Identification of a Novel SCAN Box-related Protein That Interacts with MZF1B

THE LEUCINE-RICH SCAN BOX MEDIATES HETERO- AND HOMOPROTEIN ASSOCIATIONS*

Tara L. Sander, Amy L. Haas, Michael J. Peterson, and Jennifer F. Morris‡

From the Kelly Weil Laboratory of Pediatric Molecular Oncology, Medical College of Wisconsin, Departments of Pediatrics and Biochemistry, Milwaukee, Wisconsin 53226

The SCAN box or leucine-rich (LeR) domain is a conserved motif found within a subfamily of C_2H_2 zinc finger proteins. The function of a SCAN box is unknown, but it is predicted to form α-helices that may be involved in protein-protein interactions. Myeloid zinc finger gene-1B (*MZF1B*) is an alternatively spliced human cDNA isoform of the zinc finger transcription factor, *MZF1*. MZF1 and MZF1B contain 13 C_2H_2 zinc finger motifs, but only MZF1B contains an amino-terminal SCAN box. A bone marrow cDNA library was screened for proteins interacting with the MZF1B SCAN box domain and RAZ1 (SCAN-related protein associated with MZF1B) was identified. *RAZ1* is a novel cDNA that encodes a SCAN-related domain and arginine-rich region but no zinc finger motifs. Co-immunoprecipitation assays demonstrate that the SCAN box domain of MZF1B is necessary for association with RAZ1. By yeast two-hybrid analysis, the carboxyl terminus of RAZ1 is sufficient for interaction with the MZF1B SCAN box. Furthermore, MZF1B and RAZ1 each self-associate *in vitro* via a SCAN box-dependent mechanism. These data provide evidence that the SCAN box is a protein interaction domain that mediates both hetero- and homoprotein associations.

Zinc finger genes encode an abundant class of DNA- and RNA-binding proteins that represent an estimated 5% of the genes in the human genome. Many C_2H_2 zinc finger genes have been demonstrated to function as transcriptional regulators and frequently, zinc finger genes are targeted for disruption in a variety of human diseases and cancers. The Krüppel-like subclass of mammalian C_2H_2 zinc finger proteins, first identified in the zinc finger transcription factor TFIIIA, share a conserved link between the last histidine of the preceding finger motif with the first cysteine of the next finger (H-C link) (1). Krüppel-like proteins often contain conserved modular domains outside of their zinc finger motifs. These identified domains include the KRAB (<u>Kr</u>üppel-<u>a</u>ssociated <u>b</u>ox) domains A and B, FAX (<u>f</u>inger-<u>a</u>ssociated bo<u>x</u>) domain of *Xenopus*, BTB/POZ (<u>b</u>road complex, <u>t</u>ramtrack, and <u>b</u>ric-a-brac/<u>po</u>xvirus and <u>z</u>inc finger) or ZiN (<u>z</u>inc finger <u>N</u>-terminal) domain, and the SCAN box or leucine-rich domain.

To date, the functions of the KRAB and BTB/POZ domains have been best characterized. The KRAB domain is a conserved stretch of 75 amino acids found in an estimated one-third of Krüppel-like zinc finger proteins (2). The KRAB domain, further subdivided into domains A and B, functions as a potent transcriptional repressor (3–5) and is predicted to fold into two amphipathic helices (2). The KRAB domain from KOX1 interacts with human TIF1β (also named KAP-1, <u>K</u>RAB-<u>a</u>ssociated <u>p</u>rotein-1) (6, 7) and appears to exert its transcriptional repression activity through this interaction (6, 7). In addition, the KRAB-A domain of Kid-1 interacts with KRIP-1 (<u>KR</u>AB-A <u>i</u>nteracting <u>p</u>rotein), which is likely to be the murine homologue of TIF1β and KAP-1 (8). The POZ domain defines a conserved region of approximately 120 amino acids and is found in 5–10% of zinc finger proteins. The POZ domain is a protein interaction motif (9) that mediates both homo- (10, 11) and heterodimerization (12). Several POZ proteins are transcriptional repressors, including the oncoproteins PLZF (13) and BCL-6 (14), and the POZ domain has been shown to function as an autonomous transcriptional inhibitory domain (15). The POZ domain also has been demonstrated to interact with the co-repressors N-CoR, SMRT, Sin3, and histone deacetylase (16–19), suggesting that POZ-containing proteins mediate transcriptional repression by recruiting histone deacetylase through a co-repressor complex. However, this may not be a general mechanism for POZ-containing transcription factors (20).

The leucine-rich (LeR) domain is a conserved motif present in the amino terminus of a subfamily of C_2H_2 zinc finger proteins (4). This motif has also been designated the SCAN box, which was derived from the first four proteins found to contain this domain (<u>S</u>RE-ZBP, <u>C</u>T-fin-51, <u>A</u>W-1, <u>n</u>umber 18 cDNA) (21). To date, the SCAN box has been identified in approximately 20 zinc finger proteins from human, mouse, and rat including ZNF174 (21), RLZF-Y (22), FPM315 (23), ZNF213 (24), MZF1B[1] and its murine homologue, MZF-2 (25). SCAN box domains are about 80 residues in length, and approximately two-thirds of the amino acids are highly conserved with 80–100% sequence identity. The function of the SCAN box is unknown. Based on protein sequence analysis, the SCAN box is predicted to form two or three amphipathic helices that may be involved in protein-protein interactions (4, 21). However, no proteins have been identified that interact with the SCAN box. Fusion of the SCAN box to a GAL4 DNA binding domain has

[1] M. J. Peterson and J. F. Morris, submitted for publication.

The SCAN Box Mediates Protein Interactions

TABLE I

Oligonucleotides used to PCR-amplify MZF and RAZ1 cDNA for subcloning into expression vectors

Primer	Nucleotides[a]	RE site[b]	Strand	$5' \rightarrow 3'$
1[c]	562–577	*Eco*RI	Sense	CCGGAATTCGCCACCATGAGGCCTGCGGTGC
2	2749–2763	*Hin*dIII	Antisense	CTATATAAGCTTCTCGGCGCTGTGGAC
3	1318–1332	*Hin*dIII	Antisense	CCCAAGCTTTGGGGAGAAGATGCC
4[c]	1948–1962	*Bgl*II	Sense	TGAGATCTCCACCATGGGCGATCCCCCGGGC
5	2749–2766	*Hin*dIII	Antisense	AATTAAGCTTCTACTCGGCGCTGTGGAC
6	685–699	*Kpn*I	Antisense	CGGGGTACCCACGCAGGCGTGCAGC
7	931–945	*Kpn*I	Sense	CGGGGTACCCGGGAGCCGGGCGGA
8	1618–1632	*Hin*dIII	Antisense	CTATATAAGCTTACGGCCGCCCCTAAC
9[c]	1091–1104	*Bgl*II	Sense	GGAGATCTCCACCATGAATGGTCCCCT
10[c]	700–711	*Eco*RI	Sense	CCGGAATTCGCCACCATGTTCCGGTGCTTC
11	916–930	*Hin*dIII	Antisense	CCCAAGCTTCTAGCGCAGCCCATCTAC
12	700–712	*Eco*RI	Sense	CCGGAATTCGTTTCCGGTGCTTCC
13	916–930	*Xho*I	Antisense	CCGCTCGAGCTAGCGCAGCCCATCTAC
14	115–129	*Eco*RI	Sense	CCGGAATTCATGGCGGCTACGGAG
15	640–654	*Sal*I	Antisense	CGACGTCGACTCAGCCAGTGATGCG
16	115–129	*Eco*RI	Sense	CCGGAATTCTTATGGCGGCTACGGAG
17	640–654	*Xho*I	Antisense	CCGCTCGAGTCAGCCAGTGATGCG
18	382–396	*Xho*I	Antisense	CCGCTCGAGCTATGTGGAGCGCGCTTC
19	397–414	*Eco*RI	Sense	CCGGAATTCCAGGCCCCGCCGGCTCT

[a] Corresponding nucleotide position based on GenBank™-deposited sequence for MZF1B, MZF1, and RAZ1.

[b] Flanking restriction enzyme (RE) site, single underline.

[c] Engineered Kozak consensus sequence (37), double underline.

demonstrated that the SCAN box cannot confer transactivation or repression function onto a heterologous DNA binding domain (4, 21), suggesting that the SCAN box is not an independent transcriptional regulatory domain.

To gain a better understanding of the SCAN box function, we directed our attention to the SCAN box-containing zinc finger protein, MZF1B.[1,2] MZF1B is an alternatively spliced isoform of the zinc finger transcription factor, MZF1 (26) and the human homologue of murine MZF-2 (25). MZF1 is a 485-amino acid protein that contains 13 C_2H_2 zinc finger motifs arranged in a bipartite DNA binding domain. The consensus DNA binding sites have been identified (27), and MZF may regulate the expression of specific genes in a tissue-specific manner (28, 29). MZF1 expression is both necessary for hematopoietic cell differentiation (30) and critical to the regulation of cell proliferation and apoptosis (31–33). *MZF1B* cDNA encodes a 734-amino acid protein that shares identity to the carboxyl terminus of MZF1, including the 13 C_2H_2 zinc finger motifs. However, *MZF1B* encodes an additional 257 residues at its amino terminus, which contains a SCAN box domain. Therefore, the amino-terminal domains unique to each isoform may define the distinct functions of MZF1 and MZF1B proteins. We hypothesized that the MZF1B SCAN box is a protein interaction domain. A human bone marrow cDNA library was screened for proteins interacting with the MZF1B SCAN box domain, and RAZ1 (SCAN-related protein associated with MZF1B) was identified.

EXPERIMENTAL PROCEDURES

pcDNA Expression Plasmids—MZF and RAZ1 cDNAs were subcloned into the mammalian expression vector, pcDNA3.1 (−)/Myc-His (Invitrogen; Carlsbad, CA) using oligonucleotides synthesized by Life Technologies, Inc. (Table I). Plasmid sequences were confirmed by automated sequencing using AmpliTaq DNA Polymerase FS (Applied Biosystems Inc., Foster City, CA).

Full-length *MZF1B* cDNA (amino acids 1–734) and the unique amino terminus of MZF1B (amino acids 1–257) were amplified by polymerase chain reaction (PCR)[3] methods with oligonucleotides 1 and 2 or oligonucleotides 1 and 3, respectively, and subcloned into the *Eco*RI and

[3] The abbreviations used are: PCR, polymerase chain reaction; bp, base pair; EST, expressed sequence tag; AD, activation domain; BD, binding domain; HA, hemagglutinin epitope tag; IVT, *in vitro* transcription and translation; mh, Myc/His epitope tag; RACE, rapid amplification of cDNA ends; HIV, human immunodeficiency virus.

*Hin*dIII sites of pcDNA3.1 (A) to produce plasmids MZF1B-mh/pcDNA and MZF1B N-term-mh/pcDNA, where "mh" indicates fusion of the protein to the carboxyl-terminal Myc-His epitope tag.

MZF1B lacking the Myc/His tag was created by subcloning the same PCR fragment from MZF1B-mh/pcDNA into pcDNA3.1 (C) to create MZF1B₃*/pcDNA, which added three extra amino acids (Lys, Ala, Thr) to the carboxyl terminus of MZF1B. To eliminate these extra amino acids, a *Sma*I fragment from gp-end*/pcDNA (MZF1B amino acids 463–734) was subcloned into the *Sma*I-digested MZF1B₃*/pcDNA plasmid to create MZF1B*/pcDNA. The plasmid gp-end*/pcDNA was produced by PCR amplification of MZF1B cDNA (amino acids 463–734, including the stop codon) with oligonucleotides 4 and 5 and subcloning the *Bgl*II and *Hin*dIII PCR fragment into pcDNA3.1 (A).

The unique amino terminus of MZF1B (amino acids 1–257) lacking the Myc/His tag was constructed by subcloning the same PCR fragment of MZF1B N-term-mh/pcDNA into pcDNA3.1 (C) to create MZF1B N-term*/pcDNA, which encodes an extra three amino acids (Lys, Ala, Thr) at the carboxyl-terminal end.

Full-length MZF1B with the SCAN box domain deleted, MZF1B ΔSCAN-mh/pcDNA, was constructed by a three-step subcloning procedure. First, the cDNA encoding amino acids 1–46 of MZF1B was PCR-amplified using oligonucleotides 1 and 6 and subcloned into the *Eco*RI and *Kpn*I sites of pcDNA3.1 (A), creating MZF1B 1–46-mh/pcDNA. Second, the *MZF1B* cDNA encoding amino acids 124–357 (contains an internal *Bam*HI site) was amplified by PCR using oligonucleotides 7 and 8 and subcloned into the *Kpn*I and *Hin*dIII sites of MZF1B 1–46 mh/pcDNA to create MZF1B 1–357ΔSCAN-mh/pcDNA. Third, a *Bam*HI and *Hin*dIII fragment of MZF1B-mh/pcDNA (internal *Bam*HI and 3′ *Hin*dIII site immediately following amino acid 734) was then subcloned into the *Bam*HI and *Hin*dIII sites of MZF1B 1–357ΔSCAN-mh/pcDNA to construct MZF1B ΔSCAN-mh/pcDNA.

The unique MZF1B amino terminus with the SCAN box deleted, MZF1B N-term ΔSCAN-mh/pcDNA, was constructed by subcloning a *Bgl*II fragment from MZF1B 1–357ΔSCAN-mh/pcDNA into the *Bgl*II-digested MZF1B N-term-mh/pcDNA plasmid.

Full-length *MZF1* cDNA (amino acids 1–485) was PCR-amplified with oligonucleotides 9 and 2. The PCR fragment, flanked by *Bgl*II and *Hin*dIII, was subcloned into the *Bam*HI and *Hin*dIII sites of pcDNA3.1 (A) to create MZF1-mh/pcDNA.

The cDNA isolated from yeast two-hybrid screening was removed from the pACT2 expression vector by digestion with *Bgl*II (including the hemagglutinin (HA) tag) and subcloned into the *Bam*HI site of pcDNA3.1 (A) to generate ha-RAZ1/pcDNA.

β-gal-mh/pcDNA, an expression plasmid encoding Myc-His epitope-tagged β-galactosidase, was purchased from Invitrogen (Carlsbad, CA).

Yeast Two-hybrid Expression Plasmids—MZF1B and RAZ1 cDNA were subcloned into yeast expression cloning vectors, pAS2–1 and pACT2 (CLONTECH; Palo Alto, CA), to generate fusion proteins with the GAL4 DNA binding domain (BD; amino acids 1–147) or GAL4

A

```
1   GGACGCAGAGAAGCCAGAGACTTTCGCTTCCGGCTGCCGCAGGCGCTTCGCTGGTGCAGA   60
    G   R   R   E   A   R   D   F   R   F   R   L   P   Q   A   L   R   W   C   R   20

61  CGCAGTGCTGAGCACACAGCTACCGGACAAAGAGTGACGCCCGGAGCTGGAGTTATGGCG   120
    R   S   A   E   H   T   A   T   G   Q   R   V   T   P   G   A   G   V   M   A   40

121 GCTACGGAGCCGATCTTGGCGGcCACTGGGAGTCCCGCGGCGGTGCCACCGGAGAAACTG   180
    A   T   E   P   I   L   A   A   T  [G]  S   P   A   A   V   P   P   E   K   L   60

181 GAAGGAGCCGGTTCGAGCTcAGCCCCTGAGCGTAACTGTGTGGGCTCCTcGCTGCCaGAG   240
    E  [G]  A   G   S   S  (S)  A   P   E   R   N   C   V   G   S  (S)  L   P   E   80

241 GCCTCACCGCCTGCCCCTGAGCCTTCCAGTCCCAACGCCGCGGTCCCTGAAGCCATCCCT   300
    A   S   P   P   A   P   E   P   S   S   P   N   A   A   V   P   E   A   I   P   100

301 ACGCCCCGAGCTGCGGCCTCCGCGGCCCTGGAGCTGCCTCTCGGGCCCGCACCCGTGAGC   360
   (T)  P   R   A   A   A   S   A   A   L   E   L   P   L   G   P   A   P   V   S   120

361 GTAGCGCCTCAGGCCGAAGCTGAAGCGCGCTCCACACCAGGCCCCGcCGGCTCTAGACTC   420
    V   A   P   Q   A   E   A   E   A   R   S   T   P   G   P   A   G   S   R   L   140

421 GGTCCCGAGACGTTcCGcCAGCGTTTcCGGCAGTTCCGCTACCAGGATGCGGCGGGTCCC   480
    G   P   E  (T)  F   R   Q   R   F   R   Q   F   R   Y   Q   D   A   A   G   P   160

481 CGGGAGGCTTTcCGGCAGCTGCGGGAGCTGTcCCGcCAGTGGCTGCGGCCTGAcATCCGC   540
    R   E   A   F   R   Q   L   R   E   L   S   R   Q   W   L   R   P   D   I   R   180

541 ACCAAGGAGCAGATCGTGGAGATGCTGGTGCAAGAGCAGCTGCTCGCCATCCTGCCCGAG   600
    T   K   E   Q   I   V   E   M   L   V   Q   E   Q   L   L   A   I   L   P   E   200

601 GCGGCTCGGGCCCGGCGGATCCGCCGCCGCACGGATGTGCGCATCACTGGCTGAGCGGTG   660
    A   A   R   A   R   R   I   R   R   R  (T)  D   V   R   I   T   G   *   217

661 GAGCTGCGGGCGGCCAGGGCCGGGCGCTCTGTGCGGACTGGGGCCATGATCGGGCCCGGG   720

721 GGCCTGAGCCTGGGACCCCACCCCGTGTTAATGAAAAATGAGTTTTGGCAGCGCCAAAAA   780

781 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA   840

841 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA   880
```

FIG. 1. **RAZ1 cDNA and amino acid sequence.** *A*, the isolated *RAZ1* cDNA sequence and open reading frame encoding 217 amino acids are shown. A methionine at nucleotide position 115 (amino acid 38) contains a putative translation start site based on Kozak's consensus sequence for translation initiation, (GCC)GCC(A/G)CCAUGG (37). A *solid underline* indicates the SCAN-related domain at amino acids 140–200. The arginine-rich region is at amino acids 201–217. Several putative phosphorylation (*open circles*) and myristoylation sites (*open squares*) are present within the sequence. A *schematic diagram* of RAZ1 is illustrated *below* the sequence (not drawn to scale). The cDNA sequence has been deposited into the GenBank™ data base (accession no. AF207829). *B*, *RAZ1* maps to chromosome 20. *RAZ1*, the cDNA sequence (1–775 plus poly(A)) isolated from yeast two hybrid screening. *Chrom 20*, a GenBank™-deposited human chromosome 20 sequence (accession no. AL109965) at 20q11.1–11.23, is identical to *RAZ1* cDNA at nucleotides 1–59 and 60–775 and contains an additional 108 bp between nucleotides 59 and 60 of *RAZ1*. A *dotted arrow* indicates an open reading frame (*ORF*). An *asterisk* indicates a stop codon. *AD*, activation domain.

⊞ SCAN-related domain   ■ Arginine-rich region

B
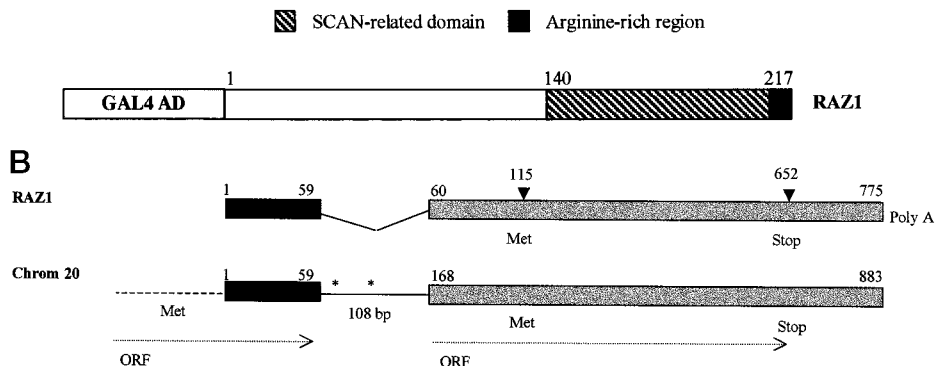
activation domain (AD; amino acids 768–881), respectively. The constructs were confirmed by automated sequencing using AmpliTaq DNA Polymerase FS (Applied Biosystems Inc., Foster City, CA).

To generate the bait plasmid, SCAN/pAS2–1, the cDNA encoding the MZF1B SCAN box (amino acids 47–123) was amplified by PCR using oligonucleotides 10 and 11 and subcloned into the *Eco*RI and *Hin*dIII sites of the mammalian expression vector CB6+ (34) to create SCAN/CB6+. SCAN/CB6+ was digested with *Eco*RI and *Bam*HI and the MZF1B SCAN box insert was subcloned into pAS2–1.

To generate fusions of the MZF1B SCAN box with the GAL4 AD and HA tag, the cDNA encoding the MZF1B SCAN box (amino acids 47–123) was amplified by PCR using oligonucleotides 12 and 13 and subcloned

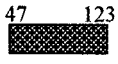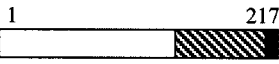into the *Eco*RI and *Xho*I sites of pACT2 to create SCAN/pACT2.
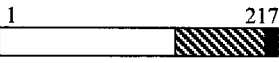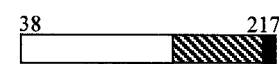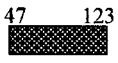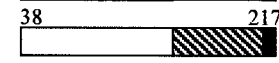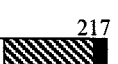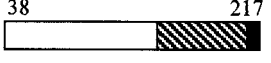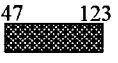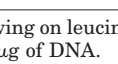
RAZ1/pACT2 is the library clone isolated from the yeast two-hybrid screen as described below. This clone contains the *RAZ1* cDNA (amino acids 1–217) library insert subcloned into the *Eco*RI and *Xho*I sites of pACT2 to generate a fusion protein containing the GAL4 AD and HA tag.

RAZ1 cDNA encoding amino acids 38–217 was amplified by PCR using oligonucleotides 14 and 15 or oligonucleotides 16 and 17 and subcloned into the *Eco*RI and *Sal*I sites of pAS2–1 or the *Eco*RI and *Xho*I sites of pACT2 to create RAZ1 38–217/pAS2–1 and ha-RAZ1 38–217/pACT2, respectively.

cDNA encoding either the amino terminus (amino acids 38–132) or

TABLE II
Identification of RAZ1, a novel cDNA library clone that interacts with the MZF1B SCAN box

Yeast Transformation		Transformation Efficiency[a]	% Positive Interaction[b]
GAL4 DNA BD	GAL4 AD		
MZF1B (47–123)	Bone Marrow cDNA Library	8.0×10^3	2.5×10^{-4}
MZF1B (47–123)		2.5×10^2	0
	RAZ1 (1–217)	4.3×10^2	0
MZF1B (47–123)	RAZ1 (1–217)	1.3×10^2	100
	RAZ1 (38–217)	8.0×10^2	0
MZF1B (47–123)	RAZ1 (38–217)	5.6×10^2	100
RAZ1 (38–217)		5.0×10^2	0
	MZF1B (47–123)	6.3×10^2	0
RAZ1 (38–217)	MZF1B (47–123)	5.0×10^2	100

[a] Number of yeast colony-forming units/μg of plasmid DNA (cfu/μg) growing on leucine and tryptophan-depleted medium: cfu $\times$ total suspension vol (μl)/vol plated (μl) $\times$ dilution factor $\times$ amount DNA used (μg) = cfu/μg of DNA.

[b] The percentage of yeast positive for both *HIS3* and *lacZ* reporter gene expression.

TABLE III
The RAZ1 SCAN domain lacks the predicted third α-helix conserved in other SCAN box proteins

SCAN box	NH$_2$–	Helix 1	Helix 2	Helix 3	–COOH
consensus:[a]		...E..R..FR.F.Y....GP.EAL..LRELC..WLRPE..TKEQILELLVLEQFL.ILP.E.Q..V....PES.EE.V...E.L			
RAZ1:[b]		LGP*TF*QR**Q*R*QDAA**R**FRO****SRO****DIR*****V*M**O**L*A***EAARARRIRRRTDVRITG			

Protein	Species[c]	Conserved motifs[d]	Accession	Reference
RAZ1	Human 20	SCAN-related	AF207829	
PGC-2/Leap1	Mouse	SCAN-related	AF106473	38
p18 (#18)[e]	Human 3	SCAN, KRAB A	Z21707	4
3c3/TRFA/p20	Human 3	SCAN, KRAB A+B	L32162	39
MZF1B	Human 19	SCAN, 13 ZF	AF055078	
ZNF174	Human 16	SCAN, 3 ZF	U31248	21
SRE-ZBP	Human 6	SCAN, 7 ZF	Z11773	40
ZNF165/Zpf165/LD65	Human 6	SCAN, 5 ZF	X84801/U78722	41, 42
ZNF192/LD5–1	Human 6	SCAN, 9 ZF	U57796	41
ZNF193 (PRD51)	Human 6	SCAN, 5 ZF	U62392	41
ZNF191/ZNF24	Human	SCAN, 4 ZF	AF016052/U68536	
KIAA0426	Human	SCAN, 11 ZF	AB007886	
KIAA0557	Human	SCAN, 5 ZF	AB011129	43
ZNF202 beta	Human 11	SCAN, 8 ZF	AF027219	44
MZF-2	Mouse	SCAN, 13 ZF	AB007407	25
Zpf-38/CTfin51/RU49	Mouse 5	SCAN, 7 ZF	X63747/D10630/U41671	45–47
Zpf96	Mouse	SCAN, 7 ZF	U62908	
Zpf-29	Mouse	SCAN, 14 ZF	X55126	48
FPM315	Human 16	SCAN, KRAB A, 9 ZF	D88827	23
ZNF213/CR53	Human 16	SCAN, KRAB A, 3 ZF	AF091512/AF017433	24
ZnF20	Human 3	SCAN, KRAB A, 20 ZF	AF011573	49
RLZF-Y1	Rat	SCAN, KRAB A, 7 ZF	AF052042	22
Zfp95	Mouse	SCAN, KRAB B, 12 ZF	U62907	

[a] The SCAN box domains of the SCAN-containing proteins listed in the table were aligned. The SCAN consensus is defined by amino acids identical in $\geq$15/22 aligned proteins. Amino acids in boldface type indicate 100% conservation. The underline indicates the predicted α-helices 1, 2, and 3.

[b] RAZ1 COOH terminus (amino acids 140–217); asterisks indicate conserved residues of the SCAN box consensus sequence.

[c] Number indicates chromosome locus.

[d] ZF, C$_2$H$_2$ zinc finger.

[e] p18 is contained within ZnF20.

carboxyl terminus (amino acids 133–217) of RAZ1 was PCR-amplified using oligonucleotides 16 and 18 or oligonucleotides 19 and 17 and subcloned into the *Eco*RI and *Xho*I sites of pACT2 to generate ha-RAZ1 N-term/pACT2 and ha-RAZ1 C-term/pACT2, respectively.

GAL4 Yeast Two-hybrid Analysis—The MATCHMAKER Two-Hy-brid System 2 and GAL4 Human Bone Marrow MATCHMAKER cDNA library were purchased from CLONTECH. A large scale, sequential transformation of the GAL4-SCAN bait fusion (SCAN/pAS2–1) and bone marrow cDNA library was carried out according to the manufacturer's directions. Briefly, the bone marrow cDNA library (50 μg) was

introduced into CG-1945 yeast previously transformed with SCAN/pAS2–1, and yeast transformants were plated onto synthetic drop-out medium depleted of histidine, tryptophan, and leucine (SD −His/−Trp/−Leu) containing 5 mM 3-amino-1,2,3-triazole. Yeast clones positive for HIS3 expression were assayed for β-galactosidase activity with the colony-lift filter assay. Yeast plasmid DNA was isolated using glass beads and phenol/chloroform extraction (Method 1 in Ref. 63) and pACT2 library plasmids were rescued via transformation of electrocompetent *Escherichia coli* KC8 (CLONTECH).

Phage Library Screening—Approximately 5×10^5 plaque-forming units from a λgt11 5′-StretchPlus cDNA library synthesized from the human erythroleukemia cell line, K562 (CLONTECH) were plated and transferred to BA-S nitrocellulose filters (Schleicher & Schuell). The filters were hybridized (35) with a ^{32}P-labeled *RAZ1* cDNA probe (nucleotides 115–396) for 20 h at 65 °C and washed under high-stringency conditions with 0.2× SSC, 0.1% SDS at 60 °C.

Northern Blot Analysis—Total RNA was isolated from tissue culture cells with TRIzol Reagent (Life Technologies). The RNA was separated by electrophoresis through a 1% agarose-formaldehyde gel and transferred to a 0.2-μm Nytran membrane (Schleicher & Schuell). The membrane was hybridized (35) with a ^{32}P-labeled *RAZ1* cDNA probe (nucleotides 115–654) for 1 h at 68 °C using QuickHyb hybridization solution (Stratagene, La Jolla, CA) and washed under high-stringency conditions with 0.1× SSC, 0.1% SDS at 60 °C.

Rapid Amplification of cDNA Ends (RACE)—First and second strand cDNA synthesis from human bone marrow poly(A)$^+$ RNA (CLONTECH), adapter ligation, and 5′ and 3′ Marathon RACE reactions were performed using the Marathon cDNA Amplification Kit (CLONTECH) according to the manufacturer's directions. Gene-specific primers were synthesized by Life Technologies for 3′ RACE (RAZ1 GSP-2 (nucleotides 361–385), 5′-GTA GCG CCT CAG GCC GAA GCT GAA G-3′) and 5′ RACE (RAZ1 GSP-1b (nucleotides 723–749), 5′-AAC ACG GGG TGG GGT CCC AGG CTC AGG-3′).

In Vitro Transcription and Translation (IVT)—The TNT Coupled Reticulocyte Lysate System (Promega, Madison, WI) was used to synthesize radiolabeled proteins for co-immunoprecipitation experiments. *In vitro* transcription and translation reactions were carried out in a final volume of 25 μl according to the manufacturer's directions using 1 μg of plasmid DNA and 10–20 μCi of [^{35}S]cysteine or methionine (Amersham Pharmacia Biotech).

Immunoprecipitation—IVT lysate (5–7 μl) and 1 μg of purified rabbit IgG (α-His, α-Myc, α-post-SCAN, α-ZF) or a 1:150 dilution of rabbit α-HA antisera were incubated at 4 °C for 2–3 h in 500 μl of immunoprecipitation buffer (150 mM NaCl, 10 mM NaPO$_4$ pH 7.2, 1% IGEPAL, 0.1 mM phenylmethylsulfonyl fluoride, 2 μg/ml leupeptin, and 2 μg/ml aprotinin). The immune complexes were immunoprecipitated with 10% protein A-Sepharose (Amersham Pharmacia Biotech) for 30 min at 4 °C, washed three times with 1 ml of immunoprecipitation buffer, and resuspended in 15 μl of 1× SDS-Laemmli buffer. The gels were fluorographed with Me$_2$SO and 2,5-diphenyloxazole and visualized by autoradiography.

Antibodies—Normal rabbit IgG was used as a negative control (Santa Cruz Biotechnology, Inc., Santa Cruz, CA). Rabbit polyclonal IgG

α-His and α-Myc are specific for the 6× histidine probe and c-Myc epitope tag, respectively (Santa Cruz Biotechnology). Rabbit polyclonal crude antiserum α-HA was raised against the HA tag (Babco, Richmond, CA). Rabbit polyclonal α-post-SCAN antiserum was generated using amino acids 124–257 of MZF1B. Rabbit polyclonal α-ZF antiserum was generated using zinc fingers 1–4 of MZF1B (amino acids 358–462) (32). IgG fractions were purified from both α-post-SCAN and α-ZF antisera by protein A chromatography (36).

Data Base Searching—Computer searches were done using the FASTA, BLAST, and MOTIFS algorithms through the Wisconsin Package software (64) or BLAST version 2.0 through the World Wide Web interface. Nucleotide sequences were compared with entries in the GenBank™ or expressed sequence tag (EST) data bases, while peptide sequences were searched against the Protein Information Resources (PIR) or Swiss-Prot data base.

Chromosome 20 Sequence—The *Homo sapiens* clone RP5–1121G12 from the RPCI5 library maps to chromosome 20q11.1–11.23 and has been assigned the EMBL/GenBank™ accession number AL109965. These data were produced by the Human Chromosome 20 Mapping and Sequencing Groups at the Sanger Center. Mapping and sequence data can be obtained on the World Wide Web.

RESULTS

The MZF1B SCAN Box Interacts with RAZ1—A human bone marrow cDNA library was screened by yeast two-hybrid analysis for potential MZF1B SCAN box interacting proteins, and one clone was identified (Table II). To confirm the positive protein interaction, we performed two-hybrid assays in the presence and absence of the MZF1B SCAN box domain. The library plasmid did not autonomously activate reporter gene expression, and a positive protein interaction was only ob-
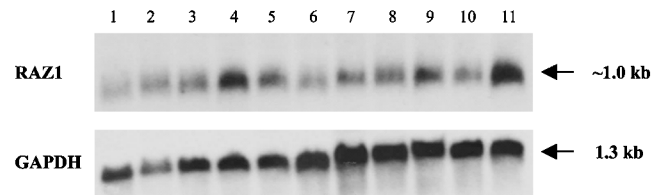
FIG. 2. **Northern blot analysis identifies a *RAZ1* transcript of ~1 kilobase in both hematopoietic and nonhematopoietic human cell lines.** *Lane 1*, HL60, human promyelocyte induced to myeloid differentiation with all-*trans*-retinoic acid; *lane 2*, KG-1, human myeloblast; *lane 3*, K562, human erythroleukemia, myeloblast-like; *lane 4*, HEL, human erythroleukemia; *lane 5*, U937, human lymphoma, monocyte-like; *lane 6*, Namawala, human Epstein-Barr virus-positive Burkitt's lymphoma; *lane 7*, DG75, human Epstein-Barr virus-negative Burkitt's lymphoma; *lane 8*, Jurkat E6, human T-lymphoblast; *lane 9*, 293T, human kidney fibroblast; *lane 10*, HeLa, human cervical adenocarcinoma; *lane 11*, Caco-2, human colon adenocarcinoma.

FIG. 3. **MZF and RAZ1 pcDNA plasmids.** The epitope-tagged and nontagged MZF and RAZ1 proteins are shown schematically (not drawn to scale). The mh epitope is fused in frame to the carboxyl terminus of MZF. The HA tag is fused to the amino terminus of RAZ1. The MZF1B SCAN box (amino acids 47–123), acidic domain (MZF1, amino acids 60–72; MZF1B, amino acids 309–321), and RAZ1 SCAN-related box (amino acids 140–200) are shown.
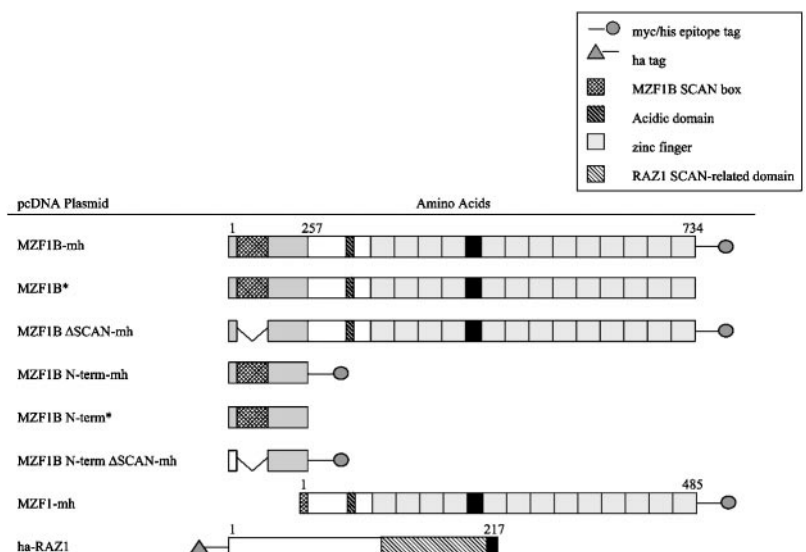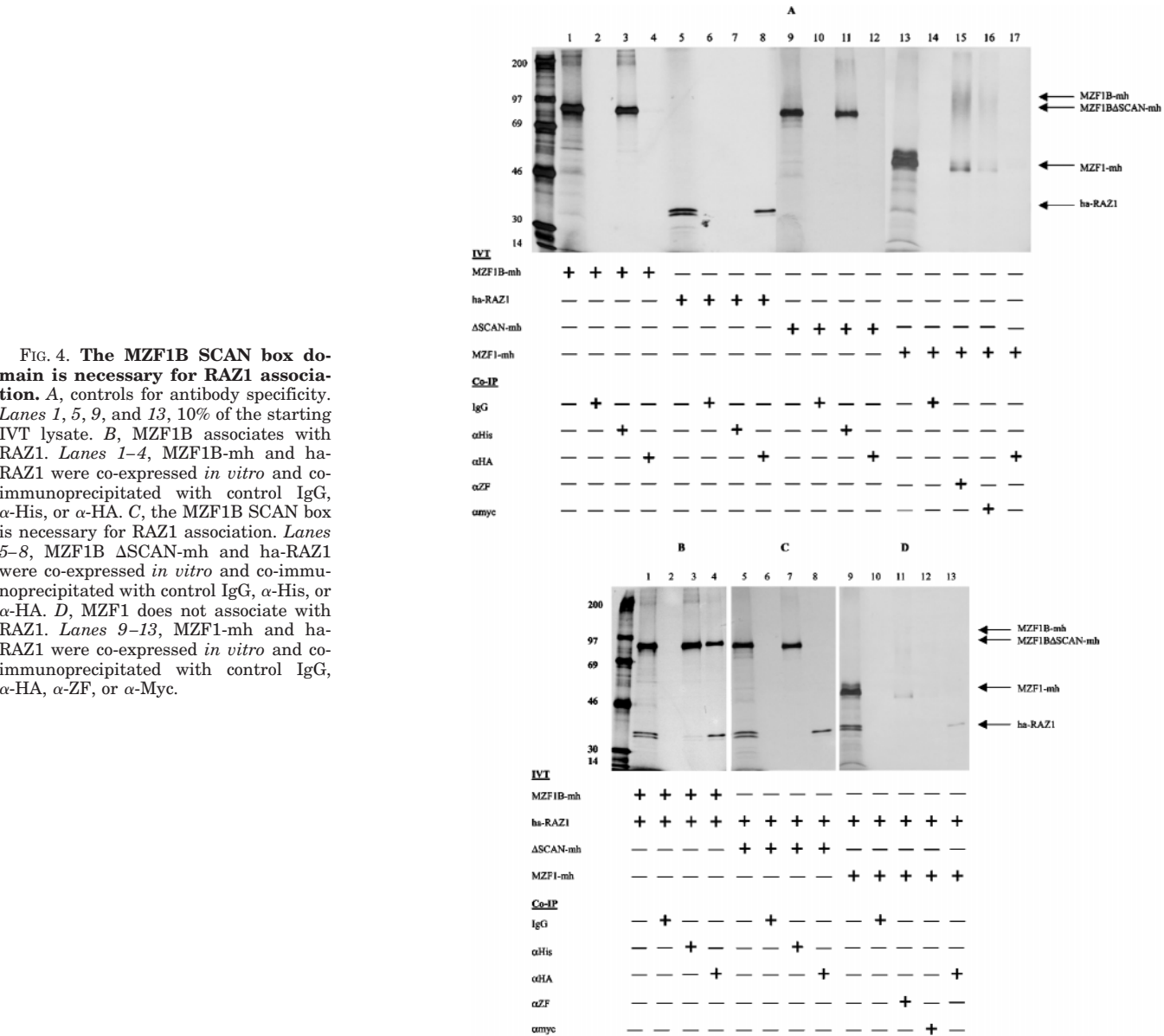
FIG. 4. **The MZF1B SCAN box domain is necessary for RAZ1 association.** *A*, controls for antibody specificity. *Lanes 1*, *5*, *9*, and *13*, 10% of the starting IVT lysate. *B*, MZF1B associates with RAZ1. *Lanes 1–4*, MZF1B-mh and ha-RAZ1 were co-expressed *in vitro* and co-immunoprecipitated with control IgG, α-His, or α-HA. *C*, the MZF1B SCAN box is necessary for RAZ1 association. *Lanes 5–8*, MZF1B ΔSCAN-mh and ha-RAZ1 were co-expressed *in vitro* and co-immunoprecipitated with control IgG, α-His, or α-HA. *D*, MZF1 does not associate with RAZ1. *Lanes 9–13*, MZF1-mh and ha-RAZ1 were co-expressed *in vitro* and co-immunoprecipitated with control IgG, α-HA, α-ZF, or α-Myc.

served when both the MZF1B SCAN box and library plasmid were co-transformed into yeast (Table II). To determine whether the interaction was an artifact of the fusion protein partner, we switched the AD and the DNA BD fusion partners for both the MZF1B SCAN box and interacting clone. The interaction between the MZF1B SCAN box and the isolated library clone was not dependent upon the fusion protein partner (Table II). These control experiments confirm that we identified a cDNA library insert positive for MZF1B SCAN box protein interaction. We have named this protein, RAZ1, a SCAN-related protein associated with MZF1B.

RAZ1 Is a Novel SCAN Box-Related Protein—The cDNA and amino acid sequences of RAZ1 are shown in Fig. 1A. The sequence is not found in the GenBank[TM] data base and appears to be a novel clone. The open reading frame for *RAZ1* is defined by fusion with the upstream GAL4 activation domain and encodes 217 amino acids, starting with a glycine and ending with a stop codon at nucleotide position 652. The first methionine is at nucleotide 115 and contains a weak Kozak consensus sequence for translation initiation (37). Thus, it is probable that we have isolated a partial cDNA clone from the GAL4 fusion library that is incomplete at the 5′-end. The predicted sequence of *RAZ1* encodes a SCAN-related domain at its car-

boxyl terminus (amino acids 140–200) but no zinc finger motifs. We refer to the domain as "SCAN-related" because the alignment with SCAN domains in other zinc finger proteins is conserved at the amino terminus and truncated at the carboxyl terminus (Table III). Approximately 20 SCAN box-containing proteins have been reported and/or deposited into the Gen-Bank[TM] data base that contain zinc finger motifs and/or KRAB domains (Table III). Thus, the SCAN box appears to be frequently associated with zinc finger motifs and sometimes with KRAB domains. It is possible that *RAZ1* is a member of a novel gene family of non-zinc finger SCAN proteins. Immediately following the SCAN-related domain is an arginine-rich region (amino acids 201–217). The *RAZ1* open reading frame also contains putative sites for post-translational modification: two casein kinase II phosphorylation sites at amino acid positions Ser^{67} and Ser^{77}, two protein kinase C phosphorylation sites at positions Thr^{101} and Thr^{144}, one cAMP- and cGMP-dependent protein kinase phosphorylation site at position Thr^{211}, and two N-myristoylation sites at positions Gly^{50} and Gly^{62} (Fig. 1A).

RAZ1 Maps to Chromosome 20—During the course of our studies, we identified a GenBank[TM]-deposited human chromosome 20 sequence at 20q11.1–11.23 with identity to *RAZ1* at nucleotides 1–59 and 60–775. This sequence also contains an

TABLE IV
The carboxyl terminus of RAZ1 is sufficient for MZF1B SCAN box interaction and RAZ1 self-association

Yeast Transformation		Transformation Efficiency[a]	% Positive Interaction[b]
GAL4 DNA BD	GAL4 AD		
	RAZ1 N-term (38–132)	5.3×10^2	0
	RAZ1 C-term (133–217)	4.0×10^2	0
MZF1B (47–123)	RAZ1 N-term (38–132)	6.9×10^2	0
MZF1B (47–123)	RAZ1 C-term (133–217)	4.2×10^2	100
RAZ1 (38–217)	RAZ1 (38–217)	3.9×10^2	100
RAZ1 (38–217)	RAZ1 N-term (38–132)	6.3×10^2	0
RAZ1 (38–217)	RAZ1 C-term (133–217)	6.9×10^2	100

[a] Number of yeast colony-forming units/μg of plasmid DNA (cfu/μg) growing on leucine and tryptophan-depleted medium: cfu $\times$ total suspension vol (μl)/vol plated (μl) $\times$ dilution factor $\times$ amount DNA used (μg) = cfu/μg of DNA.

[b] The percentage of yeast positive for both *HIS3* and *lacZ* reporter gene expression.

additional 108-bp insert between nucleotides 59 and 60 of *RAZ1*. An illustration of the chromosome clone is shown in Fig. 1*B*. To further examine the *RAZ1* gene, 5×10^5 clones from a K562 cDNA library were screened with a *RAZ1* cDNA probe. Nine unique clones were isolated, and all are identical to *RAZ1*, extending from nucleotide 21 and through the poly(A) tail (data not shown). In addition, 3′-RACE analysis identified products identical to the 3′-end of *RAZ1* that correspond to the contiguous genomic DNA sequence on chromosome 20 including the stop codon, polyadenylation signal, and poly(A) tail. We obtained six 5′-RACE products identical to *RAZ1* at the 5′-end, of which one extends from nucleotide 17. Interestingly, one 5′-RACE product contains the 108-bp insert between nucleotides 59 and 60 of *RAZ1* (data not shown). These data provide independent confirmation that two *RAZ1* transcripts may exist with divergence at the 5′-end, one that does not contain an additional 108 bp of sequence and one that does.

RAZ1 mRNA Is Expressed in Various Cell Lines—Northern blot analysis detects a *RAZ1* transcript of ~1 kilobase in total RNA isolated from both human hematopoietic and nonhematopoietic cell lines. The highest levels of RNA expression were detected in the cell lines HEL (erythroleukemia) and Caco-2 (colon adenocarcinoma) (Fig. 2). The blot was reprobed with glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) to verify equal loading of RNA (Fig. 2). Comparison of *RAZ1* with the EST data base identified 100 ESTs between 300 and 600 bp in length that are identical to *RAZ1*. The ESTs were isolated from various human tissues including brain, breast, fetal heart, kidney, melanocyte, ovarian tumor, and placenta (data not shown). Six of these EST sequences contain the additional 108-bp insert between nucleotides 59 and 60 of *RAZ1*. Northern blots and reported ESTs suggest that RAZ1 may be widely expressed.
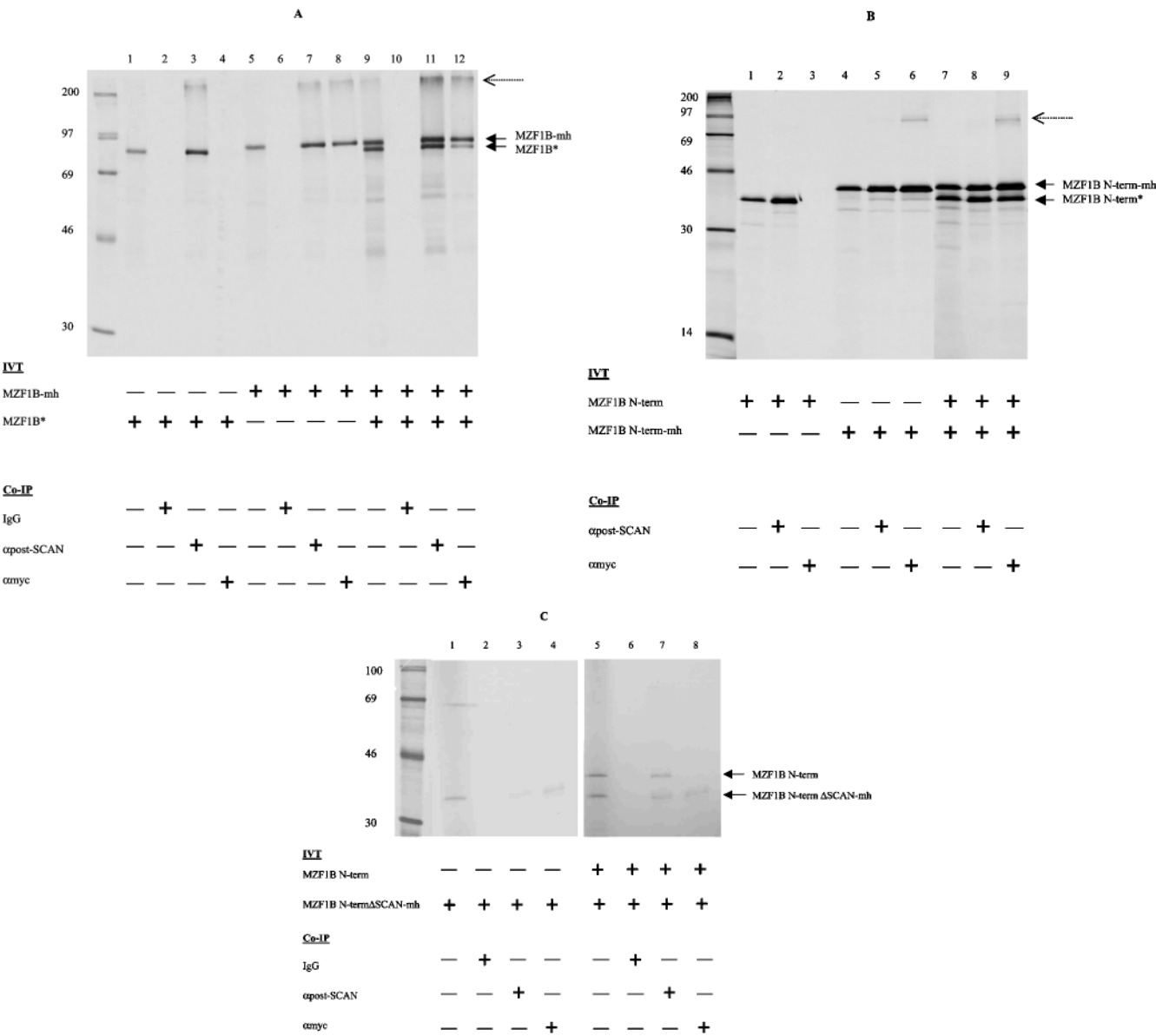
The MZF1B SCAN Box Domain Is Necessary for Interaction with RAZ1—To confirm that full-length MZF1B associates with RAZ1 and to identify MZF1B domains necessary and sufficient for interaction, MZF and RAZ1 proteins were co-expressed *in vitro* and co-immunoprecipitated with immunospecific antibodies.

As a first step, we demonstrated the immunospecificity of the nonspecific IgG, α-His, α-HA, α-ZF, and α-Myc antibodies by immunoprecipitating *in vitro* expressed MZF and RAZ1 proteins that contain either the amino-terminal HA or carboxyl-terminal mh epitope tag. The pcDNA expression plasmids used for immunoprecipitations are shown in Fig. 3. The antibodies were immunospecific, and no cross-reactivity was observed (Fig. 4*A*). Full-length MZF1B migrates as an 80-kDa protein, while the MZF1B NH$_2$ terminus migrates at approximately 42 kDa (Fig. 5). MZF1B ΔSCAN and MZF1 migrate at approximately 72 and 50 kDa, respectively. RAZ1 migrates as a doublet of approximately 35 kDa. We consistently observe that the upper band of RAZ1 is more efficiently immunoprecipitated with α-HA. Therefore, the upper band of RAZ1 might be a result of translation initiation at the methionine upstream of the HA epitope tag, while the lower migrating band may be due to internal translation initiation at methionine 38 of RAZ1, thus producing a protein that lacks the HA epitope tag.

Protein association was demonstrated by co-expressing both MZF1B and RAZ1 epitope-tagged proteins *in vitro* and co-immunoprecipitating with the immunospecific antibodies. MZF1B is detected when the lysate is immunoprecipitated with α-HA, and RAZ1 is detected with α-His (Fig. 4*B*). This suggests that full-length MZF1B and RAZ1 are being pulled down in the same immunocomplex and are interacting *in vitro*. In addition, neither RAZ1 nor MZF1B ΔSCAN proteins were detected in the same immunocomplex, suggesting that the MZF1B SCAN box is necessary for heteroassociation with RAZ1 (Fig. 4*C*). Furthermore, MZF1, which lacks the SCAN box domain, does not interact with RAZ1, verifying that the association with RAZ1 is unique to the SCAN box-containing amino-terminal region of

FIG. 5. **The MZF1B SCAN box is necessary for MZF1B self-association.** *A*, full-length MZF1B self-associates. *Lanes 1–8*, controls for antibody specificity; *lanes 9–12*, co-expression and immunoprecipitation of tagged and nontagged MZF1B. The IVT reactions for *A* and *B* were scaled down to a final volume of 10 μl, and 4.5 μl of the lysate was used for immunoprecipitations in a final volume of 250 μl. *B*, the MZF1B NH$_2$ terminus is sufficient for self-association. *Lanes 1–6*, controls for antibody specificity; *lanes 7–9*, co-expression and immunoprecipitation of tagged and nontagged MZF1B NH$_2$ terminus. *C*, the MZF1B SCAN box is necessary for self-association. *Lanes 1–4*, controls for antibody specificity; *lanes 5–8*, co-expression and immunoprecipitation of nontagged MZF1B NH$_2$ terminus and tagged MZF1B NH$_2$ terminus ΔSCAN. The *dotted arrows* indicate higher order protein complexes.

MZF1B (Fig. 4*D*).

The Carboxyl Terminus of RAZ1 Is Sufficient for MZF1B SCAN Box Interaction and RAZ1 Self-association—To identify RAZ1 domains sufficient for MZF1B SCAN box association, we performed two-hybrid assays with the MZF1B SCAN bait plasmid and either the amino terminus or carboxyl terminus of RAZ1 that contains the SCAN-related domain. As a control, we demonstrated that the individual constructs did not autonomously activate reporter gene expression (Tables II and IV). Positive protein interactions occurred when the MZF1B SCAN box domain was co-transformed with full-length RAZ1 or the carboxyl terminus of RAZ1 but not with the amino terminus of RAZ1 (Tables II and IV). This demonstrates that the carboxyl terminus of RAZ1 is sufficient for MZF1B SCAN box interaction, suggesting that the SCAN box domains from both proteins are mediating heteroprotein association.

To test whether RAZ1 self-associates via a SCAN-dependent

mechanism, we performed two-hybrid assays using RAZ1 fused to the GAL4 DNA BD and RAZ1 fused to the GAL4 AD. The individual plasmids did not autonomously activate reporter gene expression (Tables II and IV). Co-transformation of both RAZ1 fusion proteins resulted in colonies positive for protein interaction, suggesting that RAZ1 self-associates *in vitro* (Table IV). While amino acids 1–37 do not appear to be necessary for RAZ1 self-association, the carboxyl terminus of RAZ1 is necessary for self-association (Table IV). This demonstrates that the SCAN-related domain of RAZ1 mediates homo- as well as heteroprotein association.

The MZF1B SCAN Box Is Necessary for Self-association—In demonstrating that the SCAN box mediates heteroassociation between MZF1B and RAZ1 proteins as well as RAZ1 homoassociation, we reasoned that the SCAN box might also mediate MZF1B homoassociation. To first determine if MZF1B could self-associate, we performed co-immunoprecipitation assays

with epitope-tagged MZF1B and nontagged MZF1B. The Myc/His epitope tag adds ~2 kDa, and the two proteins are distinguishable by size as well as immunoreactivity with the epitope tag-specific antibodies, α-Myc or α-His. The MZF1B proteins were co-expressed *in vitro* and co-immunoprecipitated with control IgG, α-post-SCAN, or α-Myc. In lysates expressing both forms of MZF1B, we detected nontagged MZF1B when epitope-tagged MZF1B was immunoprecipitated with α-Myc, suggesting that MZF1B self-associates *in vitro* (Fig. 5*A*). To test for the possibility of nonspecific binding, we repeated the assays by co-expressing both MZF1B and epitope-tagged β-galactosidase. MZF1B did not associate with β-galactosidase, supporting our observation that MZF1B self-association is not an artifact of our co-immunoprecipitation conditions[4] (data not shown). In addition, the amino terminus of MZF1B is sufficient for self-association, since both epitope-tagged and nontagged MZF1B NH$_2$ terminus proteins were detected in the same immunocomplex (Fig. 5*B*). Finally, the MZF1B SCAN box is necessary for MZF1B self-association because nontagged MZF1B NH$_2$ terminus did not co-immunoprecipitate with MZF1B NH$_2$ terminus ΔSCAN (Fig. 5*C*). It should be noted that we consistently observe higher molecular mass bands of >200 and ~80 kDa in immunoprecipitated lysates expressing MZF1B and MZF1B NH$_2$ terminus, respectively (Fig. 5, *A* and *B*). While the identification of these bands has not been confirmed, they may represent higher order complexes of the 80-kDa MZF1B and 42-kDa MZF1B amino terminus.

DISCUSSION

RAZ1 Protein Structural Motifs—We have described the identification of *RAZ1*, a novel human cDNA clone isolated from a yeast two-hybrid screen based on interaction with the MZF1B SCAN box. The function of RAZ1 in unknown, but the predicted sequence contains conserved motifs that provide insight into RAZ1's potential role in regulating transcription factor function. *RAZ1* cDNA contains an open reading frame of 217 amino acids with a carboxyl-terminal region homologous to the SCAN box domain conserved in zinc finger proteins. Interestingly, the RAZ1 SCAN box is truncated and lacks the predicted third α-helix present in other SCAN box proteins. Thus, we have designated this as a SCAN-related domain. In contrast to other SCAN box proteins, *RAZ1* does not appear to encode zinc finger motifs based on the sequence that we have obtained.

The sequences for approximately 20 SCAN box-containing proteins have been reported and/or deposited into the GenBank[TM] data base (Table III). Of these, seven also encode a KRAB A and/or B domain, and 19 contain carboxyl-terminal zinc finger motifs. This suggests that the SCAN box is frequently associated with zinc finger motifs and sometimes with KRAB domains. The remaining SCAN box proteins that do not contain zinc finger motifs include p18, TRFA, PGC-2, and RAZ1. p18 and TRFA are partial clones that do not contain sufficient sequence to determine the presence or absence of zinc finger motifs. PGC-2 (peroxisome proliferator-activated receptor γ coactivator-2) is a murine adipogenic cofactor bound by the differentiation domain of the peroxisome proliferator-activated receptor γ (38). PGC-2 encodes a 142-amino acid protein with a carboxyl-terminal SCAN-related domain but no zinc finger motifs. The PGC-2 protein shares 76% identity to RAZ1 (49% at the NH$_2$ terminus; 97% at the COOH terminus), suggesting that PGC-2 may be the murine homologue of RAZ1.

It is possible that *RAZ1*, *PGC-2*, and potentially *p18* and

TRFA represent a novel gene family that contain SCAN box domains but lack zinc fingers. Similarly, the *SSX* gene family contains KRAB domains without zinc fingers. In addition to the KRAB domain, the *SSX* gene family encodes a novel transcription repression domain at the carboxyl terminus, SSXRD. Interestingly, this SSXRD domain exerts stronger repression than the KOX1 KRAB domain, and the KRAB-related domain fails to interact with the co-repressor TIF1β (KAP-1) (50–52). Therefore, the protein binding and repression function of the *SSX* genes that contain KRAB domains and lack zinc fingers appears to be different and distinct from KRAB proteins that contain zinc fingers. Thus, SCAN proteins that lack zinc fingers may contain other conserved domains that modify or define their function.

RAZ1 contains an arginine-rich region of 16 amino acids at the carboxyl terminus, immediately following the SCAN-related domain. Short (10–20-amino acid) arginine-rich sequences have been shown to mediate DNA and RNA binding as well as nuclear localization. The arginine-rich domain found in the amino terminus of the recombination activating gene, RAG-1, exhibits DNA binding activity (53). In addition, the arginine-rich motifs of HIV Rev and Tat proteins, bacteriophage λ N, ϕ21 N, and P22 N mediate RNA binding (reviewed in Ref. 54). Specifically, the human immunodeficiency virus (HIV) Rev binds to the Rev response element of HIV-1 as an α-helix and facilitates the nuclear export of unspliced HIV pre-mRNAs (reviewed in Ref. 54). There is an increasing amount of evidence that the arginine-rich domains present in HIV Rev, Tat, and human retroviruses T-cell leukemia virus type 1 also function as direct importin β-dependent nuclear localization signals (55, 56). Thus, the arginine-rich domain of RAZ1 may mediate DNA-RNA binding and/or function as a nuclear localization signal. Localization to the nucleus would place RAZ1 in the same cellular environment as zinc finger transcription factors, and nucleotide binding activity may allow RAZ1 to compete with other zinc finger SCAN proteins for DNA-RNA binding sites. In addition, several putative phosphorylation and *N*-myristoylation sites reside within RAZ1, suggesting that the function of RAZ1 may be regulated by post-translational modifications.

RAZ1 Gene Structure and mRNA Expression—The isolated *RAZ1* cDNA clone is 775 bp in length and contains a putative translation initiation start site, stop codon, and polyadenylation signal (5'-AAU GAA AAA-3'). Several ESTs, cDNA library clones, and RACE products share identity to *RAZ1*, suggesting that we have identified a bona fide transcript. In addition, some of the EST and 5'-RACE sequences contain an additional 108-bp insert between nucleotides 59 and 60 of *RAZ1*. It remains to be determined if both transcripts are expressed *in vivo*. Northern blot analysis detects an ~1-kilobase *RAZ1* transcript in both hematopoietic and nonhematopoietic cells, and ESTs from various tissues share identity to *RAZ1*, suggesting that the *RAZ1* gene is expressed in a variety of tissues. Based on RACE, cDNA library clones, ESTs, and chromosome 20 sequence, it is likely that we have obtained the complete 3'-end of the *RAZ1* transcript and are within a few hundred nucleotides of obtaining the entire 5'-end. We scanned the chromosome 20 sequence and found that the open reading frame upstream of *RAZ1* continues for 128 amino acids and contains a methionine with a weak Kozak consensus (Fig. 1*B*). However, further analysis is needed to confirm the complete 5'-end of the transcript.

Interestingly, the *RAZ1* gene is localized to chromosome 20q11.1–11.23. Deletion of the long arm of chromosome 20, most often 20q11.2–13, is associated with myeloid disorders, particularly myeloproliferative disorders, myelodysplastic syn-

---

[4] A. L. Haas and J. F. Morris, unpublished results. MZF1B* and β-galactosidase-mh proteins were co-expressed *in vitro* and co-immunoprecipitated with α-Myc antibodies. MZF1B* was not in the same immunocomplex as β-galactosidase-mh.

drome, acute lymphocytic leukemia, and acute myelogenous leukemia (57–59). This suggests that the genetic loss on chromosome 20q may provide a proliferative advantage to myeloid cells, possibly through the loss of a tumor suppressor gene. In addition, an increased copy number of DNA sequences from chromosome 20q has been observed in pancreatic cancers (60) and breast carcinomas (61), and trisomy of chromosome 20 is associated with the progression of papillary renal cell carcinomas (62). This indicates that a gain of chromosome 20q may facilitate uncontrolled cellular proliferation, possibly through the aberrant expression of an oncogene.

The *RAZ1* chromosome 20q11.1–11.23 location raises the question as to whether loss or gain of *RAZ1* contributes to any disorders associated with the locus. The myeloid proliferative disorders associated with loss of chromosome 20q are particularly interesting because MZF1/1B appears to be an important regulator of hematopoietic differentiation and proliferation (30–33). Therefore, *RAZ1* or *MZF1B* may be a tumor suppressor gene, and the interaction between RAZ1 and MZF1B may be necessary to elicit a tumor suppressor function. Thus, a genetic loss of RAZ1 might block tumor suppressor activity, thereby providing a proliferative advantage to hematopoietic cells.

The SCAN Box Function—Co-immunoprecipitation and yeast two-hybrid analyses demonstrate that MZF1B and RAZ1 associate *in vitro* via a SCAN box-dependent mechanism. In addition, the SCAN box domains are necessary for MZF1B and RAZ1 self-association. Therefore, we have demonstrated that the SCAN box is a protein interaction domain that mediates both hetero- and homoprotein associations. These findings suggest a novel cascade mediated by unique SCAN box protein complexes. To define cellular cascades regulated by SCAN box protein interactions, it will be necessary to identify *in vivo* SCAN box oligomers and the unique functions elicited by each complex. The identification of mechanisms regulated by SCAN box protein complexes will significantly impact our understanding of the transcriptional role of SCAN box zinc finger proteins and their associated factors.

The transcriptional activity of the SCAN box-containing zinc finger proteins MZF-2 and ZNF174 has been examined. Full-length murine MZF-2 does not activate reporter gene expression, but a truncated form of MZF-2 markedly enhances transcription (29). The transcriptionally active form of MZF-2 contains the SCAN box domain, but the SCAN box is not necessary for transcriptional activity. ZNF174 is a transcriptional repressor of reporter genes driven by the human tumor growth factor-β and platelet-derived growth factor-B promoters (21). The complete amino terminus of ZNF174, including the SCAN box domain, transcriptionally represses reporter gene expression when fused to a heterologous DNA binding domain, but the SCAN box domain is not sufficient for transcriptional repression. The ZNF174 repression domain is probably present within the remaining amino-terminal portion, and the SCAN box may or may not modify this function. Thus, the SCAN box does not appear to function as a transactivation or repression domain. These conclusions are supported by our personal observations[5] and reports by Pengue *et al.* (4) and Williams *et al.* (21), which show that the SCAN box does not confer transactivation or repression function onto a heterologous DNA binding domain. While the SCAN box is not an independent transactivation or repression domain, the SCAN box may function to

recruit co-repressors and transactivators necessary for transcriptional regulation.

MZF1B Function—We identified RAZ1 as a potential *in vivo* protein interaction partner with the SCAN box-containing zinc finger protein MZF1B, the human isoform of MZF1 previously identified by Peterson and Morris[1,2]. MZF1 was initially identified as a zinc finger transcription factor necessary for granulocytic differentiation and critical to the regulation of cell proliferation and apoptosis (30–33). In contrast to previous reports, both *MZF1B* and *MZF1* mRNA transcripts are expressed in numerous tissues.[2] Therefore, previous reports addressing MZF1 function may have been indirect measurements of MZF1B function. Thus, MZF1B may function as an important regulator of granulocytic differentiation, cell proliferation, and apoptosis. The interaction between MZF1B and RAZ1 might be necessary for mediating MZF1B function, or RAZ1 may modify intrinsic MZF1B function. It is also possible that other SCAN box proteins compete with MZF1B for binding to the same protein, thereby providing a transcriptional regulatory mechanism based on the sequestration of specific factors and availability of protein partners. Furthermore, MZF1B and RAZ1 each self-associate *in vitro*, suggesting that each protein may participate in the formation of unique complexes with distinct functions. Thus, the transcriptional activity of MZF1B is probably mediated by specific protein-protein interactions with RAZ1, MZF1B, and other SCAN box proteins. Identifying the *in vivo* MZF1B protein partners and their effect on MZF1B activity will provide insight into the possible mechanisms by which MZF1B functions to transcriptionally regulate cell development.

Addendum—During the course of review for this manuscript, Williams *et al.* (65) published their finding that the zinc finger-associated SCAN box is a conserved oligomerization domain.

REFERENCES

1. Miller, J., McLachlam, A. D., and Klug, A. (1985) *EMBO J.* **4,** 1609–1614
2. Bellefroid, 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
3. Margolin, J. F., Friedman, J. R., Meyer, W. K.-H., Vissing, H., Thiesen, H.-J., and Rauscher, F. J., III (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 4509–4513
4. Pengue, G., Calabro, V., Bartoli, P. C., Pagliuca, A., and Lania, L. (1994) *Nucleic Acids Res.* **22,** 2908–2914
5. Witzgall, R., O'Leary, E., Leaf, A., Onaldi, D., and Bonventre, J. V. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 4514–4518
6. Friedman, J. R., Fredericks, W. J., Jensen, D. E., Speicher, D. W., Huang, X.-P., Neilson, E. G., and Rauscher, F. J., III (1996) *Genes Dev.* **10,** 2067–2078
7. Moosmann, P., Georgiev, O., Le Douarin, B., Bourquin, J.-P., and Schaffner, W. (1996) *Nucleic Acids Res.* **24,** 4859–4867
8. Kim, S.-S., Chen, Y.-M., O'Leary, E., Witzgall, R., Vidal, M., and Bonventre, J. V. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 15299–15304
9. Bardwell, V. J., and Treisman, R. (1994) *Genes Dev.* **8,** 1664–1677
10. Dhordain, P., Albagli, O., Ansieau, S., Koken, M. H., Deweindt, C., Quief, S., Lantoine, D., Leutz, A., Kerckaert, J. P., and Leprince, D. (1995) *Oncogene* **11,** 2689–2697
11. Li, X., Lopez-Guisa, J. M., Ninan, N., Weiner, E. J., Rauscher, F. J., III, and Marmorstein, R. (1997) *J. Biol. Chem.* **272,** 27324–27329
12. Davies, J. M., Hawe, N., Kabarrowski, J., Huang, Q.-H., Zhu, J., Brand, N. J., Leprince, D., Dhordain, P., Cook, M., Morriss-Kay, G., and Zelent, A. (1999) *Oncogene* **18,** 365–375
13. Li, J.-Y., English, M. A., Ball, H. J., Yeyati, P. L., Waxman, S., and Licht, J. D. (1997) *J. Biol. Chem.* **272,** 22447–22455
14. Seyfert, V. L., Allman, D., He, Y., and Staudt, L. M. (1996) *Oncogene* **12,** 2331–2342
15. Deweindt, C., Albagli, O., Bernardin, F., Dhordain, P., Quief, S., Lantoine, D., Kerckaert, J. P., and Leprince, D. (1995) *Cell Growth Differ.* **6,** 1495–1503
16. David, G., Alland, L., Hing, S. H., Wong, C. W., DePinho, R. A., and Dejean, A. (1998) *Oncogene* **16,** 2549–2556
17. Dhordain, P., Lin, R. J., Quief, S., Lantoine, D., Kerckaert, J.-P., Evans, R. M., and Albagli, O. (1998) *Nucleic Acids Res.* **26,** 4645–4651
18. Huynh, K. D., and Bardwell, V. J. (1998) *Oncogene* **17,** 2473–2484
19. Wong, C. W., and Privalsky, M. L. (1998) *J. Biol. Chem.* **273,** 27695–27702
20. Deltour, S., Guerardel, C., and Leprince, D. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 14831–14836

---

[5] T. L. Sander, A. L. Haas, M. J. Peterson, and J. F. Morris, unpublished observations. When fused in frame to a GAL4 DNA binding domain, the MZF1B SCAN box does not activate the expression of a luciferase reporter gene under the control of a minimal promoter containing five GAL4 binding sites.

21. Williams, A. J., Khachigian, L. M., Shows, T., and Collins, T. (1995) *J. Biol. Chem.* **270,** 22143–22152
22. Dovat, S., Gilbert, K. A., Petrovic-Dovat, L., and Rannels, D. E. (1998) *Biochim. Biophys. Acta* **1442,** 380–388
23. Yokoyama, M., Nakamura, M., Okubo, K., Matsubara, K., Nishi, Y., Matsumoto, T., and Fukushima, A. (1997) *Biochim. Biophys. Acta* **1353,** 13–17
24. Chen, X., Hamon, M., Deng, Z., Centola, M., Sood, R., Taylor, K., Kastner, D. L., and Fischel-Ghodsian, N. (1999) *Biochim. Biophys. Acta* **1444,** 218–230
25. Murai, K., Murakami, H., and Nagata, S. (1997) *Genes Cells* **2,** 581–591
26. Hromas, R., Collins, S. J., Hickstein, D., Raskind, W., Deaven, L. L., O'Hara, P., Hagen, F. S., and Kaushansky, K. (1991) *J. Biol. Chem.* **266,** 14183–14187
27. Morris, J. F., Hromas, R., and Rauscher F. J., III (1994) *Mol. Cell. Biol.* **14,** 1786–1795
28. Morris, J. F., Rauscher, F. J., III, Davis, B., Klemsz, M., Xu, D., Tenen, D., and Hromas, R. (1995) *Blood* **86,** 3640–3647
29. Murai, K., Murakami, H., and Nagata, S. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 3461–3466
30. Bavisotto, L., Kaushansky K., Lin N., and Hromas, R. (1991) *J. Exp. Med.* **174,** 1097–1101
31. Hromas, R., Boswell, S., Shen, R.-H., Burgess, G., Davidson, A., Cornetta, K., Sutton, J., and Robertson, K. (1996) *Leukemia* **10,** 1049–1050
32. Hromas, R., Morris, J., Cornetta, K., Berebitsky, D., Davidson, A., Sha, M., Sledge, G., and Rauscher, F. (1995) *Cancer Res.* **55,** 3610–3614
33. Robertson, K. A., Hill, D. P., Kelley, M. R., Tritt, R., Crum, B., Van Epps, S., Srour, E., Rice, S., and Hromas, R. (1998) *Leukemia* **12,** 690–698
34. Patwardhan, S., Gashler, A., Siegel, M. G., Chang, L. C., Joseph, L. J., Shows, T. B., Le Beau, M. M., and Sukhatme, V. P. (1991) *Oncogene* **6,** 917–928
35. Denhardt, D. (1966) *Biochem. Biophys. Res. Commun.* **23,** 641–646
36. Fuller, S. A., Takahashi, M., and Hurrell, J. G. R. (1997) in *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K., eds) Vol. II, pp. 11.11.1–11.11.5, John Wiley & Sons, Inc., New York
37. Kozak, M. (1987) *J. Mol. Biol.* **196,** 947–950
38. Castillo, G., Brun, R. P., Rosenfield, J. K., Hauser, S., Park, C.-W., Troy, A. E., Wright, M. E., and Spiegelman, B. M. (1999) *EMBO J.* **18,** 3676–3687
39. Calabro, V., Pengue, G., Bartoli, P. C., Pagliuca, A., Featherstone, T., and Lania, L. (1995) *Hum. Genet.* **95,** 18–21
40. Attar R. M., and Gilman, M. Z. (1992) *Mol. Cell. Biol.* **12,** 2432–2443
41. Lee, P. L., Gelbart, T., West, C., Adams, M., Blaskstone, R., and Beutler, E. (1997) *Genomics* **43,** 191–201
42. Tirosvoutis, K. N., Divane, A., Jones, M., and Affara, N. A. (1995) *Genomics* **28,** 485–490
43. Nagase, T., Ishikawa, K., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N., and O'Hara, O. (1998) *DNA Res.* **5,** 31–39
44. Monaco, C., Helmer Citterich, M., Caprini, E., Vorechovsky, I., Russo, G., Croce, C. M., Barbanti-Brodano, G., and Negrini, M. (1998) *Genomics* **52,** 358–362
45. Chowdhury, K., Goulding, M., Walther, C., Imai, K., and Fickenscher, H. (1992) *Mech. Dev.* **39,** 129–142
46. Noce, T., Fujiwara, Y., Sezaki, M., Fujimoto, H., and Higashinakagawa, T. (1992) *Dev. Biol.* **153,** 356–367
47. Yang, X. W., Zhong, R., and Heintz, N. (1996) *Development* **122,** 555–566
48. Denny, P., and Ashworth, A. (1991) *Gene* (*Amst.*) **106,** 221–227
49. Gonsky, R., Knauf, J. A., Elisei, R., Wang, J. W., Su, S., and Fagin, J. A. (1997) *Nucleic Acids Res.* **25,** 3823–3831
50. Crew, A. J., Clark, J., Fisher, C., Gill, S., Grimer, R., Chand, A., Shipley, J., Gusterson, B. A., and Copper, C. S. (1995) *EMBO J.* **14,** 2333–2340
51. de Leeuw, B., Balemans, M., and Geurts van Kessel, A. (1996) *Cytogenet. Cell Genet.* **73,** 179–183
52. Lim, F. L., Soulez, M., Koczan, D., Thiesen, H. J., and Knight, J. C. (1998) *Oncogene* **17,** 2013–2018
53. Galic, Z., Alva, J. A., Lin, A., Lyon, C. J., and Aguilera, R. J. (1998) *Biochem. Mol. Biol. Int.* **45,** 535–544
54. Burd, C. G., and Dreyfuss, G. (1994) *Science* **265,** 615–621
55. Palmeri, D., and Malim, M. H. (1999) *Mol. Cell. Biol.* **19,** 1218–1225
56. Truant, R., and Cullen, B. R. (1999) *Mol. Cell. Biol.* **19,** 1210–1217
57. Johansson, B., Mertens, F., and Mitelman, F. (1993) *Genes Chromosomes Cancer* **8,** 205–218
58. Kurtin, P. J., Dewald, G. W., Shields, D. J., and Hanson, C. A. (1996) *Am. J. Clin. Pathol.* **106,** 680–688
59. Wang, P. W., Iannantuoni, K., Davis, E. M., Espinosa, R., III, Stoffel, M., and Le Beau, M. M. (1998) *Genes Chromosomes Cancer* **21,** 75–81
60. Ghadimi, B. M., Schrock, E., Walker, R. L., Wangsa, D., Jauho, A., Meltzer, P. S., and Ried, T. (1999) *Am. J. Pathol.* **154,** 525–536
61. Guan, X. Y., Xu, J., Anzick, S. L., Zhang, H., Trent, J. M., and Meltzer, P. S. (1996) *Cancer Res.* **56,** 3446–3450
62. Palmedo, G., Fischer, J., and Kovacs, G. (1999) *Lab. Invest.* **79,** 311–316
63. CLONTECH, *Yeast Protocol Handbook PT3024–1*, pp. 29–30, CLONTECH, Palo Alto, CA
64. Genetics Computer Group, *Wisconsin Package Version 10.0*, Genetics Computer Group, Madison, WI
65. Williams, A. J., Blacklow, S. C., and Collins, T. (1999) *Mol. Cell Biol.* **19,** 8526–8535

Identification of a Novel SCAN Box-related Protein That Interacts with MZF1B: THE LEUCINE-RICH SCAN BOX MEDIATES HETERO- AND HOMOPROTEIN ASSOCIATIONS

Tara L. Sander, Amy L. Haas, Michael J. Peterson and Jennifer F. Morris

Access the most updated version of this article at http://www.jbc.org/content/275/17/12857

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 63 references, 26 of which can be accessed free at
http://www.jbc.org/content/275/17/12857.full.html#ref-list-1