512–520 Nucleic Acids Research, 1998, Vol. 26, No. 2 © 1998 Oxford University Press

Interaction of the CNC-bZIP factor TCF11/LCR-F1/Nrf1 with MafG: binding-site selection and regulation of transcription

Øyvind Johnsen, Paula Murphy, Hans Prydz and Anne-Brit Kolstø*

Biotechnology Centre of Oslo, University of Oslo, PO Box 1125, Blindern N-0316 Oslo, Norway

Received September 23, 1997; Revised and Accepted November 7, 1997

ABSTRACT

We have previously shown that the widely expressed human transcription factor TCF11/LCR-F1/Nrf1 interacts with small Maf proteins and binds to a subclass of AP1-sites. Such sites are required for β-globin 5′ DNase I hypersensitive site 2 enhancer activity, erythroid porphobilinogen deaminase inducibility, hemin responsiveness by heme-oxygenase 1 and expression of the gene NAD(P)H:quinone oxidoreductase$_1$. Here we report the optimal DNA-binding sequences for TCF11/LCR-F1/Nrf1 alone and as a heterodimer with MafG, identified by using binding-site selection. The heterodimer recognises a 5′-TGCTgaGTCAT-3′ binding-site that is identical to the established NF-E2-site, the antioxidant response element and the heme-responsive element while the binding specificity of the homomer is less stringent. To investigate the activity of TCF11 through this selected site, both alone and in the presence of MafG, we have used a transient transfection assay. TCF11 alone activates transcription while MafG alone acts as a repressor. When co-expressed, MafG interferes with TCF11 transactivation in a dose dependent manner. This indicates that MafG protein, which heterodimerises efficiently with TCF11 *in vitro* (the heterodimer having a higher affinity for DNA than TCF11 alone), does not co-operate with TCF11 in transactivating transcription. We propose that since both these factors are widely expressed, they may act together to contribute to the negative regulation of this specific target site. Efficient positive regulation by TCF11 may require alternative partners with perhaps more restricted expression patterns.

INTRODUCTION

CNC-bZIP proteins are identified by an ~40 amino acid homology region immediately N-terminal to the basic region-leucine zipper (bZIP) domain. The *Drosophila melanogaster* homeotic selector gene *cap 'n' collar* (1,2) encodes the first bZIP-factor identified that contained this region. Other family members include Skn1, a basic-region transcription factor required for correct specification of certain blastomere fates in early *Caenorhabditis elegans* embryos (3,4) and three human proteins; p45 NF-E2, an erythroid-specific activator proposed to regulate the β-globins (5,6) and the more widely expressed TCF11/LCR-F1/Nrf1 (7–11) [hereafter referred to as TCF11 (transcription factor 11)] and Nrf2 (12). Homologous and related genes have also been cloned in other vertebrate species (13–15). The leucine zipper, which is responsible for homo- and heterodimerisation is not particularly conserved among CNC-bZIP family members (8). However, all CNC-bZIP family members tested so far preferentially form heterodimers with the same group of small Maf proteins (11,13,14,16). The Maf family of bZIP factors, the prototype of which (v-Maf) is responsible for the transforming activity present in an avian retrovirus (17), is subdivided into two groups based on primary sequence and ability to activate transcription. The large Maf family members are transactivators and regulate genes important in neuronal differentiation (18–20) whereas the small Maf proteins, MafF, MafG and MafK/p18, are widely expressed and transrepress transcription when bound to Maf responsive elements [MAREs; 5′-TGCTGAC(G)TCAGCA-3′], probably as homodimers (16,21–25).

We have previously shown that chicken Maf proteins MafF, -G and -K, but not the large activators v-Maf or MafB, specifically interact with several TCF11 protein isoforms *in vitro* (11). Both TCF11 and the Maf proteins -F, -G and -K/p18 are widely expressed, implying that these proteins may interact in numerous cell types (8,21,23,26,27). Heterodimers of small Maf proteins and CNC-bZIP domain family members or Fos, bind preferentially to a site containing consensus sequences for both Maf homodimer and AP1 binding called the AP1/MARE-site (5′-TGCTGAGTCAT/C-3′) (11,13,14,16,21). This is in fact a classical AP1 site with a 5′-TGC extension, as is found in the NF-E2 site implicated in the regulation of erythroid specific gene expression. In murine erythroleukemia (MEL) cells, one of the binding complexes that specifically recognises this site was isolated and shown to consist of a heterodimer between the CNC-bZIP domain family member p45 NF-E2 and the small Maf homologue p18 (22). Several other AP1/MARE-sites require the 5′-TGC triplet for their correct function, and as yet unidentified binding-activity to these sites has been detected in a range of cell-types (6,22,27–30) including F9 cells (28) which do not express NF-E2 or AP1 factors.

Using DNA affinity chromatography we observed sequence-specific binding activity of the endogenous TCF11 isoforms p47/49 to the NF-E2-site in K562 cells (11). *In vitro* binding-site selection experiments presented here show that heterodimerisation dramatically increases the DNA-binding potential of TCF11. We

*To whom correspondence should be addressed. Tel: +47 22 95 84 60; Fax: +47 22 69 41 30; Email: a.b.kolsto@farmasi.uio.no

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

observed a 5′-TGCTgaGTCAT-3′ binding-site that is identical to the NF-E2-site (6), the antioxidant response element (ARE) (31) and the heme response element (HRE) (30). TCF11 alone shows only limited sequence-specificity. The Maf halfsite recognised by the heterodimer contains the previously described 5′-TGC AP1 site-extension. This TCF11:Maf-site represents the first selected CNC-bZIP heterodimer binding-site and can aid in identifying genes regulated by this presumably widespread heterodimer. The functional relevance of this binding has been tested in transfected COS 1 cells and it was found that while TCF11 can transactivate a reporter construct through the NF-E2 site when expressed alone, MafG co-expression interferes with this activation in a dose dependent manner. This is in contrast to the co-operative activation observed for NF-E2:small Maf heterodimers in NIH3T3 cells (16) and Ech:small Maf heterodimers in quail fibroblasts (21) demonstrating that different heterodimeric combinations of CNC-bZIP factors and small Mafs may have different activities through the same or similar target sequences.

MATERIALS AND METHODS

Standard methods in molecular biology were used (32).

Plasmid constructs and fusion proteins

MBP-TCF11-A is a fusion protein between *Escherichia coli* maltose-binding protein (MBP) and the 300 C-terminal amino acids of human TCF11, including the CNC-bZIP region, constructed in the plasmid pMALc [New England Biolabs (NEB)] as previously described (11). MBP-MafG is an MBP fusion protein with chicken MafG (21), which has only one conserved amino acid substitution compared to human MafG in the bZIP domain. Production of fusion proteins and their isolation by amylose affinity chromatography were performed according to NEB protocols. Proteins were stored at –20°C in protein elution buffer/25% ethylene glycol (Pierce) (11).

The reporter constructs for cell transfections were produced by cloning; (i) the 900 bp *Eco*RI–*Bam*HI fragment (PBGD1.5Luc), (ii) the 320 bp *Acc*I–*Bam*HI fragment (PBGD3.2Luc) or (iii) the 180 bp *Pvu*II–*Bam*HI fragment (PBGD5.1Luc), of the porphobilinogen deaminase (PBGD) gene erythroid-specific promoter (33) in front of the firefly luciferase gene in the pGL3 enhancer vector (Promega). Site directed mutagenesis of the NF-E2 site within construct PBGD3.2Luc was performed using the Stratagene 'Quick change' mutagenesis kit. Expression constructs were produced by cloning the full length coding sequence of TCF11 (8) (5′ to the *Eco*RV site at bp 3550) or MafG wild-type and mutant form (ΔL2PM4P, 21) into the expression vector pCDNA3 (Invitrogen).

Binding-site selection

The DNA library R76 (36) consists of 26 randomised nucleotides flanked by 25 bp constant regions used for PCR amplification and subcloning: 5′-CAGGTCAGTTCAGCGGATCCTGTCG(G/A/T/C)$_{26}$GAGGCGAATTCAGTGCAACTGCAGC-3′, synthesised by Dr Eshrat Babaie, The Biotechnology Centre of Oslo. It was rendered double-stranded using the Klenow fragment and primer F (5′-GCTGCAGTTGCACTGAATTCGCCTC-3′). Primer R is 5′-CAGGTCAGTTCAGCGGATCCTGTCG-3′. Double-stranded R76 was purified by polyacrylamide electrophoresis (32) and suspended in TE (10 mM Tris–HCl, 1 mM EDTA, pH 8.0). MBP-TCF11-A (1.3 pmol) (and MBP-MafG when heterodimer binding-sites were selected) was incubated in binding buffer [5 mM Tris–HCl, pH 8.0, 75 mM NaCl, 2.5 mM $MgCl_2$, 0.5 mM EDTA, 5% glycerol, 2.5% ethylene glycol, 1% Tween-20, 1 mM DTT and 100 μg/ml poly(dIdC)] for 10 min at room temperature before the addition of 0.1 nmol dsR76 in a total volume of 20 μl. DNA-binding was allowed at room temperature for 20 min. Protein–DNA was electrophoresed through a 5% polyacrylamide gel (acrylamide:bisacrylamide 36:1) in 0.4× TBE (1× TBE is 89 m Tris–borate, pH 8.3, and 2 mM EDTA) at 10 V/cm for 10 min. The upper 0.5 cm of the gel, containing protein–DNA complexes, was excised. Protein-bound oligonucleotides were isolated by crushing the gel-slice and incubating in 400 μl NEMS solution (0.5 M NH_4Ac, 10 mM MgAc, 1 mM EDTA and 0.1% SDS) for 12–15 h at 37°C with agitation. Oligonucleotides present in the supernatant were precipitated by the addition of 1 ml ice-cold ethanol (2.5 vol) and 20 μg dextran, as carrier. The DNA was collected by centrifugation (13 000 r.p.m. for 30 min). Pelleted DNA was washed briefly in cold 70% ethanol and suspended in 10 μl ddH_2O and amplified by nine cycles of PCR using 0.1 nmol of each primer and Taq polymerase (Gibco) in a total volume of 20 μl. NEMS solution was added to the PCR-mix (final volume 200 μl), and DNA precipitated and washed as described. Binding-site selection was then repeated as described, including 1 μl (10 μCi/ml) [α-^{32}P]dATP (Amersham) in the PCR reaction for radiolabelling of DNA. Electrophoresis time in the second and subsequent rounds was extended to 75 min to resolve MBP-TCF11-A:MBP-MafG heterodimers from MBP-MafG homodimers. Protein-bound DNA was identified by autoradiography. DNA from round three was digested with *Eco*RI and *Bam*HI and ligated into the corresponding sites of pBluescript SKII+ (Stratagene). Plasmids were transformed into *E.coli* DH5α and the nucleotide sequences of individual clones were analysed (Tables 1 and 2).

A note on library design: of the 36 TCF11:MafG-selected oligonucleotides analysed (Table 2A), 21 molecules utilised nucleotides flanking the randomised region (18 used primer F; 5′-C**GC**CCTC and three used primer R; 5′-C**TG**TC**G**, Table 2D) instead of the MafG consensus halfsite nucleotides 5′-TGCTga (Table 2B). As a substrate, the GC doublet seems to be more important than the TG doublet, since in the first case, a mismatch of four bases is tolerated. This is not evident, however, for sites selected inside the randomised sequence where 13, 13 and 12 molecules contain T, G and C, respectively (Table 2D, positions –3, –2 and –1). In the design of future oligonucleotide libraries for the analysis of Maf containing complexes it will be important to avoid such potential protein binding to the constant primer regions.

Transient transfection assay

The various cell lines (acquired from ATCC) were grown in the recommended media at 37°C in 5% CO_2. At ~60–70% confluence the cells were transfected with plasmid DNA by a standard calcium phosphate precipitation method (34) using a total of 10–15 μg of DNA per 9 cm culture dish. The DNA mixture consisted typically of 2 μg of luciferase reporter construct, 1 μg of internal control plasmid (either pRSV-CAT or pEFβ-gal), various amounts of the TCF11 and/or MafG expression constructs and empty vector to the required total weight. The cultures were grown for 48 h after transfection before harvesting and enzyme assays were performed. Luciferase activity was measured on a Lumat LB 9507 luminometer

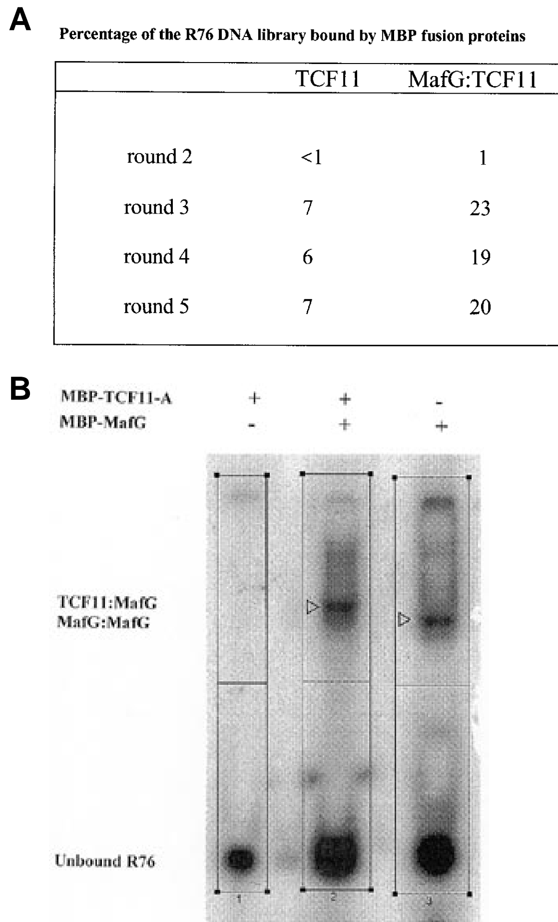

Figure 1. DNA binding of selected R76 oligonucleotide pools bound to MBP-TCF11-A and MBP-TCF11-A:MBP-MafG. (**A**) Percentage binding of selected oligonucleotides to the indicated protein(s). (**B**) Protein bound dsR76 oligonucleotide pools after three rounds of selection, separated on a polyacrylamide gel and viewed in a phosphoimager. Proteins used in the binding-reaction are indicated. Arrowheads indicate the complexes corresponding to MBP-TCF11-A:MBP-MafG heterodimer and MBP-MafG homodimer bound to the selected oligonucleotides.

Table 1. Repertoire of *in vitro* selected oligonucleotides bound by MBP-TCF11-A

(**A**) Sequences selected by the TCF11 homomer. AP1-like sequences are in bold.

	Primer F	Selected sequences	Primer R
A1	CGCCTC	CTAAAAAGGACGAACACAA**TGACTC**	CGACAG
A2	CGCCTC	TATAAACCAGTNC**TCA**ATA	CGACAG
A3	CGCCTC	AAGCTAAATCCCCGAAACCGCTGCGT	CGACAG
A4	CGCCTC	AATGCACAT**TCA**AACGCCA**TCA**CTTT	CGACAG
A5	CGCCTC	**GTCA**TGACATCTCTATGTAACCTAAA	CGACAG
A6	CGCCTC	ATCTTAACCTTTAATA	CGACAG
A7	CGCCTC	GACCACTAAACCAGA	CGACAG
A8	CGCCTC	CAATTATAATTATA**TCA**AT**TGA**TCTA	CGACAG
A9	CGCCTC	TAGTACCTA**TCA**TAACAACGTCCCCA	CGACAG
A11	CGCCTC	CGCCCACCCA**TGACTC**TACGCACAAT	CGACAG
A12	CGCCTC	TAGA**TGA**ATTGTTCACTCAAACGCTA	CGACAG
A13	CGCCTC	A**TCA**TCAAGATTTCC**TCA**GTATAGTA	CGACAG
A14	CGCCTC	ACCTTACACATAACGCAACACC	CGACAG
A15	CGCCTC	AATATCCTGGGCAGCTCTTAATTAAA	CGACAG
A16	CGCCTC	TGTCCCCAAAGC**TGACT**AAGCAATAA	CGACAG
A17	CGCCTC	ATA**GTCA**CATTCTCACAA**TCA**AACAA	CGACAG
A18	CGCCTC	TACC**TCA**CGTGTACGCCTCCTTA	CGACAG
A21	CGCCTC	AATAACAACTAWCCTTTAAGATCTAC	CGACAG
A22	CGCCTC	GCTAGTATACCCAAACCCGTTCCCA	CGACAG
A24	CGCCTC	CCATACCAATAACT**TCA**ATTTCCGAA	CGACAG
A25	CGCCTC	GTTAACGTGCTT**GTCA**CGCACACAAT	CGACAG
A26	CGCCTC	AACGAATAACTGGGAGC**TCA**ACT	CGACAG
A27	CGCCTC	TCCTCCCATTCTCCTTCATTT**TCA**NN	CGACAG
A28	CGCCTC	CTAAACAAGGTCC**TGAC**CTTTCGAGA	CGACAG
A29	CGCCTC	AA**TGAC**AATAGGAC**TCA**ACTCTA**TGA**	CGACAG
A30	CGCCTC	TTAATGGACGACACCGTTGCTAA**TGA**	CGACAG
A31	CGCCTC	AGAA**GTCA**AAATTCTAAATACCAAA	CGACAG
A32	CGCCTC	ACCATATATTA**TCA**CCTTACTATGGA	CGACAG
A35	CGCCTC	TCCTACCCTAAAATCTTTAAAAAT	CGACAG
A36	CGCCTC	TGTTTTACCGTACCCTCAGTTACN	CGACAG
A37	CGCCTC	AAACCAATTTAGCAATTATAGCTCTC	CGACAG
A38	CGCCTC	**GTCATCA**TTCTTANC	CGACAG
A39	CGCCTC	ATGTAGTATATCGACCAGAAC**TGA**TA	CGACAG
A40	CGCCTC	TATCATGCTACTCCCCCTGCCTACA	CGACAG

(**B**) Relative abundance of TCA-containing half-sites in the sequences shown in (A).

TCA-containing halfsites	observed	expected	obs/exp
GTCAG:	2.0	1.6	**1.3**
GTCAA:	1.0	1.6	**0.6**
GTCAT:	7.0	1.6	**4.4**
GTCAC:	2.0	1.6	**1.3**
TCA:	31.0	25.8	**1.2**

using 5–50 μl of cell extract (from a total of 700 μl) brought to a total volume of 200 μl in buffer containing 10 mM Mg(OAc), 50 mM Tris–MES, pH 7.8, 2 mM ATP. Aliquots (100 μl) of 1 mM luciferin (Sigma L6882) were added for each measurement. The Luciferase activity for each culture was normalised to the activity of the internal control, either CAT or β-galactosidase. The β-galactosidase control was used in experiments with MafG expression vector since MafG was found to interfere with CAT expression in the pRSV-CAT construct. CAT activity was measured using a standard protocol (35) and β-galactosidase activity was measured using ONPG as substrate and a colorimetric assay.

RESULTS

Binding sites selected by TCF11

The approach used to select optimal binding sequences involved incubating the protein with a pool of degenerate oligonucleotides followed by isolation of the bound complexes on a polyacrylamide gel and PCR amplification of the retained oligonucleotide fraction. This binding, selection and amplification is repeated cyclically before selected oligonucleotides are cloned, sequenced and analysed. The procedure has been widely applied to identify transcription factor binding-sites (24,36,37). To retain maximum complexity of the selected fraction we have limited the numbers of selection-amplification cycles. The formation of protein–DNA complexes was monitored throughout the experiment. We observed a strong increase in binding from the second to the third round and after three rounds of selection and amplification, no further increase in DNA-binding was obtained (Fig. 1). The amplified binding-sites were therefore cloned after the third round.

The TCF11 homomer, bound to the library DNA, did not enter a polyacrylamide gel upon electrophoresis but remained in the well-region. This was consistent with previous observations using specific oligonucleotides and confirmed that it was not an artefact resulting from the nature of the synthetic oligonucleotide probes that were previously used (11). The limited mobility of the TCF11 binding complex suggests that the fusion protein MBP-TCF11-A binds to DNA in a multimeric form. The sequences of oligonucleotides within this complex showed that the AP1

Table 2. MBP-TCF11-A:MBP-MafG selected sequences

(**A**) Sequences of *in vitro* selected DNA molecules that bound to MBP-TCF11-A:MBP-MafG. TCF11:MafG consensus sequences (B) are shown in bold.

	Primer F	Selected sequences	Primer R
AG1	CGCCTC	**GTCA**TTATTGCAGTGGGCAACCTCT	CGACAG
AG2	CGCCTC	**GTCA**TTCAGGGCCGCTACCTA	CGACAG
AG3	CGCCTC	**GTCA**TCCTAGTCCAATTCTTACTAT	CGACAG
AG4	CGCCTC	**ATCA**TGCATTGATTCAATACAAACAA	CGACAG
AG5	CGCCTC	ACACAAC**TGC**ACT**GTCA**TTCAGTCG	CGACAG
AG6	CGCCTC	ACATA**TGCTGCGTCA**TATCATTACTC	CGACAG
AG7	CGCCTC	**GTCA**TATAACCTTCACTATAGAACA	CGACAG
AG8	CGCCTC	AAG**TGCTTACTCA**TCCATCTTAAAA	CGACAG
AG9	CGCCTC	**GTCA**TCATGCTGACAAGGTTTCTGT	CGACAG
AG10	CGCCTC	AGAAC**TGCTGACTCA**ATATGTGGCT	CGACAG
AG12	CGCCTC	**ATCA**TCCACCCTGATGGACACATA	CGACAG
AG13	CGCCTC	**GTCA**TCTTAGCATCCTCACTT	CGACAG
AG14	CGCCTC	AACCAGCATTACGCAGCACATGTA	CGACAG
AG16	CGCCTC	GTACGATAACGTTGCATGATGAA	CGACAG
AG17	CGCCTC	AAATA**TGACT**T**AGCA**ATTACGTA	CGACAG
AG18	CGCCTC	GTATATGATATAATAGTCTCC	CGACAG
AG19	CGCCTC	CACCTACCACATAAT**TGCT.AGTCA**	CGACAG
AG20	CGCCTC	TAATCACTCAACAACTATA**TGACT**	CGACAG
AG22	CGCCTC	**GTCA**TTCTAAACATAACACCA	CGACAG
AG23	CGCCTC	**GTCA**TCACTCCCAACAGTTTAGCATT	CGACAG
AG27	CGCCTC	**GTCA**TGCTTCAACTCGCTCC	CGACAG
AG28	CGCCTC	GACTAACCTTGTCTCACCTCA**TGACT**	CGACAG
AG36	CGCCTC	**GCTGAGTCA**CCTTCCACGAATTGGCCG	CGACAG
AG38	CGCCTC	TACCA**TGAC**G**CAGCA**ACTCCTGAGAC	CGACAG
AG39	CGCCTC	GCCCTTTTC**T**ACCA**AGTCA**CTTATCC	CGACAG
AG41	CGCCTC	AAACATCACAATGACATCACACTTTT	CGACAG
AG42	CGCCTC	AA**TGACACTGCA**ATTCCTCTTTANNG	CGACAG
AG43	CGCCTC	A**TCA**TTCTCATACCTTAACACAATAC	CGACAG
AG44	CGCCTC	ATCCCAAAATTCACGACGTCATGTAT	CGACAG
AG45	CGCCTC	AACAAACA**TGAC**GA**AGCA**CTTATAAT	CGACAG
AG46	CGCCTC	AATATTTCCCG**TGACTCAACA**TAATT	CGACAG
AG47	CGCCTC	ATCTCACATA**ATGACA**CTA**CATGACT**	CGACAG
AG48	CGCCTC	**GTCA**TCGTGAATTGCCATCTAGTTCA	CGACAG
AG49	CGCCTC	ACTCCAAAAGCGCTTGCTCAGCAAAC	CGACAG
AG50	CGCCTC	CTA**TGACTCAGCA**ATAAAGT	CGACAG
AG52	CGCCTC	**GTCA**CTTATTACTAGGCNCACTTAGT	CGACAG

(**B**) Calculation of the TCF11:MafG consensus sequence from sites present inside the library DNA (A).

	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	12	13
G	2	3	5	0	1	2	0	13	0	0	8	0	12	0	0	0	0	4	4	2	4	6
A	2	7	1	9	8	3	0	1	1	2	2	8	0	0	0	14	1	3	3	4	5	1
T	1	3	5	4	3	6	13	0	1	11	2	3	0	14	0	0	9	5	5	6	4	4
C	8	1	3	1	2	3	1	0	12	1	1	3	2	0	14	0	4	2	2	2	1	3
	n	**n**	**n**	**a**	**a**	**n**	**T**	**G**	**C**	**T**	**g**	**a**	**G**	**T**	**C**	**A**	**t**	**n**	**n**	**n**	**n**	**n**

(**C**) Calculation of the TCF11 half-site using all sequences in (A).

	+4	+5	+6	+7	+8	+9	+10	+11	+12	+13	+14	+15	+16
G	26	-	-	-	-	8	5	3	10	10	7	7	10
A	3	-	-	31	1	5	7	11	9	4	8	8	5
T	0	31	-	-	25	10	10	14	9	9	7	7	6
C	2	-	31	-	5	8	9	3	3	8	8	8	8
	G	**T**	**C**	**A**	**T**	**N**	**N**	**T/A**	**N**	**N**	**N**	**N**	**N**

(**D**) Alignment of the oligonucleotides in (A) using the TCF11:MafG consensus sequence (B).

Binding-sites including primer-sequences (underlined):

	Primer F	
AG1	CGCCTC	**GTCA**TTATTGCAG
AG2	CGCCTC	**GTCA**TTCAGGGCC
AG3	CGCCTC	**GTCA**TCCTAGTCC
AG4	CGCCTC	**ATCA**TGCATTGAT
AG7	CGCCTC	**GTCA**TATAACCTT
AG9	CGCCTC	**GTCA**TCATGCTGA
AG12	CGCCTC	**ATCA**TCCACCCTG
AG13	CGCCTC	**GTCA**TCTTAGCAT
AG22	CGCCTC	**GTCA**TTCTAAACA
AG23	CGCCTC	**GTCA**TCACTCCCA
AG27	CGCCTC	**GTCA**TGCTTCAAC
AG43	CGCCTC	**ATCA**TTCTCATAC
AG48	CGCCTC	**GTCA**TCGTGAATT
AG52	CGCCTC	**GTCA**CTTATTACT

	Primer R	
AG20	TGTCG	**AGTCA**TATAGTTGT
AG28	TGTCG	**AGTCA**TGAGGTGAG
AG47	TGTCG	**AGTCA**TGTAGTGTC

Binding-sites not including primer sequences:

Position	-5-4	-3	-2	-1	1	2	3	4	5	6	7	8910
AG5	CACAAC	T	G	C	*A*	*C*	*T*	G	T	C	A	TTCAGTCGC
AG6	CACATA	T	G	C	T	G	*C*	G	T	C	A	TATCATTAC
AG8	CTCAAG	T	G	C	T	*T*	A	*C*	T	C	A	TCCATCTTA
AG10	CAGAAC	T	G	C	T	G	A	*C*	T	C	A	ATATGTGGC
AG17	CGTAAT	T	G	C	T	*A*	A	G	T	C	A	TATTTGAGG
AG19	CATAAT	T	G	C	T	.	A	G	T	C	A	CGACAG
AG36	CGCCT	C	G	C	T	G	A	G	T	C	A	CCTTCCACG
AG38	AGGAGT	T	G	C	T	G	*C*	G	T	C	A	TGGTAGAGG
AG39	CTTTTC	T	*A*	C	*C*	*A*	A	G	T	C	A	CTTATCCCG
AG42	AGGAAT	T	G	C	*A*	G	*T*	G	T	C	A	TTGAGGCG
AG45	TATAAG	T	G	C	T	*T*	*C*	G	T	C	A	TGTTTGTTG
AG46	GAATTA	T	G	*T*	T	G	A	G	T	C	A	CGGGAAATA
AG47	GAGTCA	T	G	*A*	T	G	*T*	G	T	C	A	TTATGTGAG
AG50	CTTTAT	T	G	C	T	G	A	G	T	C	A	TAGGAGGCG

halfsite 5′-GTCAT was represented four times more frequently than expected at random, indicating some specific interaction with DNA (Table 1), however palindromic or semi-palindromic sites were not abundant. It is interesting to note that similar halfsites have been selected using *C.elegans* Skn-1, a transcription factor with 65% similarity to TCF11 in the DNA-binding region that lacks a leucine zipper and does not dimerise (4). It appears, therefore, given the absence of palindromic sites, that MBP-TCF11-A *in vitro* does not interact with DNA as a bipartite structure formed by conventional dimerisation through the bZIP domain, but that a homomeric complex is capable of interacting with AP1 halfsites. The fact that we did not select any tandem repeated sites, suggests that at least *in vitro*, MBP-TCF11-A alone has no strong preference for repeated motifs.

Binding sites selected by TCF11:MafG

The MBP-TCF11-A:MBP-MafG heterodimer migrates slightly slower than the MBP-MafG homodimer (Fig. 1B) on a polyacrylamide gel. It was, therefore, possible to isolate sequences bound to either of the complexes. When the two proteins are mixed they preferentially form heterodimers, the preference being obvious by the second round of selection (not shown). Three rounds of selection generated an oligonucleotide pool that bound strongly to the proteins (Fig. 1B). Of the 36 DNA sequences that were inspected (Table 2A), 31 sites (86%) contained a sequence consistent with the deduced TCF11:MafG core consensus sequence 5′-TGCTgaGTCAT-3′ (Table 2B and D), showing that binding of the heterodimer is highly specific.

Of the TCF11:MafG-selected oligonucleotides, 55% comprised nucleotides in the constant (primer) region. The primer sequences were utilised only as Maf halfsites indicating that the MafG half of the heterodimer is more promiscuous than TCF11 in its sequence specificity (see Materials and Methods for further details). The 5′-TCA triplet at positions 5–7 was 100% conserved in all oligonucleotides. This shows that heterodimerising with a small Maf protein renders TCF11 more stringent in AP1-site binding preference. Maf binding sites have previously been divided into two groups; cyclic AMP responsive element (CRE)-type (TGACGTCA) and TPA-responsive element (TRE)-type (TGAGTCA) (24). In the current experiments all TCF11:MafG-selected sites showed TRE-like halfsite spacing. This is in contrast to the previously reported EMSA competition experiment in K562 cell nuclear extracts which showed that TCF11 p47/49 also bound specifically to an NF-E2-site with a CRE-like halfsite (11). The two methods used are likely to differ in sensitivity but the observation could also be explained by TCF11 forming heterodimers with bZIP-factors other than Maf proteins in these cell extracts, or alternatively, the endogenous TCF11:Maf heterodimer may display sequence-specificity different from that of the corresponding heterodimer formed *in vitro*.

The consensus heme-responsive element (30) contains a T at position +11 (Table 2C) which is only partially conserved and which can be substituted to an A without loss of heme responsivity in mouse L929 fibroblasts. The same position has 45% T and 35% A in the heterodimer selected TCF11 halfsites (Table 2C). This suggests that TCF11 prefers an A/T pair located at this position which was shown to be protected in an α-globin NF-E2-site in K562 cells (38). The CNC motif, being located immediately N-terminal to the basic DNA-binding domain, could therefore have a role in contacting DNA. Targeted point-mutations in this region of TCF11 may reveal its role in stabilising protein binding to the TCF11:Maf-site.

Our binding site-selected consensus site is also interesting in that 15 of the 20 positions (75%) deviating from the TCF11:MafG consensus nucleotides selected within the degenerate library are located at positions 1, 2 and 3 (Table 2). This 'hot-spot' may therefore be a central sub-element inside the 11 bp TCF11:Maf-site which is used to discriminate it from a strong AP1 (Jun/Fos) element. This possibility is now being tested.

Table 3. The occurrence of the selected TCF11:MafG binding-site in association with various genes. The TCF11:MafG consensus binding sequence (5′-TGCTgaGTCAT-3′) was used in a computer search of the EMBL database and of the eukaryotic promoter database (EPD)

EMBL-HUMAN			**Sequence**	
human		Alpha-globin gene cluster	GACTGCTGAGTCATCCT	5' region
human		AMBP gene	TTCTGCTGAGTCATGCC	regulatory region
human		Apolipopr. (a) -related gene B	ATTTGCTGAGTCAGTAT	5' region
human		Beta globin region	CAATGCTGAGTCATGAT	locus control region
human		Cdc25B mRNA	CCCTGCTGAGTCATCTG	60 nt. from 3'
human		CpG island, clone 151a5	AAATGCTGAGTCAAGTA	CpG island
human		Dopamine D1A receptor gene	TGCTGCTGAGTCAATGC	5' region
human		Erythrocyte band 7 int. memb. mRNA	GGTTGCTGAGTCAGCGT	440 nt from 5'
human		Ferrochelatase gene	ATTTGCTGAGTCATGGC	5' region
human		GST gene	CAGTGCTGAGTCAGCCG	5' region
human		Hp2 gene 3' end and Hpr* gene 5' end	CATTGCTGAGTCAATCA	3' and 5' region
human		ITI gene	TTCTGCTGAGTCATGCC	exon 1
human		Lymphoma proprotein conv. mRNA	CTTTGCTGAGTCAACAC	nt. 30
human		Oligodendrocyte myelin glycoprotein	AAATGCTGAGTCAGATC	exon 1-2
human		Phenol sulfotransferase gene (STP)	TTTTGCTGAGTCATCAG	promoter
human		Phenol sulfotransferase gene (STP2)	TTTTGCTGAGTCATCAC	promoter
human		Ribonuc. diP reductase M1	AAATGCTGAGTCAGGGT	intron 1
human		Salivary statherin gene	ACTTGCTGAGTCAGAAT	5' region
human		Tissue-type plasminogen activator	TTATGCTGAGTCATACC	enhancer
EMBL-RODENT				
rat		Beta-alanine synthase mRNA	ATTTGCTGAGTCAGCAG	nt. 450
mouse		Beta-globin HS2	CAGTGCTGAGTCATGCT	LCR
mouse		Beta-globin HS2	TCATGCTGAGTCATGCT	LCR
mouse		Beta-globin HS3	GTTTGCTGAGTCAAAGT	LCR
rat		Brain glucose transporter	GGCTGCTGAGTCATACT	exon 2
mouse		CRG-2 (IP-10) gene	CCTTGCTGAGTCATCTC	5' region
rat		Cytochrome P450 2B3 gene	GGTTGCTGAGTCATTCC	exon 9
mouse		Gene for CD7 antigen	GCCTGCTGAGTCATGAC	exon 1
mouse		Heme oxygenase	TTTTGCTGAGTCACCCT	5' enhancer
rat		Heparin-binding growth factor 1, mRNA	TGCTGCTGAGTCATGGC	nt. 260
mouse		Ig active H-chain deletion mutant IF2	AGTTGCTGAGTCAAAAT	J-C intron
mouse		Lymphotoxin-beta gene	CTTTGCTGAGTCACATG	5' region
rat		Menkes protein (MNK) mRNA	GATTGCTGAGTCACTGG	nt. 850
mouse		Testosterone 16-alpha-hydroxylase gene	CATTGCTGAGTCACTCA	intron 1
EPD-PROMOTERS				
human	EP026018	Aldolase A E2P2	TCCGGCTGAGTCACGAT	promoter
Xl	EP007078	Aldolase A E3P2	TCTGGCTGAGTCACGGC	promoter
Gg	EP016048	b' I-globin larva	CCATGCTGAGTCACACC	promoter
Mm	EP007118	Ferritin H	AGGTGCTGAGTCACGGC	promoter
human	EP017081	HSP 70K	GGTTGCTGAGTTAGTCA	promoter
Mm	EP027013	Ig k' MOPC41	CCATGCTGAGTCCTGAC	promoter
human	EP030062	Ig k' T	CCATGCTGAGTCCTGGC	promoter
human	EP026008	IL-6 (BSF-2)	AAGTGCTGAGTCACTAA	promoter
Gg	EP041006	MCK	AAGTGCTGAGTCTAGCT	promoter
Rn	EP016034	MT-IB	CCCTGCTGAGTCTGCAC	promoter
human	EP017063	Myosin H sk-m.	GTGTGCTGAGTCTGTCT	promoter
Xl	EP014069	Nucleolar Ag p120	GAGTGCTGAGTGAGTTG	promoter
Mm	EP011154	PBGD erythroid	CTGTGCTGAGTCACTGG	promoter

Potential TCF11/Maf target genes

It was immediately apparent that the binding site selected for the TCF11:MafG heterodimer is identical to the NF-E2 site that mediates erythroid specific gene expression of, for example, the β-globin gene cluster and the PBGD gene (5,6,39,40). We carried out a search of the sequence database to gain an idea of the overall distribution of the selected sites and found that a number of genes contain potential binding-sites for the heterodimer in their regulatory regions (Table 3). Many of these putative target genes can be classified into genes involved in haemoglobin and iron metabolism and genes important in cellular detoxification. Transcriptional responses to antioxidants and several xenobiotics act through AREs (5′-GCnnnGTCA-3′) (31) and AREs from heme oxygenase 1, NAD(P)H:quinone oxidoreductase (NQO_1), glutathione S-transferase and phenol sulfotransferase 1 and 2 show similarly positioned nucleotides that together define a consensus site identical to that of the TCF11:small Maf binding-site (Table 3). Indeed, TCF11 has been shown to positively regulate chloramphenicol acetyl-transferase (CAT) gene expression when linked to an ARE derived from the NQO_1 gene (41).

TCF11 activity through the NF-E2 site

Based on the similarity between TCF11 and p45 NF-E2, it was previously suggested that TCF11 may act through this site and such binding has been demonstrated *in vitro* (9,11). The results of the binding-site selection assay further underlined this possibility. We analysed the activity of TCF11 in a transient transfection assay where the firefly luciferase gene, under the control of the PBGD erythroid-specific promoter (39,40) (chosen for the simple context in which a single NF-E2 site is presented), acts as a reporter driven by TCF11 produced from a full length TCF11 cDNA. Correct expression and nuclear localisation of the TCF11 protein were confirmed by immunofluoresence staining using a polyclonal TCF11 antibody (11, data not shown). The activity of TCF11 was assayed in a number of different cell lines and was

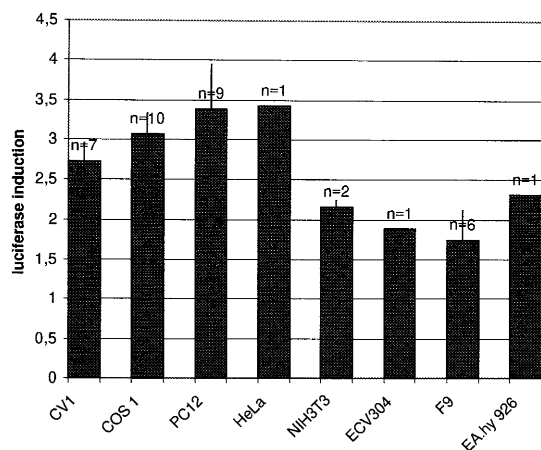

Figure 2. TCF11 transactivates expression through the PBGD erythroid-specific promoter in a variety of cell lines as indicated The values for luciferase induction show fold induction over background activity (in the absence of TCF11 expression vector) set to 1. Each value for luciferase activity was normalised to the activity of an internal control (CAT). The induction shown is the average over a number of experiments (n) and the error bars reflect standard deviation for each mean value.

found to transactivate expression in COS 1 and CV 1 monkey kidney cells, in human HeLa cells and endothelial cells (EA.hy 926 and ECV304), murine NIH3T3 fibroblasts and rat PC12 cells (Fig. 2). The level of transactivation was low and somewhat variable from experiment to experiment but consistently positive against a high background level of activity in the absence of transfected TCF11 (not shown). Murine F9 embryonal carcinoma cells showed the highest variability and in some experiments no transactivation was observed. To demonstrate that the NF-E2 site within the PBGD promoter mediated the transactivation, a number of deletions of the promoter region were analysed (Fig. 3A), this showed that transactivation was lost when the NF-E2 site was deleted (Fig. 3B). The background levels also dropped 3–10-fold (not shown). To demonstrate the role of the NF-E2 site more directly, its sequence was mutated within the context of the shortest active promoter (PBGD3.2Luc) (Fig. 3A). The two mutations that were assayed both reduced or abolished transactivation (Fig. 3C). This implies that TCF11 can bind to and transactivate expression through the NF-E2 site.

The activity of TCF11 in the presence of MafG

DNA binding assays showed that TCF11 preferentially forms heterodimers with small Maf proteins (including MafG) *in vitro* (11). The binding site selection studies reported here indicate that the preferred DNA binding sequence for the TCF11:MafG heterodimer is a perfect NF-E2 site. Furthermore, these experiments showed that TCF11 alone did not form simple homodimeric binding units and while it did bind DNA with a sequence specificity, this specificity was limited compared to the very clear preference shown by the heterodimer. We therefore wished to compare the activities of TCF11 alone and TCF11 co-expressed with MafG. This comparison was performed in COS 1 cells. As described above, TCF11 transactivated expression and the level of transactivation was found to be dependent on the amount of TCF11 transfected (Fig. 4B). In such a transfection assay, which involves transcription through an NF-E2 site that harbours a core AP1 site, background expression of the reporter gene is high. As has been found previously (16,21) the expression of MafG alone efficiently repressed this background level (Fig. 4). This is not surprising since small Maf proteins do not contain a known transactivation domain and so binding of the Maf homodimer may block access of endogenous factors (possibly AP1) responsible for the background activity. There is also evidence that small Maf proteins may block endogenous activation indirectly (29). Expression of a mutant form of MafG, which harbours a single amino acid change within the leucine zipper and so cannot dimerise (MafGΔ L2PM4P; 21) showed no such repression. Surprisingly, when TCF11 and MafG were co-expressed in the same cell, MafG blocked the weak transactivation observed with TCF11. Different relative amounts of the expression vectors for TCF11 and MafG were transfected in an effort to titrate the interaction. It was found that even at low relative amounts of MafG to TCF11 (2:7 µg of vector DNA) a significant drop in transactivation was already obvious (Fig. 4A). Even lower relative amounts of MafG (down to 10 ng of vector DNA) showed no evidence of co-operative transactivation with TCF11 (Fig. 4B). It is clear that the presence of MafG interferes with TCF11 mediated transactivation but it is not known whether this is due to the TCF11:MafG heterodimer lacking transactivation ability or whether MafG preferentially forms homodimers in this cellular context which compete for binding site access. The first possibility is suggested by the fact that TCF11 clearly and preferentially forms heterodimers with MafG *in vitro* (11; Fig. 1B).

MafG interferes with expression from the SV-40 promoter

In this series of experiments it was also observed that MafG can repress expression from the SV-40 promoter. This is important to note for the design of future transfection experiments involving small Maf proteins since vectors using this promoter cannot, therefore, be used as independent internal controls. In early experiments the plasmid pRSV-CAT, where CAT gene expression is driven from the constitutive SV-40 promoter, was used as an internal control to correct for differences in transfection efficiency. It was found that whenever the MafG expression vector was used in these experiments the level of CAT activity was greatly reduced (in the region of 10-fold, results not shown). Therefore, if the luciferase activity were corrected for CAT activity, the results were highly variable, inconsistent and uninterpretable. From these experiments the luciferase activity not corrected for CAT showed consistently the same trend observed in later experiments using an alternative internal control (pEFβ-gal). The internal control used in subsequent experiments expresses β-galactosidase under the control of the elongation factor Iα promoter and did not appear to be influenced by MafG (29,42).

DISCUSSION

The PCR-assisted approach of cloning transcription factor binding-sites following *in vitro* selection, is a powerful tool in the identification of regulatory sequences (24,36,37). We have shown that the TCF11:MafG heterodimer shows a clear preference for a site identical to a number of known regulatory elements including the NF-E2 site, the ARE and the HRE. The sequences of the selected oligonucleotides show a number of interesting features that are discussed in the Results section. It is interesting that TCF11 alone does not form simple homodimers *in vitro* but apparently

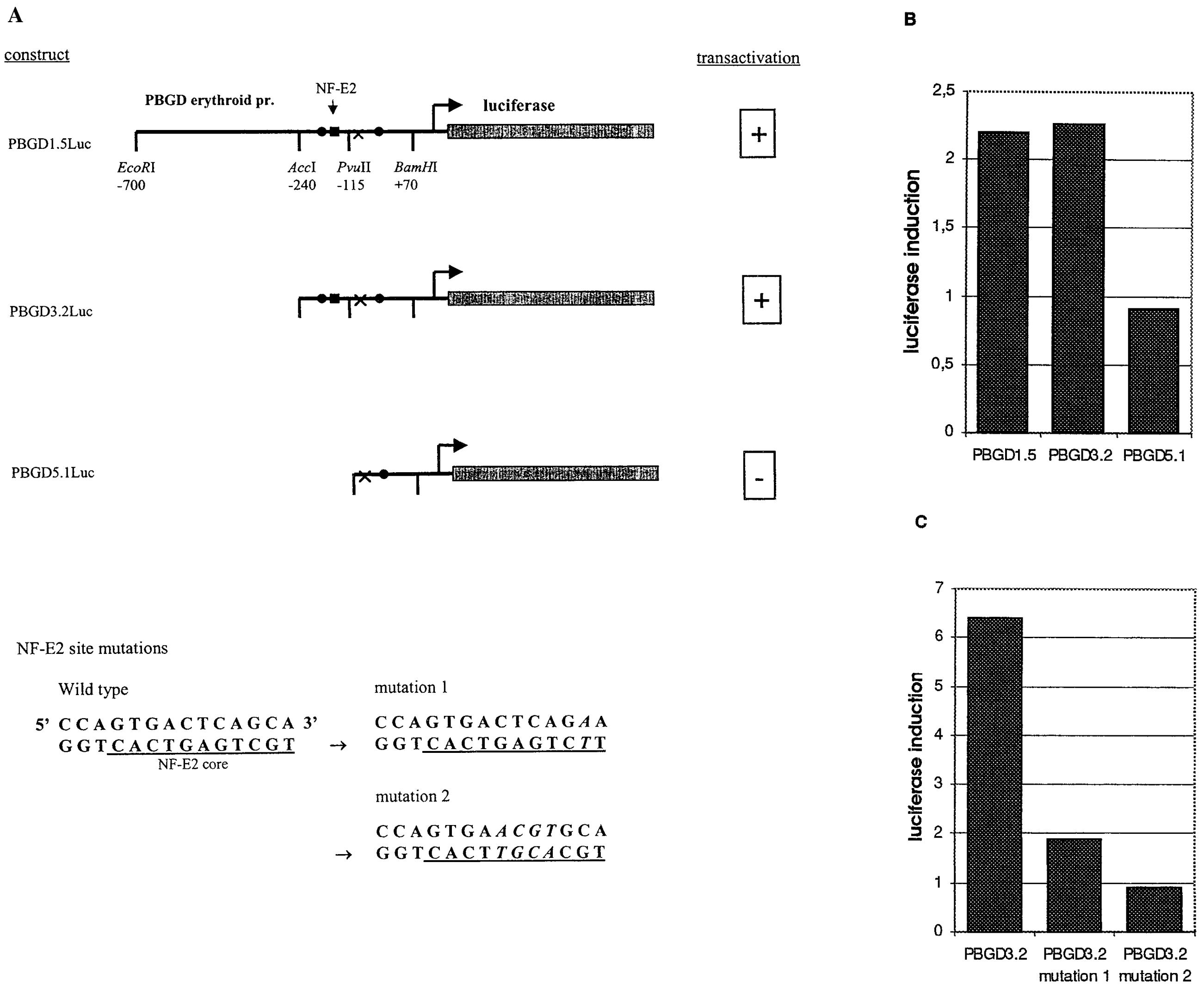

Figure 3. The NF-E2 site mediates transactivation by TCF11. (**A**) Reporter constructs used in the transfection experiments presented in (B) where the luciferase gene is under the control of various deletions or mutations (in italics) of the PBGD erythroid-specific promoter. The restriction sites show relative positions within the promoter (36,37). ■ represents the NF-E2 site, ● represents NF-E1 sites and × represents the CACC box. The NF-E2 core site is underlined. (**B**) Luciferase induction through the different reporter constructs shown in (A) upon co-transfection with TCF11. Note that both constructs PBGD1.5 and PBGD3.2 are transactivated to the same level whereas PBGD5.1 is not responsive to TCF11. (**C**) Comparison of transactivation through the wild-type PBGD (PBGD3.2) promoter and two versions where the NF-E2 site has been mutated (as shown in A). A typical example of a single experiment is shown in (B) and (C). The experiments were repeated at least four times. Each luciferase value shows fold induction over background and has been normalised to CAT activity.

binds as a multimeric complex to AP1 half sites, as indicated by its retarded electrophoretic mobility. The reduced specificity seen in the selection of TCF11 sites, together with the preference shown for the formation of heterodimers when TCF11 and small Maf proteins are co-expressed *in vitro*, indicate that a functional form of TCF11 is as a heterodimer with small Maf and/or, perhaps, with other unidentified bZIP partners. However, our transfection experiments show that TCF11 can transactivate expression when transfected alone and the fact that a dose dependent increase is observed would suggest that this effect is not dependent on TCF11 heterodimerisation with limiting amounts of endogenous factors.

The transactivation observed with TCF11 alone is inhibited by co-expression of MafG. Since these factors are both widely expressed and are likely to be co-expressed in a variety of cell types *in vivo*, and since they preferentially form heterodimers *in vitro*, it seems likely that the factors may commonly exist as an inactive or repressive heterodimeric form. Positive regulation of the target sites to which they bind may depend on heterodimerisation of these factors with alternative partners, perhaps with more tissue restricted expression. This study has provided us with a system in which to test alternative TCF11 partners. The other small Maf proteins and CNC-bZIP family members are candidates but the identification of new, tissue restricted partners for TCF11 is also pertinent.

The absence of co-activation by TCF11 and MafG shows that TCF11 acts differently to other CNC-bZIP family members in similar transfection assays. It has been observed that while all three small Maf proteins (-K, -F and -G) repress expression through the NF-E2 site in NIH3T3 cells or quail fibroblasts,

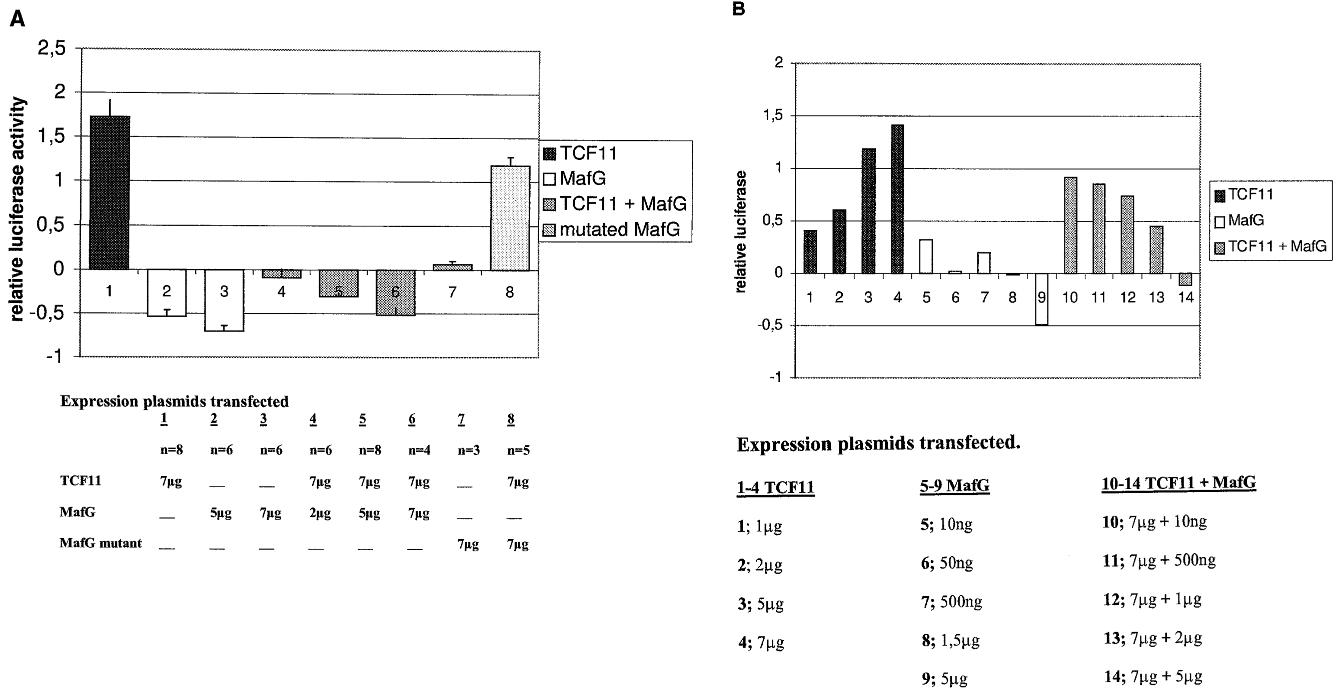

Figure 4. MafG interferes with TCF11 transactivation in COS 1 cells. Luciferase activity relative to background levels (set to 0) is represented. All luciferase values were normalised to β-galactosidase activity, expressed from an internal control plasmid. (**A**) Mean values from a number of experiments (represented by n). Error bars represent standard deviation of the mean. Each point on the x-axis represents a different combination of expression plasmids as shown under the graph. Note the dose dependent interference with TCF11 transactivation by MafG (4–6). (**B**) The results of a single experiment (mean of duplicate transfections) including a greater range of transfected expression vectors than shown in (A). Note that the transactivation by TCF11 (1–4), the repression by MafG (5–9) and the interference by MafG of TCF11 transactivation (10–14) are dose dependent.

co-expression with NF-E2 p45 interferes with Maf repression and leads to co-transactivation of the reporter (16,21,26). The NF-E2 related chicken gene, Ech (most similar to Nrf2), transactivates very efficiently in quail fibroblast cells in the presence of MafK, also overcoming repression by MafK alone (13). It is important to note that homology between the CNC-bZIP family members is largely restricted to the CNC-bZIP region involved in DNA-binding and heterodimerisation. Therefore functions mediated by other domains, hypothetically co-activator interactions and contact with the transcriptional complex, may differ between, for example, p45 NF-E2:small Maf and TCF11:small Maf heterodimers. This suggests that different CNC-bZIP factors, in partnership with different (or perhaps the same) small Maf proteins can act differentially through the same or similar regulatory elements indicating a complex network of competitive interactions when these factors are co-expressed.

Two recent articles report the cloning of the human homologue of MafG and, in contrast with our results, claim that a very small relative amount of MafG co-expressed with a TCF11 isoform (Nrf1) leads to co-operative transactivation (43,44). However, the data presented show that the slight increase in reporter activity in the presence of the two expression vectors falls well within (44), or just outside (43), the range of errors for the experiment. The slight increase reported in the latter case is detected in the range of 2 µg Nrf1 expression vector: 1–10 ng MafG expression vector, the effect being lost at 100 ng MafG. It is difficult to understand why such a low relative amount of MafG would have a positive effect that is lost so rapidly. We have attempted to repeat these observations in our assay system and see no such effect (Fig. 4), especially given the large variability inherent in this kind of experiment. It is possible that this discrepancy can be explained by differences in our assay systems although Toki *et al.*, used the same cell line (COS 1) for their transfections. They have, however, expressed the human MafG cDNA whereas we used the chicken homologue (94% identity at the protein level).

TCF11 is a transcription factor that has been implicated in the regulation of erythroid-specific expression because of its ability to bind the NF-E2 site, but it is not erythroid specific, showing widespread expression. No target genes for TCF11 have as yet been identified. This study has shown that potential targets, based on the presence of optimal binding-sites, fall into a number of groups of genes (Table 3) which are co-regulated in response to a specific signal; antioxidant response, heme biosynthesis and erythroid differentiation, implicating TCF11 in specific biological processes. The role of TCF11 in these processes can now be tested. It is not known how many of these potential TCF11 binding sites represent real targets of TCF11 but it is interesting to speculate that the ubiquitously expressed proteins TCF11, Maf and AP1 all participate in gene regulation through these sites and thereby are connected in a network regulating a broad range of genes. The number of sites that represent real targets may be limited by competition from alternative factors or physical unavailability of sites due to protein binding to flanking sequences. In the search for TCF11 target genes, it has become important to also consider genes involved in early embryonic development, specifically in the gastrulation process, since the work of Farmer *et al.* (45) has shown that mice lacking TCF11 (LCR-F1) are blocked during the early steps of gastrulation. The identification of functional target genes

for TCF11 is now of primary importance for the further understanding of TCF11 activity and function *in vivo*.

ACKNOWLEDGEMENTS

This work was supported by grants to A.-B. K. from the Research Council of Norway and the Norwegian Cancer Society and to P. M. from the Norwegian Cancer Society. Ø. J. is a Research Fellow of the Research Council of Norway and the University of Oslo. The MBP-MafG construct, MafG cDNA clones and pEFβgal were kindly provided by Dr Makoto Nishizawa, University of California at San Diego.

REFERENCES

1 Mohler, J., Vani, K., Leung, S. and Epstein, A. (1991) *Mech. Dev.*, **34**, 3–10.
2 Mohler, J., Mahaffey, J.W., Deutsch, E. and Vani, K. (1995) *Development*, **121**, 237–247.
3 Bowerman, B., Eaton, B.A. and Priess, J.R. (1992) *Cell*, **68**, 1061–1075.
4 Blackwell, T.K., Bowerman, B., Priess, J.R. and Weintraub, H. (1994) *Science*, **266**, 621–628.
5 Talbot, D., Philipsen, S., Fraser, P. and Grosveld, F. (1990) *EMBO J.*, **9**, 2169–2178.
6 Andrews, N.C., Erdjument-Bromage, H., Davidson, M.B., Tempst, P. and Orkin, S.H. (1993) *Nature*, **362**, 722–728.
7 Chan, J.Y., Han, X.-L. and Kan, Y.W. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 11371–11375.
8 Luna, L., Johnsen, Ø., Skartlien, A.H., Pedeutour, F., Turc-Carel, C., Prydz, H. and Kolstø, A.B. (1994) *Genomics*, **22**, 553–562.
9 Caterina, J.J., Donze, D., Sun, C.-W., Ciavatta, D.J. and Townes, T.M. (1994*) Nucleic Acids Res.*, **22**, 2383–2391.
10 Luna, L., Skammelsrud, N., Johnsen, Ø., Abel, K.J., Weber, B.L., Prydz, H. and Kolstø, A.B. (1995) *Genomics*, **27**, 237–244.
11 Johnsen, Ø., Skammelsrud. N., Luna, L, Nishizawa, M., Prydz, H. and Kolstø, A.-B. (1996) *Nucleic Acids Res.*, **24**, 4289–4297.
12 Moi, P., Chan, K., Asunis, I., Cao, A. and Kan, Y.W. (1994) *Proc. Natl. Acad. Sci. USA.*, **91**, 9926–9930.
13 Itoh, K., Igarashi, K., Hayashi, N., Nishizawa, M. and Yamamoto, M. (1995) *Mol. Cell. Biol.*, **15**, 4184–4193.
14 Oyake, T., Itoh, K., Motohashi, H., Hayashi, N., Hoshino, H., Nishizawa, M., Yamamoto, M. and Igarashi, K. (1996) *Mol. Cell Biol.*, **16**, 6083–6095.
15 McKie, J., Johnstone, K., Mattei, M-G. and Scrambler, P. (1995) *Genomics*, **25**, 716–719.
16 Igarashi, K. Kataoka, K., Itoh, K., Hayashi, N., Nishizawa, M. and Yamamoto, M. (1994) *Nature*, **367**, 568–572.
17 Nishizawa, M., Kataoka, K., Goto, N., Fujiwara, K.T. and Kawai, S. (1989*) Proc. Natl. Acad. Sci. USA*, **86**, 7711–7715.
18 Pouponnot, C., Nishizawa, M., Calothy, G. and Pierani, A. (1995) *Mol. Cell. Biol.*, **15**, 5563–5575.
19 Cordes, S.P. and Barsh, G.S. (1994) *Cell*, **79**, 1025–1034.
20 Kurschner, C. and Morgan, J.I. (1995) *Mol. Cell. Biol.*, **15**, 246–254.
21 Kataoka, K., Igarashi, K., Itoh, K., Fujiwara, K.T., Noda, M., Yamamoto, M. and Nishizawa, M. (1995) *Mol. Cell. Biol.*, **15**, 2180–2190.
22 Ney, P.A., Andrews, N.C., Jane, S.M., Safer, B., Purucker, M.E., Weremowicz, S., Morton, C.C., Goff, S.C., Orkin, S.H. and Nienhuis, A.W. (1993) *Mol. Cell. Biol.*, **13**, 5604–5612.
23 Fujiwara, K., Kataoka, T.K. and Nishizawa, M. (1993) *Oncogene*, **8**, 2371–2380.
24 Kataoka, K., Noda, M. and Nishizawa, M. (1994) *Mol. Cell. Biol.*, **14**, 700–712.
25 Kerppola, T.K. and Curran, T. (1994) *Oncogene*, **9**, 3149–3158.
26 Igarashi, K., Itoh, K., Motohashi, H., Hayashi, N., Matuzaki, Y., Nakauchi, H., Nishizawa, M. and Yamamoto, M. (1995) *J. Biol. Chem.*, **270**, 7615–7624.
27 Motohashi, H., Igarashi, K., Ohtani, H., Nishizawa, M., Engel, J.D. and Yamamoto, M. (1996) *Genes Cells*, **1**, 223–238.
28 Walters, M., Andrews, N.C., Magis, W. and Martin, D.I.K. (1996) *Exp. Hematology*, **24**, 445–452.
29 Kataoka, K., Noda, M. and Nishizawa, M. (1996) *Oncogene*, **12**, 53–62.
30 Inamdar, N.M., Ahn, Y.I. and Alam, J. (1996) *Biochem. Biophys. Res. Commun.*, **221**, 570–576.
31 Rushmore, T.H., Morton, M.R. and Pickett, C.B. (1991) *J. Biol. Chem.*, **266**, 11532–11639.
32 Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
33 Beaumont, C., Porcher, C., Picat, C., Nordmann, Y. and Grandchamp, B. (1989*) J. Biol. Chem.*, **264,** 14829–14834.
34 Chen, C. and Okayama, H. (1987*) Mol. Cell Biol*., **7**, 2745–2752
35 Seed, B. and Sheen, J-Y (1988) *Gene*, **67**, 271–277
36 Pollock, R. and Treisman, R. (1990) *Nucleic Acids Res*., **18**, 6197–6204.
37 Wright, W.E., Binder, M. and Funk, D. (1991) *Mol. Cell Biol*., **11**, 4104–4110.
38 Strauss, E.C. and Orkin, S.H. (1997) *A Companion to Methods in Enzymology*, **11**, 164–170.
39 Mignotte, V., Wall, L., deBoer, E., Grosveld, F. and Romeo, P.H. (1989) *Nucleic Acids Res.*, **17**, 37–54.
40 Mignotte, V., Eleouet, J.F., Raich, N. and Romeo, P.-H. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6548–6552.
41 Venugopal, R. and Jaiswal, A.K (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 14960–14965.
42 Mizushima, S. and Nagata, S. (1990) *Nucleic Acids Res*., **18,** 5322 .
43 Toki, T., Itoh, J., Kitazawa, J., Arai, K., Hatakeyama, K., Akasaka, J., Igarashi, K., Nomura, N., Yokoyama, K., Yamamoto, M. and Ito, E. (1997) *Oncogene*, **14**, 1901–1910.
44 Marini, M.G., Chan, K., Casula, L., Kan, Y.W., Cao, A. and Moi, P. (1997) *J. Biol. Chem*., **272**, 16490–16497.
45 Farmer, S.C., Sun, C.-W., Winnier, G.E., Hogan, B.L.M. and Townes, T.M. (1997) *Genes Dev*., **11**, 786–798.