An unusual arrangement of 13 zinc fingers in the vertebrate gene *Z13*

Thomas C. SCHULZ, Blair HOPWOOD,* Peter D. RATHJEN† and Julian R. E. WELLS‡
Department of Biochemistry, University of Adelaide, Adelaide, S.A. 5005, Australia

The zinc finger is a protein domain that imparts specific nucleic acid-binding activity on a wide range of functionally important proteins. In this paper we report the molecular cloning and characterization of a novel murine zinc-finger gene, mZ13. Analysis of mZ13 cDNAs revealed that the gene expresses a 794-amino-acid protein encoded by a 2.7 kb transcript. The protein has an unusual arrangement of 13 zinc fingers into a 'hand' of 12 tandem fingers and a single isolated finger near the C-terminus. This structural organization is conserved with the probable chicken homologue, cZ13. mZ13 also contained an additional domain at the N-terminus which has previously been implicated in the regulation of zinc-finger transcription factor DNA-binding, via protein–protein interactions. mZ13 expression was detected in a wide range of murine embryonic and adult tissues. The structural organization of mZ13 and its expression profile suggest that it may function as a housekeeping DNA-binding protein that regulates the expression of specific genes.

INTRODUCTION

Nucleic acid-binding proteins can be classified on the basis of protein motifs and/or structures which are required for binding activity. Members of the C2-H2 zinc-finger family conform to the consensus sequence Cys-Xaa$_{2-4}$-Cys-Xaa$_3$-Phe-Xaa$_5$-Leu-Xaa$_2$-His-Xaa$_{3-5}$-His [1,2]. Two cysteine and histidine residues tetrahedrally coordinate a Zn^{2+} ion, providing a 'miniglobular' structural unit for individual fingers [3] of an antiparallel β ribbon and an α helix. Sequence-specific DNA-binding is generated by insertion of the α helix into the major groove at the target site, with residues in the N-terminus of the α helix contacting the bases [4].

Many proteins contain multiple C2-H2 zinc fingers. The arrangement of individual fingers within different proteins is variable although the most common format is a single cluster of multiple tandem repeats organized as a 'hand' (reviewed in [5]). Other patterns have been described, including proteins with individual zinc fingers [6], multiple discrete groups of zinc-finger clusters organized as separate 'hands' [7], and individual fingers dispersed throughout the protein [8]. Tandem repeats of C2-H2 zinc fingers are separated by 7–8 amino acids which are termed the His/Cys link. In proteins that are members of the Krüppel subclass of zinc-finger proteins these residues are a variant of the consensus TGEKP(F,Y) [9]. Other subfamilies of C2-H2 zinc-finger proteins can be defined on the basis of conserved structural or sequence motifs outside the finger regions. These include the N-terminal KRAB [10], FAX [11] and POZ domains [12,13], and the presence of homeodomains [14]. It is not known whether zinc-finger subfamilies define functionally related proteins.

The versatility of the C2-H2 zinc finger in sequence and motif organization, including combination with other functional motifs, enables it to be involved in a wide range of processes involving binding to both DNA and RNA. While zinc-finger proteins have been demonstrated to have functions as diverse as heterochromatin packaging [8] and nucleocytoplasmic RNA transport [15], the majority of C2-H2 zinc-finger proteins appear to be transcription factors. These include housekeeping factors such as Sp1 [16] and developmentally regulated factors such as *Krüppel* [17], *hunchback* [18] and Krox-20 [19]. Functional studies of various zinc-finger transcription factors have revealed versatile actions including both positive and negative transcriptional regulation, and complex interactions with other transcription factors (reviewed in [5]).

As part of a screen for DNA-binding proteins we identified a novel chicken zinc-finger protein termed Z13 (B. Hopwood and J. R. E. Wells, unpublished work; Genbank U14555). In this paper we report the isolation of murine Z13 cDNA and genomic clones, and preliminary expression analysis of the protein in embryonic and adult mice.

MATERIALS AND METHODS

Isolation, subcloning and sequencing of mZ13 genomic and cDNA clones

To generate murine-specific sequence and probes for the isolation of mZ13 cDNAs, a Balb/C EMBL3 murine genomic library (Clontech) was screened using standard procedures [20] with the 2.2 kb cZ13 cDNA (B. Hopwood and J. R. E. Wells, unpublished work; Genbank U14555), oligolabelled with [α-^{32}P]dATP (Gigaprime kit, Bresatec Ltd). A third-round duplicate positive plaque was isolated and phage DNA was purified by the liquid lysate method [20], giving a 13 kb mZ13 genomic clone (λmZ13). *Bam*HI and *Sal*I/*Bam*HI (Pharmacia) restriction fragments of λmZ13 that hybridized to cZ13 by Southern blot [20] were subcloned into pBluescript SK or pBluescript II KS (Stratagene) for sequencing (T7 sequencing reagent kit, Bresatec Ltd; T7 Sequenase, USB). Murine sequence corresponding to cZ13 was identified and used to design primers for reverse transcription and PCR amplification of mZ13 cDNAs from 16 days post coitum (d.p.c) embryonic kidney RNA from CBA strain mice.

Abbreviations used: d.p.c., days post coitum; RACE, rapid amplification of CDNA ends; RT, reverse transcriptase; SSC,0.15 M NaCl/0.015 M sodium citrate.

* Present address: Department of Gene Regulation, Roche Institute of Molecular Biology, Nutley, NJ, U.S.A.
† To whom correspondence should be addressed.
‡ Deceased.

The 5′ region of mZ13 was amplified by a modification of rapid amplification of cDNA ends (RACE)-PCR using the 5′-ampliFINDER RACE kit (Clontech). Poly(A)$^+$ RNA (2 μg) was reverse transcribed with primer 5′-TCTTGCCACACTCCTCA-CAC-3′ corresponding to nucleotide residues 1279–1298 of the mZ13 cDNA. The anchor sequence 3′-NH_3-GGAGACTTCC-AAGGTCTTAGCTATCACTTAAGCAC-P-5′ was ligated on to the 3′ end of the cDNA. The primers used for PCR were a 3′ primer: 5′-CTTGTGGATGATGGAGCCGT-3′ (nucleotides 1092–1111 of the mZ13 cDNA), and a 5′ anchor primer: 5′-CTGGTTCGGCCCACCTCTGAAGGTTCCAGAATCG-ATAG-3′ (Clontech) which hybridized to the anchor sequence. DNA amplification was performed with Pfu DNA polymerase (Stratagene) according to manufacturer's instructions, in a capillary thermal cycler (Corbett research) for 35 cycles of denaturation (94 °C, 5 s), annealing (60 °C, 5 s), and extension (72 °C, 5 min). After gel electrophoresis, the mZ13 cDNA (5′) fragment was extracted (Geneclean kit, Bio101) and subjected to a further 35 cycles of DNA amplification under the same conditions, except the annealing temperature was 50 °C and the 3′ primer was modified to contain a *Xho*I site: 5′-GGCCCTC-GAGCTTGTGGATGATGGAGCCGT-3′. The product was digested with *Eco*RI/*Xho*I, gel purified and cloned into pBluescript II KS. The cDNA insert was restricted with *Alu*I, *Hpa*II, *Hae*III and *Pvu*II, and subcloned into pBluescript II KS for sequencing.

The 3′ region of mZ13 was amplified from poly(A)$^+$ RNA by reverse transcriptase (RT)-PCR [21] using AMV reverse transcriptase (Molecular Genetic Resources) and Pfu DNA polymerase. First-strand cDNA synthesis was primed with the oligonucleotide 5′-CCACAGGCAACATTAGAAAATAATC-3′ (mZ13 cDNA nucleotides 2689–2701; the unmatched 5′ end of the oligonucleotide corresponded to the sequence of the genomic clone in this region). PCR amplification was carried out using the 5′ primer 5′-ACGGCTCCATCATCCACAAG-3′ (mZ13 nucleotides 1092–1111) and the same 3′ primer as used for reverse transcription. DNA amplification was performed as described, for 35 cycles of denaturation (94 °C, 5 s), annealing (54 °C, 5 s), and extension (72 °C, 4 min). The mZ13 cDNA 3′ fragments were end-filled with T4 DNA polymerase [20] and cloned into pBluescript II KS for sequencing.

Exon-containing regions of the mZ13 genomic clone (λmZ13) were identified by hybridization to mZ13 cDNAs. These regions were subcloned into pBluescript II KS and sequenced, or sequenced with mZ13-specific oligonucleotides. Exon X (see Figure 4a) was not identified by this analysis, and was amplified by PCR from λmZ13 using the conditions described for RT-PCR using the 5′ primer: 5′-GAAGCAGTTCACCACCTCAG-3′ (mZ13 cDNA position 1529–1649) and the 3′ primer 5′-ACAGGTCCCCCACATTC-3′ (mZ13 cDNA position 1977–1994). The 90 bp *Bam*HI-digested product was cloned into pBluescript KS and sequenced.

Sequence comparison

The mZ13 nucleotide and conceptually translated open reading frame sequences were compared with the GenBank, PIR and Swiss-Prot databases using Hitashi MacDNASIS software.

Isolation of embryonic and adult RNAs

Tissue samples from 10.5, 12.5 and 16.5 d.p.c. embryos and adult tissues were isolated from CBA strain mice. Total RNA was isolated from these samples by the method of Chomczynski and Sacchi [22]. Poly(A$^+$) RNA was isolated directly from tissue samples with the On Track mRNA isolation kit (Biotecx).

Genomic Zooblot analysis

A genomic Southern Zooblot (Clontech) was the kind gift of Professor Robert B. Saint (Department of Genetics, University of Adelaide, Australia). The filter was prehybridized in 50% (v/v) deionized formamide, 1% SDS (w/v), 10% (w/v) poly-(ethylene glycol), 1 M NaCl, 50 mM Tris/HCl, pH 7.5, 5× Denhardt's, 100 μg/ml sonicated denatured salmon sperm DNA, at 37 °C for 48 h. A 253 bp *Kpn*I/*Bam*HI probe, containing the mZ13 exon XIV and the 5′ portion of exon XV (position 2231–2405), was isolated from a *Bam*HI subclone of the mZ13 genomic clone. This fragment encoded the 13th zinc finger (see Figure 4a). The probe was oligolabelled with [α-^{32}P]dATP and hybridized to the filter at 37 °C for 48 h. The filter was washed at low stringency in 6× SSC (1× SSC: 0.15 M NaCl, 0.015 sodium citrate)/0.1% SDS at 42 °C for 10 min and in 2× SSC/0.1% SDS at 65 °C for 20 min, followed by high stringency washing in 0.5× SSC/0.1% SDS at 65 °C for 20 min, and was exposed to Konica medical grade X-ray film at −80 °C with an intensifying screen for 7 days.

Antisense riboprobes and RNase protection analysis

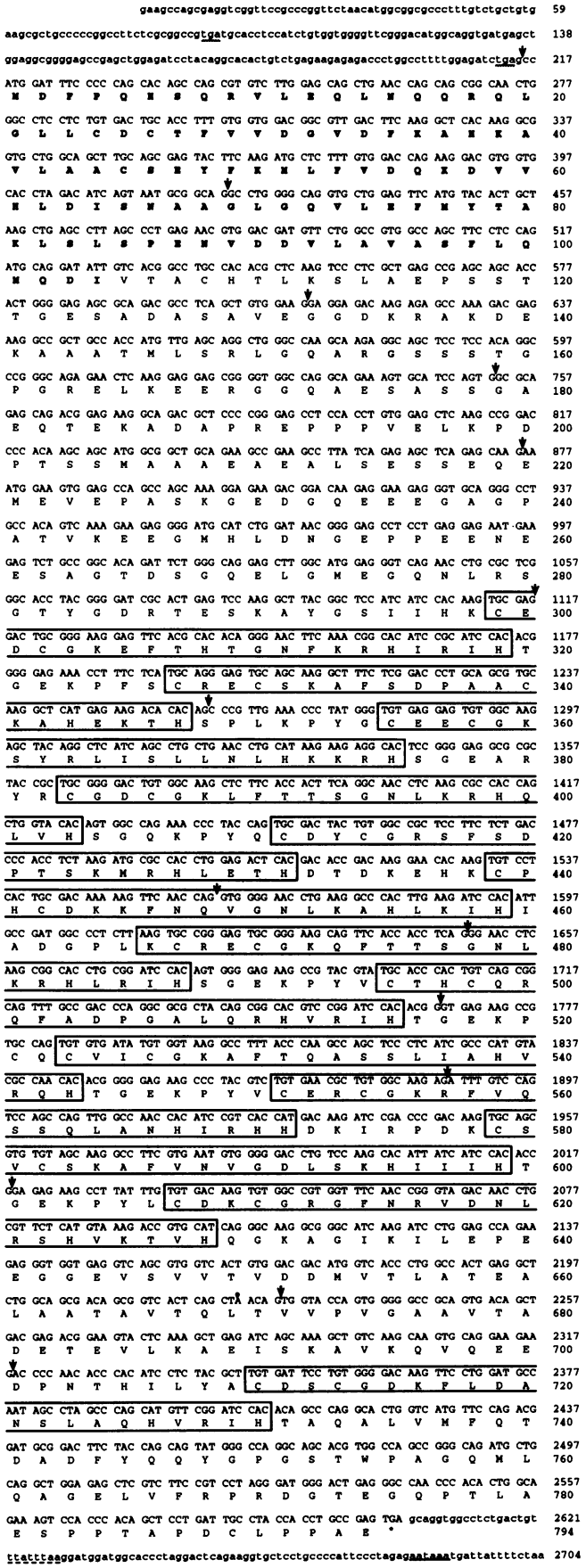
An mZ13-specific riboprobe was generated from the mZ13 genomic clone. Exon XIII (see Figure 4a) was amplified from a λmZ13 *Bam*HI subclone by PCR using the 5′ primer 5′-GGAGAGGATCCTTATTTGTGTGAC-3′ (cDNA nucleotides 2018–2041) and the 3′ primer 5′-GTGAGCTCGAGTGACCG-CTGTCGCTG-3′ (cDNA nucleotides 2202–2226). PCR amplification was performed with *Taq* DNA polymerase (Bresatec Ltd) according to the manufacturer's instructions, in a thermal cycler (MJ research) for 30 cycles of denaturation (94 °C, 1 min), annealing (54 °C, 30 s), and extension (72 °C, 1 min). The product was digested with *Bam*HI/*Xho*I, giving a 194 bp fragment, gel-purified and cloned into *Bam*HI/*Xho*I-digested pBluescript II KS (Stratagene). An mZ13-specific antisense riboprobe of 192 bp (plus 30 bp of pBluescript II KS polylinker) was generated by transcription with T3 RNA polymerase (Promega) [23] after linearization with *Bam*HI in transcription reactions containing 100 μCi of [α-^{32}P]rUTP (Bresatec Ltd). Unincorporated radioactive label was removed from the reaction products by denaturing PAGE [20]. The purified antisense probe was eluted in 50% (v/v) formamide for 2 h at 37 °C and used in RNase protection assays as described by Krieg and Melton [23], except that 120000 c.p.m. of probe was added to each RNA sample. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as a loading control [24]. RNase digestion products were resolved by denaturing PAGE [20], and visualized by exposure to Konica medical-grade X-ray film at −80 °C with an intensifying screen for 14 days.

RESULTS

mZ13 cDNA isolation and characterization

A murine Z13 (mZ13) genomic clone was isolated from a Balb/C EMBL3 genomic library by cross-hybridization with a cZ13 cDNA. mZ13-specific sequence was derived from this clone and used for the isolation of 5′ and 3′ regions of mZ13 from 16 d.p.c., CBA strain, embryonic kidney poly(A)$^+$ RNAs by RACE-PCR and RT-PCR respectively. The compiled sequence of the mZ13 cDNA (Genbank accession U14556) and the deduced amino acid sequence of the mZ13 protein are presented in Figure 1.

An open reading frame of 794 amino acids extended from position 218 bp to position 2600 bp. The identity of the initiation codon for this open reading frame in mZ13 was confirmed by the



presence of two in-frame TGA stop codons (light underline, Figure 1) located upstream of the ATG, at positions 89 bp and 212 bp, and a consensus Kozak sequence [25] surrounding the ATG.

Conceptual translation of the mZ13 open reading frame revealed the presence of 13 C2-H2-type zinc fingers. Twelve of these were clustered in repeated fashion, separated by seven amino acids conforming to the consensus TGEKP(F,Y). This arrangement, which shared 96% amino acid identity with the chicken Z13 (Figure 2a, 297 to 628 of mZ13, 205 to 536 of cZ13), is typical of many clustered fingers and identifies Z13 as a member of the *Krüppel*-like subfamily of zinc-finger proteins [17]. A region of lesser but significant sequence similarity downstream of the 12th zinc finger included the 81-amino-acid spacer region in mZ13 between the 12th and 13th zinc fingers, and the 13th zinc finger. The isolated position of the 13th zinc finger within the Z13 open reading frame was conserved between mouse and chicken. This conservation of sequence and organization within the putative DNA-binding domain indicates that the cZ13 and mZ13 proteins are probably sequence homologues and share a common evolutionary origin.

Sequence similarity between cZ13 and mZ13 upstream of the first zinc finger and downstream of the 13th zinc finger was sporadic. However, using the best alignment of these proteins (DNASIS) the termination codon for the mZ13 open reading frame corresponded to that of cZ13 (Figure 2a), the poly(A) addition signal was present in the same position (Figure 2b), and an AU motif, implicated in mRNA instability [26,27], and resembling the consensus TTATTTAT [28], was conserved in sequence and position in the 3′ untranslated regions of mZ13 and cZ13 (dashed underline in Figure 2b, position 2622–2629 bp). Thus the gross organization of the Z13 expression unit was common to the murine and chicken genes.

The 2701 bp mZ13 sequence shown in Figure 1 lacks the poly(A) tail, which was not cloned in this analysis. However, the close structural relationship between important domains and sequences in mZ13 and cZ13 indicates that the murine poly(A) tail would be located immediately downstream of nucleotide 2701. Thus the predicted size of the complete mZ13 transcript is 2701 nucleotides plus the length of the poly(A) tail.

mZ13 contained an additional region of sequence similarity at the N-terminus (amino acids 1–104) to a functional domain found in other zinc-finger proteins, viral non-zinc-finger proteins and cellular non-zinc-finger proteins [12,13]. This domain, which has been termed the POZ domain [13], can modulate zinc-finger DNA-binding activity through dimerization. The most related POZ domain to mZ13 was KUP [29], a human zinc-finger gene with 39.4% amino acid identity to mZ13 over the POZ domain (Figure 3). Within both mZ13 and KUP the POZ domain was located at the extreme N-terminus of the protein. The N-terminus of the cZ13 protein has not been isolated but alignment of the 12 N-terminal cZ13 amino acids with the POZ domain consensus sequence (Figure 3) indicated that the POZ domain structure was

Figure 1 Sequence of the mZ13 cDNA

The 794-amino-acid open reading frame spans nucleotide positions 218 to 2600. The in-frame termination codons (light underline) upstream of the initiation codon are located at nucleotide positions 89 and 212. The poly(A) addition signal (underlined) is located at position 2681 to 2686 and the conserved AU element (dashed underline) is located at position 2622 to 2629. The POZ domain is represented in bold type and individual zinc fingers are boxed. The positions of introns within the genomic clone are indicated by arrows. The silent C (Balb/C)/A (CBA) polymorphism at nucleotide position 2224 is represented with a dot. The termination codon is indicated by an asterisk. The sequence was derived from the ampliFINDER RACE-PCR clone (nucleotides 1 to 1111) and the RT-PCR clone (nucleotides 1092 to 2701).

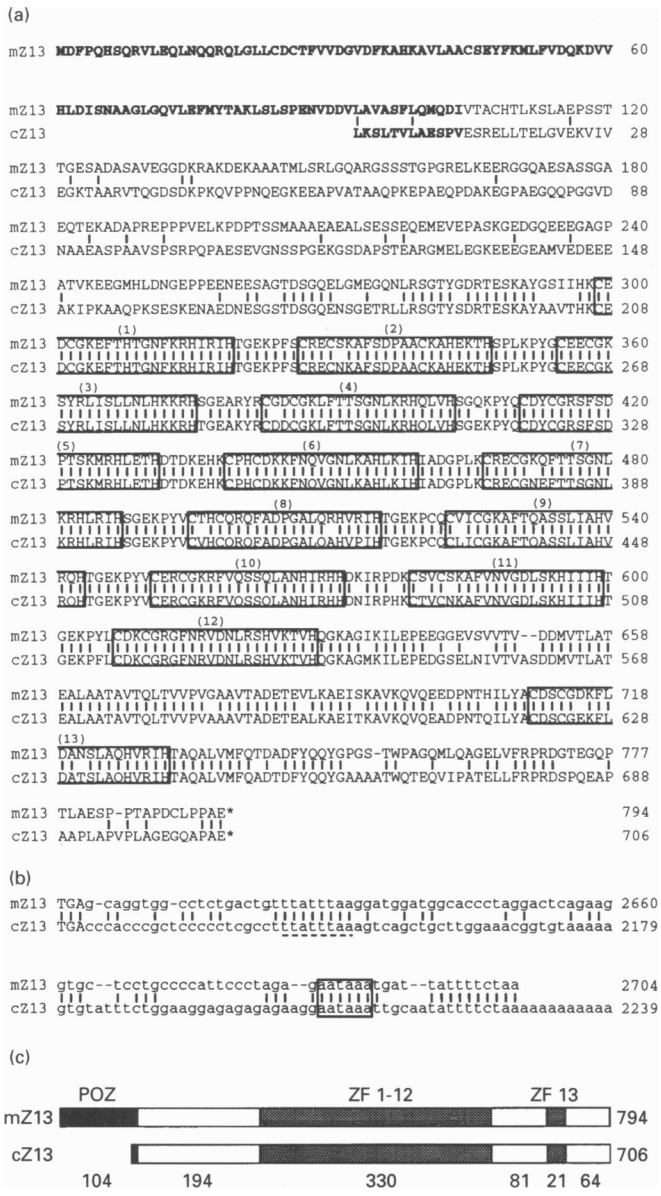


Figure 2 Sequence comparison of mZ13 and cZ13

(**a**) Alignment of the murine and chicken Z13 sequences. Amino acid positions of mZ13 (Figure 1) and cZ13 are indicated on the right. The POZ domain is shown in bold type and individual zinc fingers are numbered and boxed. Identical residues are indicated by vertical lines. Termination codons are indicated by asterisks. The chicken sequence is incomplete at the N-terminus which has not been cloned. (**b**) Alignment of the 3′ untranslated regions of mZ13 and cZ13. The termination codons are in capital text, and numbering refers to nucleotide position within the respective cDNAs. The poly(A) addition signal (boxed), and the conserved AU motif (dashed underline) are indicated. (**c**) Schematic alignment of mZ13 and cZ13 open reading frames. The POZ domain is shaded black, zinc fingers 1 to 12 and zinc finger 13 are shaded. The size of each ORF is indicated to the right. Numbering at the bottom indicates the size of each domain and the intervening regions in the mZ13 protein.

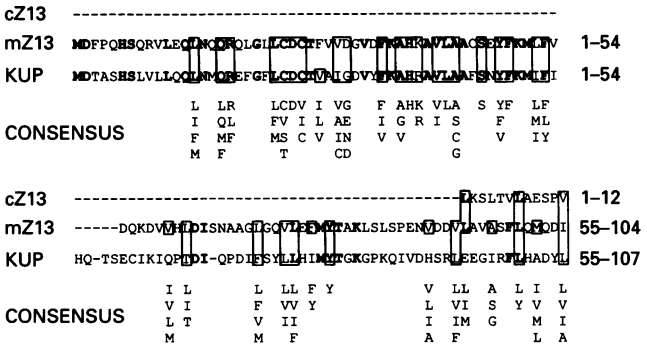


Figure 3 A conserved POZ domain in Z13

Sequence alignment of the POZ domains of mZ13 and KUP [29], and the partial cZ13 POZ domain. Boxed residues correspond to the POZ domain consensus sequence [13]. Residues conserved in mZ13, KUP and cZ13 are shown in bold. Numbering is with respect to the initiation codon of mZ13 and KUP and the start of the cZ13 partial sequence.

also likely to be conserved in the chicken Z13 protein. The existence of this domain within mZ13 has important implications for the biological function of the protein.

Figure 2(c) indicates that the organization of structural and functional domains within the Z13 open reading frame was conserved between the mouse and chicken proteins.

mZ13 genomic structure

A Balb/C EMBL3 murine genomic library (Clontech) was screened with cZ13 cDNA, and a 13 kb mZ13 genomic clone (λmZ13) was isolated. The detailed genomic organization of the mZ13 locus deduced by comparison of the mZ13 cDNA and genomic sequences is shown in Figure 4(a). The genomic clone, derived from Balb/C strain mice, contained a silent base substitution at position 2224 of the CBA cDNA sequence (Figure 1). The substitution of a C for an A residue at this position did not alter the predicted protein and is thought to represent a sequence polymorphism. 213 bp of 5′-untranslated region was not present on the genomic clone. This suggests the existence of an additional intron larger than 3 kb located 4 bp upstream of the mZ13 initiation codon. The entire mZ13 genomic region is thus predicted to occupy at least 8.4 kb. The open reading frame was contained within 14 exons covering 5.4 kb. The 13 introns ranged in size from 79 bp to 550 bp and were flanked by splice donor and splice acceptor sequences (Figure 4b) which conformed to defined consensus sequences [30].

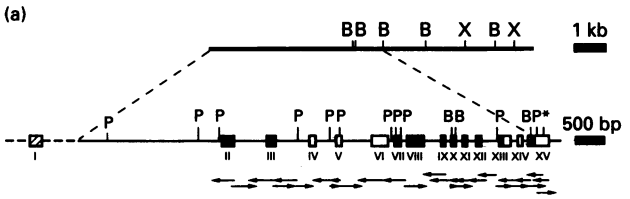
The positions of the introns within the cDNA sequence are indicated by arrows in Figure 1. There was little correlation between the positions of individual functional domains and the positions of introns. Indeed both the N-terminal POZ domain and individual zinc fingers were interrupted by introns.

Conservation of mZ13 homologues

The distribution of Z13 homologues throughout eukaryotes was investigated by Southern blot analysis of *Eco*RI-digested genomic DNAs (Zooblot, Clontech) from a variety of eukaryotic species (Figure 5). Specific hybridization was detected after high stringency washing in rat, mouse, canine, bovine, rabbit and chicken genomic DNAs. A much fainter signal was detected in the yeast DNA. This indicated a widespread distribution of Z13 sequence homologues among eukaryotes.

mZ13 expression in murine tissues

mZ13 expression was investigated by RNase protection analysis in RNA samples generated from 10.5 d.p.c., 12.5 d.p.c. and 16



(b)

EXON 5′ SPLICE SITE		INTRON			EXON 3′ SPLICE SITE	
I	GAG	.....		...tctggctttctgcag	C	II
II	CAG	gtacca...	(550)	...tcctgtcctcgacag	G	III
III	AAG	gtgagg...	(541)	...gtgttctgtctccag	G	IV
IV	GTG	gtgagt...	(264)	...cccgtccttcctcag	G	V
V	AAG	gtgccc...	(440)	...ctctttttctgtgcag	A	VI
VI	GAG	gtgcgg...	(81)	...ctcgtgcgcctgcag	G	VII
VII	CAG	gtaccc...	(86)	...ccttctgacccacag	C	VIII
VIII	CAG	gtgtgc...	(256)	...cctgtctgggtgcag	G	IX
IX	CAG	gtaggc...	(87)	...gcctttcatccaaag	G	X
X	CGG	gtgggt...	(81)	...gtgcctctgccccag	G	XI
XI	GAG	gtagcg...	(134)	...tttgtgaccttgtag	A	XII
XII	CCG	gtgagc...	(240)	...ttgcctccccaccag	G	XIII
XIII	CAG	gtgtgc...	(116)	...tgactctgcccacag	T	XIV
XIV	AAG	gtgaga...	(79)	...cccccacctccacag	A	XV
	C/A AG	GT A/G AGT	CONSENSUS	$(^{T}_{C})_{11}N^{C}_{T}AG$	G	

Figure 4 The genomic structure of mZ13

(**a**) Restriction map and intron/exon structure of mZ13. Restriction fragments containing exons were identified by Southern blot and subcloned, Exon X was isolated by PCR (see Materials and methods section). Subclones were sequenced as indicated by arrows. mZ13 exons are represented by boxed regions, exon I (striped) consists of the 5′ untranslated region and was not identified on the genomic clone. The open reading frame initiates at the start of exon II and the termination codon located in exon XV is indicated by an asterisk. The POZ domain is encoded within exons II and III (black), the 12 clustered zinc fingers (shaded) are encoded within exons VII to XIII, and the 13th zinc finger is encoded within exon XV (shaded). The exact positions of introns on the cDNA are shown in Figure 2. Restriction sites are: B, *Bam*HI; X, *Xho*I; P, *Pst*I. (**b**) Intron/exon boundaries in the mZ13 genomic clone. The 5′ donor and 3′ acceptor sites for the 14 mZ13 introns are listed with consensus donor and acceptor sites [30]. The length of each intron in nucleotides is shown in parentheses. Exons are as indicated in (**a**). The intron between exon I and II is at least 3 kb, The 5′ splice donor sequence of this intron has not been characterized fully.

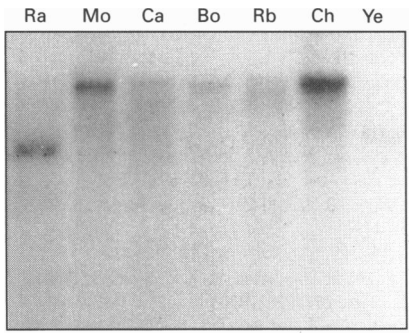


Figure 5 Identification of mZ13 sequences in eukaryote species

A Southern blot of *Eco*RI-digested eukaryotic genomic DNAs was probed with an mZ13 genomic fragment, encoding the 13th zinc finger. Ra, rat; Mo, mouse; Ca, canine; Bo, bovine; Rb, rabbit; Ch, chicken; Ye, yeast.

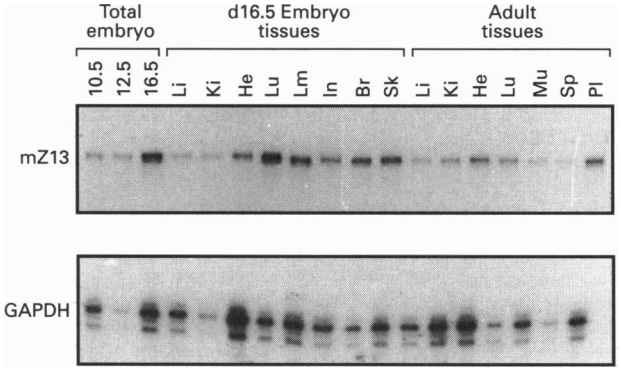


Figure 6 mZ13 expression in embryonic and adult tissues

RNase protections were performed on 20 μg of total embryonic RNA from 10.5, 12.5, and 16.5 d.p.c. CBA embryos, tissue-specific RNA from 16.5 d.p.c CBA embryos and tissue-specific RNA from adult CBA mice. Abbreviations: Li, liver; Ki, kidney; He, heart; Lu, lung; Lm, limbs; In, intestine; Br, brain; Sk, skin; Mu, muscle; Sp, spleen; Pl, placenta. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as a loading control.

d.p.c. murine embryos, and RNA isolated from specific tissues of 16 d.p.c. murine embryos and adult mice (Figure 6). mZ13 transcripts were detected in all the samples assayed, although the levels of expression were variable. The widespread expression profile suggests that the mZ13 gene is not developmentally or temporally regulated and may be a housekeeping gene. Further analysis is required to determine whether mZ13 expression is restricted to subsets of cells, is cell cycle regulated, or is ubiquitous.

DISCUSSION

Sequence and structural analysis of mZ13 indicated that it was a novel murine C2-H2 zinc-finger gene expressed as a transcript of 2.7 kb. Comparison of the mZ13 and cZ13 open reading frames demonstrated a conserved structure of 12 clustered and one isolated C2-H2-type zinc fingers. Homology between mZ13 and cZ13 within the zinc-finger regions was 96%, indicating that the proteins share a common evolutionary origin and may be homologues. As this generates nearly identical zinc-finger domains, it is plausible that mZ13 and cZ13 have very similar DNA target sequences. Although DNA-binding is the most likely function of the C2-H2 zinc-finger domains of mZ13 and cZ13, other functions such as RNA binding cannot be eliminated without experimental analysis.

C2-H2 zinc fingers clustered as 'hands', separated multiple 'hands', individual fingers, or dispersed individual fingers are common arrangements of this motif within a protein [5]. The combination in mZ13 and cZ13 of a 12 zinc-finger 'hand' and an additional isolated finger is unusual. A similar unconventional arrangement in the *Saccharomyces cerevisiae* TFIIIA [31] and murine ΦAP3 [32] proteins has been previously described. However, unlike Z13, the yeast TFIIIA spacer region is not a conserved feature of TFIIIA genes and therefore is not likely to exhibit functional significance. In ΦAP3 [32], the spacer regions between 'hand' and individual finger domains appear to contain vestigial or inactivated fingers in which mutation has caused the loss of the first of the zinc-coordinating histidines (Figure 7). The residual spacing and sequence are consistent with the previous existence of a zinc finger at this position. The spacer region between zinc-finger domains in mZ13 and cZ13 does not appear to have been generated in this manner, as the sequence of this

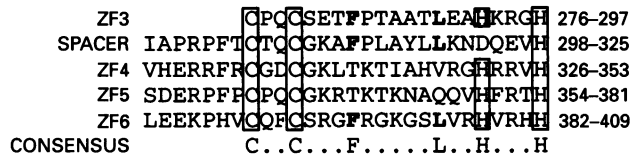


Figure 7 The spacer region between zinc-finger domains of ΦAP3 contains a vestigial zinc finger

Amino acid residues 276 to 409 of ΦAP3 [32] are aligned with the zinc-finger consensus sequence. The invariant cysteine and histidine residues are boxed, and conserved phenylalanine and leucine residues shown in bold. Zinc fingers (ZF) 3, 4, 5 and 6 are indicated. The spacer between ZF3 and ZF4 is of the correct size to represent a zinc finger and contains all the appropriate conserved residues except the first invariant histidine.

region does not bear any hallmark of C2-H2 zinc fingers and does not correlate in length with the existence of inactivated fingers. The conservation of size and sequence of the Z13 spacer region between mouse and chicken genes indicates that it may have functional significance, possibly as a structural feature involved in interaction between the separated finger domains during DNA binding.

mZ13 contained an N-terminal POZ domain [13], placing it in this subclass of *krüppel*-like zinc-finger proteins. The closest related sequence to the POZ domain of mZ13 within this subclass was KUP [29], which shared 39.4% amino acid identity within this domain. Experiments with truncated, chimeric or intact POZ domains of other zinc-finger proteins revealed that the presence of this domain can markedly reduce DNA binding *in vitro* [13]. This effect was demonstrated to result from homodimerization between POZ domains, although some POZ domains were able to form heterodimers. Dimerization and the reduction of DNA binding was demonstrated to be an effect independent of the DNA-binding domain. The presence of a POZ domain in the Z13 sequence therefore indicates that Z13 DNA-binding activity is likely to be modulated by this domain.

Analysis of the genomic structure of mZ13 revealed the presence of numerous introns. While introns positioned directly between zinc fingers were identified, other introns interrupted individual fingers and an intron interrupted the N-terminal POZ domain, thus appearing to contradict the domain hypothesis of intron/exon organization [33].

mZ13 was expressed in all embryonic and adult tissues examined, although there was minor variation in expression levels. It is therefore likely that Z13 is a housekeeping factor in common with other zinc-finger proteins such as Sp1. However, the possibility of cell cycle regulation, or differential expression in subsets of cells cannot be excluded by the RNase protection assay used in this analysis.

The presence of an AU instability sequence in the 3′ untranslated regions of mZ13 and cZ13 indicates the possibility of post-transcriptional regulation. This motif is responsible for the specific targeting of certain lymphokine, cytokine and proto-oncogene mRNAs for degradation. The conserved sequence and location of the AU motif in mZ13 and cZ13 (Figure 2b) indicates that it may have functional activity. The C2-H2 zinc-finger proteins ZFY and ZFX also contain AU repeats in their 3′ untranslated regions [34].

The generalized expression of mZ13 and the presence of transcription factor DNA-binding and protein–protein interaction motifs in mZ13 suggest that mZ13 functions as a housekeeping DNA-binding protein involved in the regulation of specific genes. This assignment is supported by the identification of conserved mZ13 homologues in several eukaryotic genomes, including rat, canine, bovine, rabbit and chicken, which gives an indication of the functional importance of the gene. The availability of full-length cDNA clones and genomic clones will be of importance for elucidation of the biological role of mZ13 by detailed genetic analysis in conjunction with molecular investigation of mZ13 binding sites and target genes.

This work has been supported by an Australian Research Council Grant. This paper is dedicated to the memory of Julian Wells, an inspirational supervisor and friend who is sorely missed.

REFERENCES

1 Miller, J., McLachlan, A. D. and Klug, A. (1985) EMBO J. **4**, 1609–1614
2 Brown, R. S., Sander, C. and Argos, P. (1985) FEBS Lett. **186**, 271–274
3 Lee, S. M., Gippert, G. P., Soman, K. V., Case, D. A. and Wright, P. E. (1989) Science **245**, 635–637
4 Pavletich, N. P. and Pabo, C. O. (1991) Science **252**, 809–817
5 El-Baradi, T. and Pieler, T. (1991) Mech. Dev. **35**, 155–169
6 Baldarelli, R. M., Mahoney, P. A., Salas, F., Gustavson, E., Boyer, P. D., Chang, R.-F., Roark, M. and Lengyel, J. A. (1988) Dev. Biol. **135**, 85–95
7 Ruiz i Altaba, A., Perry-O'Keefe, H. and Melton, D. A. (1987) EMBO J. **6**, 3065–3070
8 Reuter, G., Giarre, M., Farah, J., Gausz, J., Speirer, A. and Speirer, P. (1990) Nature (London) **344**, 219–223
9 Schuh, R., Aider, W., Gaul, U., et al. (1986) Cell **47**, 1025–1032
10 Belleftoid, E. J., Poncelet, D. A., Lecocq, P. J., Revelant, O. and Martial, J. A. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 3608–3612
11 Knöckel, W., Pöting, A., Köster, M., El-Baradi, T., Nietfeld, W., Bouwmeester, T. and Pieler, T. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 6097–6100
12 Bihui, H. Y., Lista, F., Lo Coco, F., Knowles, D. M., Offit, K., Chaganti, R. S. K. and Dalla-Favera, R. (1993) Science **262**, 747–750
13 Bardwell, V. J. and Treisman, R. (1994) Genes Dev. **8**, 1664–1667
14 Fortini, M. E., Lai, Z. and Rubin, G. M. (1991) Mech. Dev. **34**, 113–122
15 Guddat, U., Bakken, A. and Pieler, T. (1990) Cell **60**, 619–628
16 Kadonaga, J. T., Carner, K. R., Masiarz, F. R. and Tijan, R. (1987) Cell **51**, 1079–1090
17 Rosenberg, U. B., Schröder, C., Preiss, A., Kienlin, A., Côté, S., Riede, I. and Jäckle, H. (1986) Nature (London) **319**, 336–339
18 Tautz, D., Lehmann, R., Schnürch, H., Schuh, R., Seifert, E., Kienlin, A., Jones, K. and Jäckle, H. (1987) Nature (London) **327**, 383–389
19 Chavrier, P., Lemaire, P., Revelant, O., Bravo, R. and Charnay, P. (1988) Mol. Cell. Biol. **8**, 1319–1326
20 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
21 Austbel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (1994) Current Protocols in Molecular Biology, vol 2., Current Protocols, U.S.A.
22 Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. **162**, 156–159
23 Krieg, P. A. and Melton, D. A. (1987) Methods Enzymol. **155**, 397–415
24 Lints, T. J., Parsons, L. M., Hartly, L., Lyons, I. and Harvey, R. P. (1993) Development **119**, 419–431
25 Kozak, M. (1989) J. Cell Biol. **108**, 229–241
26 Shaw, G. and Kamen, R. (1986) Cell **46**, 659–660
27 Brawerman, G. (1987) Cell **48**, 5–6
28 Kruys, V., Wathelet, M., Poupart, P., Contraries, R., Fliers, W., Content, J. and Huez, G. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 6030–6034
29 Chardin, P., Courtois, G., Mattei, M.-G. and Gisselbrecht, S. (1991) Nucleic Acids Res. **19**, 1431–1436
30 Mount, S. M. (1982) Nucleic Acids Res. **10**, 459–472
31 Archambault, J., Milne, C. A., Schappert, K. T., Baum, B., Friesen, J. D. and Segall, J. (1992) J. Biol. Chem. **267**, 3282–3288
32 Fognani, C., Della Valle, G. and Babiss, L. E. (1993) EMBO J. **12**, 4985–4992
33 Gilbert, W. (1985) Science **228**, 823–824
34 North, M., Sargent, C., O'Brien, J., Taylor, K., Wolfe, J., Affara, N. A. and Ferguson-Smith, M. A. (1991) Nucleic Acids Res. **19**, 2579–2586

Received 13 March 1995/24 May 1995; accepted 30 May 1995