobin, Nature 358 (1992) 499–502.
[17] Q. Li, Z.J. Duan, G. Stamatoyannopoulos, Analysis of the mechanism of action of non-deletion hereditary persistence of fetal hemoglobin mutants in transgenic mice, EMBO J. 20 (2001) 157–164.
[18] D.J. Whyatt, E. deBoer, F. Grosveld, The two zinc finger-like domains of GATA-1 have different DNA binding specificities, EMBO J. 12 (1993) 4993–5005.
[19] R. Mantovani, The molecular biology of the CCAAT-binding factor NF-Y, Gene 239 (1999) 15–27.
[20] M. Crossley, A.P. Tsang, J.J. Bieker, S.H. Orkin, Regulation of the erythroid Krüppel-like factor (EKLF) gene promoter by the erythroid transcription factor GATA-1, J. Biol. Chem. 269 (1994) 15440–15444.
[21] A.C. Oates, S.J. Pratt, B. Vail, Y. Yan, R.K. Ho, S.L. Johnson, J.H. Postlethwait, L.I. Zon, The zebrafish klf gene family, Blood 98 (2001) 1792–1801.
[22] S.F. Tsai, E. Strauss, S.H. Orkin, Functional analysis and in vivo footprinting implicate the erythroid transcription factor GATA-1 as a positive regulator of its own promoter, Genes Dev. 5 (1991) 919–931.
[23] K. Nishikawa, M. Kobayashi, A. Masumi, S.E. Lyons, B.M. Weinstein, P.P. Liu, M. Yamamoto, Self-association of gata1 enhances transcriptional activity in vivo in zebra fish embryos, Mol. Cell. Biol. 23 (2003) 8295–8305.
[24] W. Zhou, D.R. Clouston, X. Wang, L. Cerruti, J.M. Cunningham, S.M. Jane, Induction of human fetal globin gene expression by a novel erythroid factor, NF-E4, Mol. Cell. Biol. 20 (2000) 7662–7672.

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*HI, *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

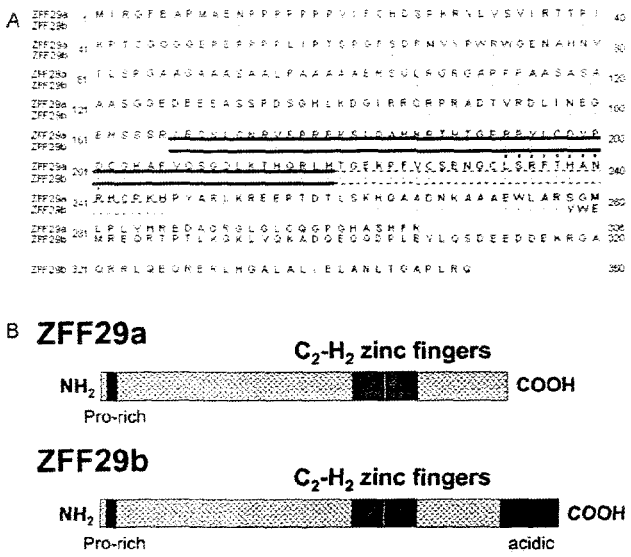

Figure 1 (A) Deduced polypeptide sequences of ZFF29a and ZFF29b, and (B) schematic representation of the deduced structures of ZFF29a and ZFF29b proteins

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.

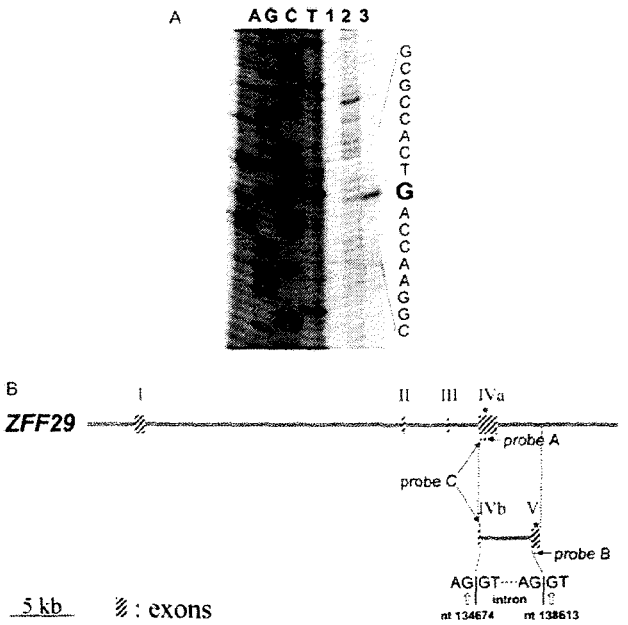

Figure 2 Transcription of the *ZFF29* gene

(**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

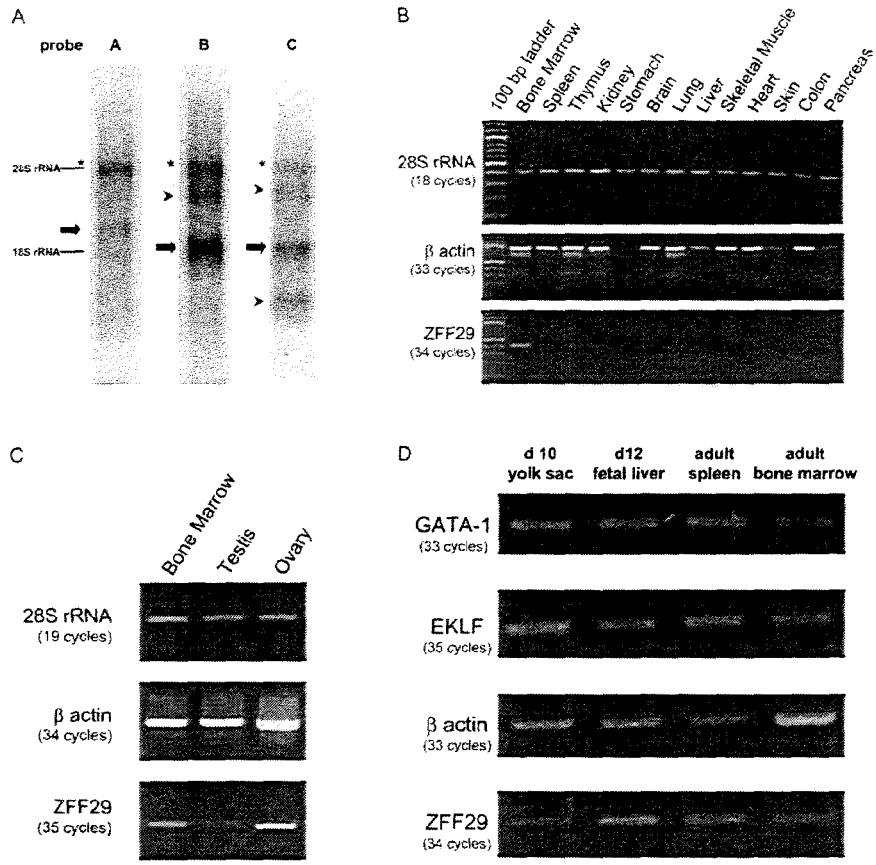

Figure 3 Expression of ZFF29 mRNA

(**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™ 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

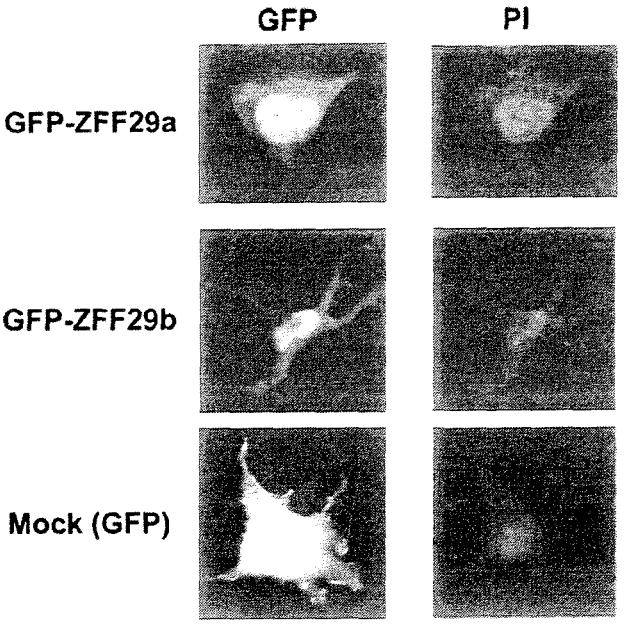

Figure 4 Nuclear localization of ZFF29 proteins

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.

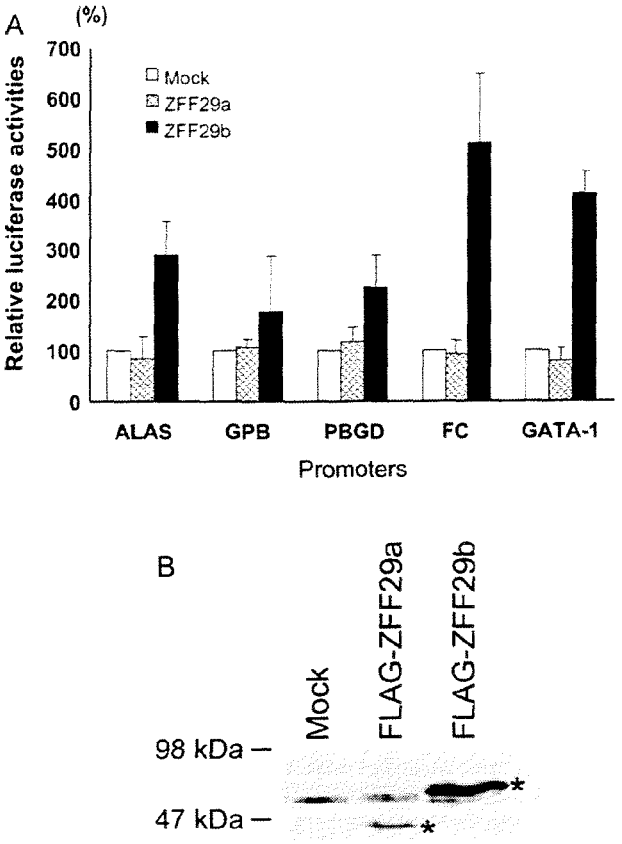

Figure 5 (A) *Trans*-activation by ZFF29 of promoters of various genes expressed in erythroid cells, and (B) expression of ZFF29 proteins in transiently transfected K562 cells

(**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.

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.

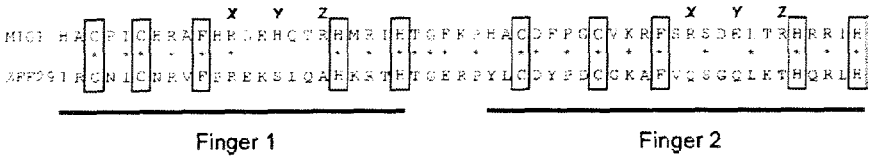

Figure 6 Comparison of zinc fingers between yeast MIG1 and ZFF29

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

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 320 (2004) 920–926
BBRC
www.elsevier.com/locate/ybbrc

Expression cloning of oligomerization-activated genes with cell-proliferating potency by pseudotype retrovirus vector

Akihiro Abe,[a,*] Nobuhiko Emi,[a] Tadaharu Kanie,[a] Shizuka Imagama,[a] Yoshie Kuno,[b] Masahide Takahashi,[c] Hidehiko Saito,[d,e] and Tomoki Naoe[a]

[a] *Department of Hematology, Nagoya University Graduate School of Medicine, Nagoya, Japan*
[b] *Tajimi Prefectural Hospital, Gifu, Japan*
[c] *Molecular Pathology, Nagoya University Graduate School of Medicine, Nagoya, Japan*
[d] *Nagoya Medical Center, Nagoya, Japan*
[e] *Aichi Blood Disease Research Foundation, Nagoya, Japan*

Received 6 May 2004
Available online 25 June 2004

Abstract

We developed a method of clone proliferation promoting fusion genes whose proteins were activated by protein oligomerization through the helix–loop–helix region (PNT domain) of TEL. We inserted a cDNA library downstream of the PNT domain with a retrovirus vector. The resulting retrovirus infected cytokine-dependent 32D cells and cells with cytokine-independent growth were analyzed for the inserted cDNA. We cloned 25 independent fusion genes including seven kinds of partner genes. Six of the seven were a fusion of TEL with protein tyrosine kinase, LYN, HCK, FGR, SYK, FLT3, and TYK2. A serine/threonine kinase, ARAF1, was also found to fuse with TEL. These kinase fusion proteins included kinase domains with proper reading frames. These fusions may be a useful model for clarifying the downstream signal transduction of constitutive active kinase and this expression cloning method may provide a new tool with which to study cell proliferation signalling.

Keywords: Expression cloning; Retrovirus vector; TEL; Tyrosine kinase; Serine/threonine kinase

Expression cloning of cellular oncogenes or proto-oncogenes has been developed by detecting their ability to transform and deregulate the growth of cells [1–4]. Cell growth and the morphological transformation of NIH3T3 fibroblasts have largely been used to clarify single-hit oncogenes. The development of an efficient gene-transfer technique with a retrovirus vector makes it possible to clone relatively rare cDNA by expression cloning [5–8] and to introduce cDNA to a variety of cells. The strategy of converging hematological cell lines to cytokine-independent growth by expression cloning is applicable for identifying oncogenes, however, the applications of hematological cell lines are limited. One limitation of using cytokine-dependent hematopoietic cell lines to clone oncogenes is that the transfecting cDNAs do not produce active oncogene products without specific structural changes, such as mutation, duplication or truncation. To overcome these limitations, we designed an oligomerization-activated gene library using a retrovirus vector and infected a cytokine-dependent cell line to obtain cytokine-independent clones.

Retrovirus functional cloning is a powerful and efficient way to obtain specific cDNA based on the function of the encoding protein [8–10]. The advantages of using the retrovirus vector are the stable expression of the inserted gene and the high efficiency of gene transduction for a variety of cells, especially those with a pseudotype envelope of vesicular stomatitis virus glycoprotein (VSV-G) [11,12]. Single infection makes it possible to express the inserted gene successively and select cells with the desired phenotype with any functional assay. With these advantages, hematopoietic cell

* Corresponding author. Fax: +81-52-744-2161.
E-mail address: aakihiro@med.nagoya-u.ac.jp (A. Abe).

0006-291X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2004.06.040

lines have become available for expression cloning. Many strategies to screen specific cDNA have been reported based on cell proliferation, immunological screening, protein localization, or phenotype complementation [13–16].

Protein oligomerization is critical to the conformational change of a functionally active protein for cell proliferation, differentiation, and the prevention of apoptosis. Receptor tyrosine kinase with adaptation to ligand or non-receptor protein tyrosine kinase with adaptation to the intracellular domain of tyrosine kinase-associated receptors induces an oligomer (dimer) and activates their function [17,18]. TEL is a transcriptional repressor [19,20] that fuses with receptor or non-receptor tyrosine kinase in many hematological malignancies [21–28]. These fusion proteins form an oligomer mediated by the *TEL* pointed self-association motif (PNT) and lead to the constitutive activation of the tyrosine kinase. In this report, we describe a method to clone constitutively active fusion genes which were artificially constructed by fusing the TEL PNT domain to a cDNA library and used to transform a cytokine-dependent hematopoietic cell line, 32D, into a cytokine-independent line. We termed this novel method OLAG-REX (oligomerization activated gene cloning by retrovirus-mediated expression screening).

Materials and methods

Cell culture. Interleukin-3 (IL-3)-dependent murine myeloid cells, 32D, were maintained in Iscove's modified Dulbecco's medium (IMDM) with 10% fetal bovine serum (FBS) (Sanko Junyaku, Tokyo, Japan) and 2 ng/ml of murine IL-3 (mIL-3) [29]. Recombinant mIL-3 was kindly provided by Kirin Brewery (Tokyo, Japan). 293 or 208F cells were maintained in IMDM with 10% FBS. Cultures were performed at 37 °C in a 10% CO_2-humidified atmosphere in an incubator.

DNA constructs and expression plasmids. A CMV early promoter was inserted into a pLRNL [12] retrovirus vector cleaved with blunted *Eco*RI and *Sac*I resulting in pCLRNL. pCLNRL was digested with *Hin*dIII and self-ligated to make pCLNL. The *Bst*BI–*Cla*I fragment of the GAPDH promoter was inserted into pCLNL cleaved with *Bst*BI and *Cla*I resulting in pCLNGAP. The replication origin of pCLNGAP was replaced with that of the pBluescript KS(+) vector (Stratagene, La Jolla, CA) in order to make a high-copy plasmid and T3–T7 fragment including the multi-cloning site of pBluescript KS(+) vector inserted in the blunted *Cla*I site. This vector was designated pHNGAP. The GAPDH promoter was created by polymerase chain reaction (PCR) with the primers (GAPDHPRF; 5′-ATTTCGAACAATTGCCTGTT CCCACCGTGTGCCCAAGA-3′, GAPDHPRR; 5′-ATCGATTGTC GAACAGGAGGAGCAGAGAGC-3′). The PCR fragment was digested with *Bst*BI–*Cla*I for ligation. Flag-PNT (FPNT) or HA-PNT (HPNT) was cloned from pF-TS or pH-TS [26], including the helix–loop–helix oligomerization domain of TEL by PCR amplification, respectively, and inserted into pHNGAP cleaved with *Sal*I and *Eco*RI, resulting in pHNGAP/FPNT or pHNGAP/HPNT. The Flag-tagged TELΔPNT mutant was generated by PCR amplification as described previously [26] and inserted into pHNGAP, resulting in pHNGAP/FΔPNT.

PNT-deleted fusion genes were made by inserting an *Eco*RI–*Not*I fragment of the fusion genes into the *Eco*RI–*Not*I site of pHNGAP/FΔPNT. The blasticidin *S* deaminase (BSD) coding region was cloned by PCR amplification from pUCSV-BSD (Kaken, Urayasu, Chiba, Japan) with the primers (BSDF; 5′-TTAAGCTTATGGCCAAGCC TTTGTCTCA-3′, BSDR; 5′-AACAATTGTTAGCCCTCCCACAC ATAAC-3′). The neomycin resistance gene of pHNGAP/HPNT was replaced with BSD, resulting in a pHBGAP/HPNT vector. For the construction of pCGCGP, an *Eag*I fragment of pHCMV-G, including a CMV early promoter and vesicular stomatitis virus envelope glycoprotein (VSV-G), was inserted into pHCMV-GP [30,31] cleaved with *Not*I.

Construction of the cDNA library. Total RNA was isolated from a CMMoL patient's bone marrow mononuclear cells with informed consent by the guanidium thiocyanate method. Poly(A)$^+$ RNA was purified using oligo(dT)-Latex (Daiichikagaku, Tokyo, Japan). Double-stranded cDNA was synthesized by the method of Gubler Hoffman with minor modifications using the *Not*IT18 primer, 5′-AATGCGGC CGCTTTTTTTTTTTTTTTTT-3′. After attachment of the *Eco*RI adapter (Promega, Madison, WI), double-stranded cDNA was digested with the *Not*I enzyme and size-fractionated by electrophoresis on a 1.2% agarose gel. cDNAs longer than 800 bp were ligated to the pHNGAP/FPNT vector. Double-stranded sequencing of the plasmid insert was performed using a ABI PRISM310 Genetic Analyzer (Applied Biosystems, Foster City, CA).

Infection and transient expression of recombinant retrovirus. 293 cells (5×10^6) were seeded onto 10-cm dishes 1 day before the transfection. To generate pseudotype viruses, we co-transfected 10 μg the cDNA library plasmid and 10 μg of pCGCGP into 293 cells using calcium phosphate co-precipitation. The culture medium was replaced with 8 ml of fresh medium 8 h after transfection and the pseudotype virus was collected 48 h after transfection. The virus titer was determined on rat 208F fibroblasts as described [32]. To screen the retrovirus library, 1×10^6 32D cells were infected with 8 ml of these virus supernatants in the presence of 5 μg/ml protamine. To isolate IL-3-independent 32D, cells were washed with PBS twice, resuspended in IL-3-free medium, and seeded on 96-well plates 24 h after infection. Actively proliferating cells were subjected to further analysis 2–4 weeks after depletion of IL-3. G418-resistant 32D cells expressing the indicated proteins and control cells were grown in RPMI 1640 medium supplemented with 10% FBS in IL-3. Cells were washed twice in the same IL-3-free medium and then seeded in IL-3-free medium in 12-well plates at a concentration of 1×10^5 cells/ml. The number of cells showing proliferation in either the absence or presence of IL-3 was scored every day for 3 days in culture.

Isolation of cDNA fragments from 32D transformed with integrated retroviruses. Genomic DNAs extracted from 32D clones with proliferation without IL-3 were subjected to PCR to recover the integrated cDNAs by using vector primers (T3; 5′-AATTAACCCTCACTAA AGGGA-3′, T7; 5′-GTAATACGACTCACTATAGGGC-3′). PCR amplification was performed with *Thermus aquaticus* DNA polymerase (Perkin–Elmer) for 35 cycles (40 s at 94 °C for denaturation, 40 s at 60 °C for annealing, and 3 min at 72 °C for extension). The resulting PCR fragments were sequenced using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) and analyzed as described above. cDNAs recovered by PCR amplification from transformed 32D cells were introduced into 32D with a pHNGAP/FPNT retrovirus vector and the transforming capability was confirmed.

Antibodies. Anti-phosphotyrosine MoAb horseradish peroxidase (HRP)-conjugated PY-20 was purchased from ICN Biochemicals (Aurora, OH), anti-FLAG MoAb was from Sigma (St. Louis, MO), and HRP-conjugated anti-HA MoAb was from Roche Molecular Biochemicals (Indianapolis, IN).

Immunoprecipitation and Western blotting. Cells were solubilized in lysis buffer (20 mM Tris–HCl, pH 7.4, 0.1% sodium dodecyl sulfate (SDS), 1% Triton X-100, and 1% sodium–deoxycholate). The cell lysates were pre-cleared by incubation with protein G–Sepharose (Amersham–Pharmacia Biotech, Tokyo, Japan) overnight and then centrifuged. The supernatants were incubated with the indicated

antibodies for 1 h followed by 20 μl of 50% (v/v) protein G–Sepharose for 2 h at 4 °C. The immunoprecipitates were washed three times in lysis buffer. For immunoblotting, the precipitates were boiled with electrophoresis SDS sample buffer for 3 min. Whole cell lysates or immunoprecipitates were separated by SDS–PAGE and transferred onto polyvinylidine difluoride (PVDF) membranes (Bio-Rad; Hercules, CA). The membrane blots were blocked with 5% skim milk in Tris-buffered saline (TBS) containing 0.1% Tween 20 (TBS-T) or with a commercial blocking reagent (Roche Molecular Biochemicals) to detect PY-20 for 1 h at 37 °C and then incubated with the primary antibodies in TBS-T for 2 h at room temperature. Following washing, the membranes were incubated with an HRP-linked whole anti-mouse IgG antibody (Amersham–Pharmacia Biotech) in TBS-T for 2 h at room temperature. After washing, an enhanced chemiluminescence (ECL) assay (Amersham–Pharmacia Biotech) was performed and the positive bands were detected on X-ray films.

Results

Retrovirus vector constructs

A retrovirus vector plasmid, pHNGAP/FPNT, which enabled expression of the fusion gene of TEL with inserted cDNA, was constructed for transferring the cDNA library and expression cloning (Fig. 1). We designed a novel retrovirus packaging system in which a high-titer VSV-G pseudotype retrovirus was produced without the need for stable packaging cells. Retrovirus vector plasmid and pCGCGP, which expresses VSV-G and retrovirus *gag-pol* genes on a single plasmid, were transfected into 293 cells followed by virus harvest 48 h post-transfection. The estimated titers of the retrovirus were 1×10^6 colony forming units/ml based on G418-resistant colony formation of the infected 208F cells.

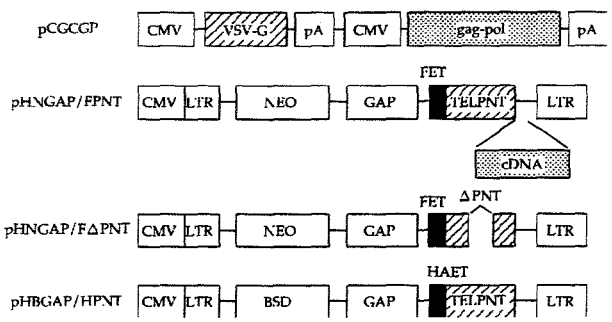

Fig. 1. A schematic representation of pCGCGP, pHNGAP/FPNT, pHNGAP/FΔPNT, and pHBGAP/HPNT vectors. pHNGAP/FPNT, which was constructed by inserting Flag epitope-tagged TEL under the control of the GAPDH promoter, was designed to construct a TEL-cDNA fusion library. cDNAs can be inserted into *Eco*RI–*Not*I sites for fusion with TEL. VSV-G and retrovirus *gag-pol* were expressed from pCGCGP under the control of a separate CMV promoter. LTR, long terminal repeat; NEO, neomycin resistance gene; GAP, GAPDH promoter; FET, Flag epitope tag; CMV, CMV promoter; VSV-G, vesicular stomatitis virus glycoprotein; pA, SV40 polyadenylation signal; BSD, blasticidin *S* deaminase resistance gene; and HAET, HA epitope tag.

This retrovirus production system is of general use for the rapid production of high-titer viral supernatants.

Cloning of oligomerization-activated genes

cDNA encoding Flag-tagged partial TEL including the PNT domain was inserted downstream of the GAPDH promoter. The *Eco*RI and *Not*I sites downstream of PNT enabled the site directional insertion of cDNA fragments, creating TEL-cDNA fusion libraries. Fusion genes with transforming potential were screened according to the cytokine-independent growth of 32D. The complementary cDNAs cloned by OLAG-REX are listed in Fig. 2. All fusion proteins have proper reading frames with TEL. We identified 25 independent fusions proteins with seven kinds of protein kinase genes. Six out of the 7 protein kinases were tyrosine kinase, TEL/LYN, TEL/HCK, TEL/FGR, TEL/SYK, TEL/FLT3, and TEL/TYK2. A serine/threonine kinase, ARAF1, was also identified to fuse with TEL. TEL/SYK, identified in a hematological malignancy, showed CMMoL-like myelodysplastic syndrome [26] and TEL/FLT3, TEL/TYK2 have been artificially constructed and described [33,34].

Proliferation assay

To confirm the transforming capacities of these proteins, each fusion gene was transduced into 32D cells using pHNGAP retroviral vector. We selected one fusion protein from each and determined the number of fused amino acids. These transduced cells showed cytokine-independent proliferation although the proliferating ability varied (Fig. 3). TEL/LYN-232 and TEL/FGR-221 showed similar growth to IL-3-positive 32D, and TEL/SYK-365, TEL/FLT3-605, and TEL/TYK2-776 grew slightly slower, maybe because some populations had an adhering feature. TEL/HCK-227 and TEL/ARAF1-303 showed slower proliferation. This transforming capacity depended on the PNT domain of TEL analyzed in TEL/LYN-232, TEL/SYK-365, and TEL/FLT3-605.

Constitutive phosphorylation and oligomerization

Tyrosine phosphorylation was examined in TEL/LYN-232, TEL/SYK-365, and TEL/FLT3-605 immunoprecipitated by the Flag antibody. Constitutively, the tyrosine phosphorylation of TEL/LYN-232, TEL/SYK-365, and TEL/FLT3-605 was observed in 32D cells transformed with these fusion genes, which were dependent on the PNT domain of TEL (Fig. 4A). The tyrosine phosphorylation of TEL/HCK-227, TEL/FGR-221, and TEL/TYK2-776 was also detected by Western blotting of whole cell lysates with an anti-phosphotyrosine antibody, however, the tyrosine phosphorylation of TEL/ARAF1-303 was not detected since ARAF1 is a

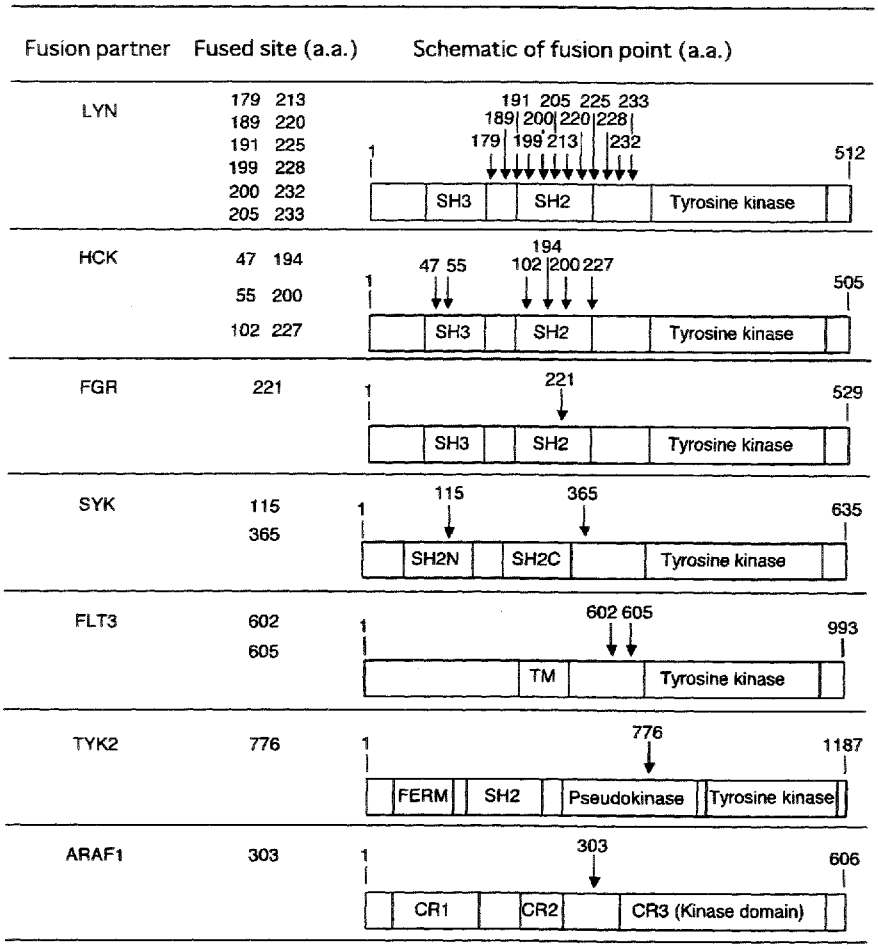

Fig. 2. cDNAs isolated using OLAG-REX. Fused sites are represented by the amino acid number and indicated on protein structural schema. All protein kinase fusions include kinase domains and reading frames are matched with TEL. SH2, src homology-2; SH3, src homology-3; TM, transmembrane; FERM, 4.1, ezrin, radixin, and moesin homology region; and CR, conserved region.

serine/threonine kinase (Fig. 4C). The intensity of the bands detected with an anti-Flag antibody was dependent on the fusion protein, probably due to the difference in conformational affinity for the Flag epitope or stability for protein degradation.

We compared the TEL-mediated oligomerization properties of TEL/LYN-232 and TEL/SYK-365 (Fig. 4B). Flag-tagged TEL/LYN-232 (FTL) or PNT-deleted TEL/LYN-232 (ΔFTL) was co-transduced into 32D with HA epitope-tagged TEL/LYN-232 (HTL) to establish FTL/HTL or ΔFTL/HTL. FTS/HTS and ΔFTS/HTS were similarly established. Transduced proteins were immunoprecipitated using the Flag-specific monoclonal antibody and immunoprecipitated proteins were analyzed with HRP-conjugated anti-HA MoAb. HTL or HTS was co-immunoprecipitated with FTL or FTS, however, neither HTL nor HTS was co-immunoprecipitated with ΔFTL or ΔFTS. These results suggest that TEL/LYN-232 and TEL/SYK-365 are oligomerized through the PNT domain.

Discussion

We have established a retrovirus-mediated cDNA expression cloning method, OLAG-REX, by which cDNAs can be isolated based on oligomerization-mediated activation resulting in the cytokine-independent proliferation of 32D cells. Protein oligomerization is contrived with the helix–loop–helix region of TEL, which is frequently fused in the chromosomal translocation in hematological malignancies. A pseudotype retrovirus was produced transiently using the original system in which a retrovirus vector plasmid and pCGCGP expressing VSV-G and the *gag-pol* gene were cotransfected into 293 cells, therefore, the establishment of packaging cell lines was not required. We showed here that our expression cloning strategy is a powerful and efficient method to isolate a constitutively active protein kinase fused to TEL.

To obtain oligomerization-activated genes which promote the cytokine-independent growth of 32D, two