

# Small Maf proteins interact with the human transcription factor TCF11/Nrf1/LCR-F1

Øyvind Johnsen, Nina Skammelsrud, Luisa Luna, Makoto Nishizawa<sup>1</sup>, Hans Prydz and Anne-Brit Kolstø\*

Biotechnology Centre of Oslo, University of Oslo, PO Box 1125 Blindern N-0316 Oslo, Norway and <sup>1</sup>Department of Molecular Oncology, Faculty of Medicine, Kyoto University, Yoshida-Konoecho, Sakyo-ku, Kyoto 606, Japan

Received June 20, 1996; Revised and Accepted September 11, 1996

## ABSTRACT

The human *TCF11* gene encodes a ubiquitously expressed bZIP transcription factor of the cap n' collar (CNC) domain family. It has a high sequence similarity to the erythroid-specific bZIP factor p45 NF-E2 in the CNC domain, which is involved in DNA binding. LCR-F1, a *TCF11* isoform, is a more potent transcriptional activator than p45 NF-E2 in erythroid cells. We show here that the *TCF11* protein interacts to form heterodimers with small Maf proteins, previously shown to dimerize with p45 NF-E2, ECH and Fos. Such heterodimerization significantly alters the DNA binding characteristics of *TCF11*. While *TCF11* alone binds *in vitro* to the tandem NF-E2 site derived from 5' DNase hypersensitive site 2 in the  $\beta$ -globin locus control region and to the single NF-E2 site in the porphobilinogen deaminase gene promoter, stronger binding is detected in the presence of small Maf proteins. Using antibodies, *TCF11* isoforms bound to the single NF-E2 site were detected in K562 erythroid cell nuclear extracts. These findings place *TCF11* as a good candidate for the proposed widely expressed factor(s) known to interact with small Maf proteins and bind NF-E2 sites in a sequence-specific manner resembling NF-E2.

## INTRODUCTION

The erythroid-specific transcription factor NF-E2 (1,2) activates transcription from a discrete set of AP1-related sites. Such NF-E2 sites have been identified in promoters and enhancers of several erythroid-specific genes, including the 5' DNase hypersensitive site 2 (5'HS2) in the  $\beta$ -globin locus control region (LCR) (3–5) and the porphobilinogen deaminase (PBGD) erythroid-specific promoter (1,6). Purification and molecular cloning of NF-E2 revealed an erythroid-specific 45 kDa bZIP protein which dimerized with a ubiquitous 18 kDa partner (2). This partner, necessary for DNA binding activity and transcriptional activation, was shown to be a member of the small Maf proto-oncprotein family, presumably MafK (7–9). The small Maf

family proteins MafF, MafG and MafK are bZIP transcription factors which can form homo- and heterodimers with each other, in addition to heterodimers with Fos (10–13) and the cap n' collar (CNC) domain family members p45 NF-E2 (7) and ECH (14). Maf homodimers or intra-subfamily heterodimers recognize 13 or 14 bp elements (Maf-responsive elements; MAREs) of the 12-*O*-tetradecanoylphorbol-13-acetate-responsive element (TRE) or the cyclic AMP-responsive element (CRE) type [T-MARE (TGCTGA(C/G)-TCAGCA) and C-MARE (TGCTGACGTCAGCA) respectively], with strongest requirements for the first triplets in the half-sites (13,15,16). The NF-E2 site, (T/C)GCTGA(C/G)TCA(C/T), is composed of a T-MARE half-site [(T/C)GCTGA(C/G)] and an AP1 half-site [(C/G)TCA(T/C)], suggesting that native NF-E2 (p18:p45) binds to this site in a fixed orientation, with Maf recognizing the long half-site (7,8). A G→T mutation at –2 relative to the AP1 core, which affects NF-E2 but not AP1 binding, results in loss of 5'HS2 enhancer activity (3) and of PBGD erythroid promoter inducibility (6). The Maf recognition of this extended half-site requires an ancillary DNA binding region on the N-terminal side of the Maf bZIP domain (16).

Small Maf proteins lack a transactivation domain and repress the strong NF-E2 site-driven transcription observed in transient assays, probably by competing with endogenous activator(s) for binding to this site (7,13). Co-expression of small Maf proteins and CNC domain family members, however, transactivates NF-E2 site-driven transcription by the formation of heterodimers (7,13,14). Heterodimers of Fos and the small Maf proteins also bind to NF-E2 sites, but such heterodimers act as trans-suppressors, like the small Maf homodimers (13). It therefore appears that the expression of genes containing NF-E2-like regulatory sequences, like  $\beta$ -globin and erythroid expressed PBGD, depends on the effective concentrations of Maf proteins and interacting partners. Since small Maf proteins are widely expressed (8,11), they could in theory act as regulators in non-erythroid tissues as homodimers and heterodimers with other proteins yet to be characterized. Strong endogenous binding activity and activation through NF-E2 sites have been observed in cells that express neither NF-E2 nor AP1 and the protein(s) which represents this activity has been shown to interact with the small Maf proteins (17,18). This suggests that small Maf proteins and associated partners may

\*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

regulate other genes (15) through NF-E2-like sites in non-erythroid tissue as well.

Molecular cloning of cDNA (19) and genomic DNA (20) for the bZIP transcription factor TCF11 revealed a gene consisting of seven exons spanning 15 kb with a number of alternative splicing and polyadenylation variants and an open reading frame of 772 amino acids. Nrf1 (21) and LCR-F1 (22), which are TCF11 isoforms transcribed from the *TCF11* gene, were cloned from cDNA libraries made from the erythroleukemia cell line K562 using functional screening assays. Translation of Nrf1 *in vitro* resulted in two isoforms of 110 and 65 kDa (21). The short isoform is probably translated from an internal start site(s) (19,21). Analysis of the TCF11 amino acid sequence reveals a C-terminal domain which is homologous to the *Drosophila* segmentation gene product CNC (23), the erythroid-specific transcription factor p45 NF-E2 (2,24), the ubiquitously expressed Nrf2 (25) and the more tissue restricted ECH (14). This region, the CNC domain, includes the basic DNA binding region and suggests that TCF11 belongs to a distinct subfamily of bZIP factors recognizing related DNA elements.

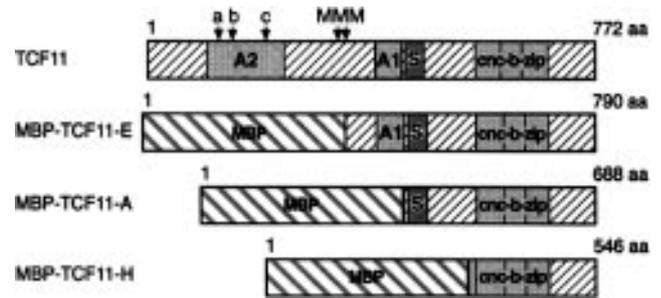
In this study, we investigated whether TCF11 could interact with Maf proteins and bind to NF-E2 sites. Our results suggest that TCF11, like p45 NF-E2, ECH and Fos, competes for small Maf proteins to form heterodimers with higher affinity than either homodimer for binding to NF-E2 sites. In combination with the strong transactivation observed by LCR-F1 in K562 cells (22), this suggests that TCF11/Nrf1/LCR-F1 has a functional role as a transcriptional activator by binding to NF-E2 sites in erythroid tissue. Since both TCF11 and small Maf proteins are ubiquitously expressed, TCF11-Maf heterodimers may also regulate other genes through MARE/NF-E2-like sites in other tissues.

## MATERIALS AND METHODS

Restriction enzymes were obtained from New England Biolabs (NEB) and radioactive compounds from Amersham. Oligonucleotides were synthesized at the Biotechnology Centre, Oslo, by Dr Eshrat Babaie. Anti-p45 NF-E2 and anti-Nrf2 polyclonal antibodies were obtained from Santa Cruz Biotechnology. Standard methods in molecular biology were used (26).

### Construction of pMALc plasmids and isolation of fusion proteins

The plasmid vector pMALc (NEB) encoding the maltose binding protein (MBP) was used to generate the three C-terminal MBP-TCF11 constructs -E, -A and -H (Fig. 1). The TCF11 cDNA clone pZcEA6, which contains the 1720-4200 TCF11 fragment in pBluescript SK (19) was digested with *Sma*I. *Sal*I linkers were ligated to the resulting *TCF11* fragment and the fragment then digested with *Eco*RI and *Sal*I. This *Eco*RI-*Sal*I fragment was ligated to the corresponding sites in the linearized pMALc vector to generate the final construct MBP-*TCF11*-E, encoding a 790 amino acid fusion protein where 398 amino acids are the C-terminal part of TCF11. MBP-*TCF11*-A was constructed by digestion of pZcEA6 with *Apo*I followed by ligation of the 2.1 kb 3' TCF11 fragment into the *Eco*RI site of pMALc. Translation of MBP-*TCF11*-A results in a 688 amino acid fusion protein including 300 amino acids of the TCF11 C-terminal sequence. MBP-*TCF11*-H was constructed by digestion of pZcEA6 with *Bsi*HKAI and ligation of the 1.1 kb 3' fragment into



**Figure 1.** Schematic representation of TCF11 and TCF11-maltose binding protein (MBP) fusion proteins. Acidic transactivation regions are indicated by A (19,22). The most C-terminal acidic region (A1) has been characterized (22), whereas the N-terminus (A2) is a putative transactivation domain (19). A serine-rich region is indicated by S. CNC-b-ZIP represents the CNC domain, the basic DNA binding region and the leucine zipper respectively. a, b and c indicate regions of variability in the TCF11 protein, resulting from alternative exon usage (19). A methionine cluster that probably directs internal translation initiation of TCF11 (19,21) is represented by MMM. MBP indicates the region consisting of *E. coli* maltose binding protein.

*Pst*I-digested pMALc. Translation of MBP-*TCF11*-H results in a protein of 546 amino acids of which 158 amino acids are the TCF11 C-terminal part. The construction of MBP-Maf fusion proteins is described elsewhere (12,13,15). Briefly, the MBP-v-Maf is a fusion of MBP and the 129 C-terminal amino acids of chicken v-Maf. The corresponding number of amino acids for the other chicken Maf proteins are 117 (MafB), 138 (MafF) and 151 (MafG), all containing the complete bZIP region. Competent *Escherichia coli* BK2118 or Y1089 cells were transformed with the various pMALc constructs by electroporation followed by induction of fusion proteins and their isolation by amylose affinity chromatography. The MBP-*TCF11*-E fusion protein did not bind to the amylose resin and was purified by preparative SDS-PAGE and subsequent electroelution. Protein concentrations were determined by a modified Lowry procedure (27) and Coomassie staining. The protein solutions were diluted with an equal volume of 50% ethylene glycol (Pierce) and stored in small aliquots at  $-20^{\circ}\text{C}$ .

### *In vitro* transcription and translation

For protein dimerization assays, TnT-coupled transcription-translation rabbit reticulocyte extract (Promega) was used for the synthesis of full-length  $^{35}\text{S}$ -labelled TCF11 with plasmid pZcEA48 (19) as template. Alternatively, the transcription was performed using Ribomax (Promega) followed by RNA purification using the RNeasy kit (Qiagen). A 3.5 kb *Eco*RI-*Eco*RV *TCF11* cDNA fragment derived from pZcEA56 (19) comprising the protein coding region of *TCF11* was subcloned into pcDNA3 (Invitrogen) in the corresponding sites. The resulting plasmid, pcA56, was linearized with *Sma*I and used as a template for T7 RNA polymerase. The Maf pRAM-GEM subclones (11) were linearized with *Eco*RI and used as templates for T7 RNA polymerase.

### Protein dimerization assay

Around 300 ng MBP-*TCF11*-A or MBP-Maf fusion proteins were bound to amylose resin by incubating 100  $\mu\text{l}$  bacterial lysate with 100  $\mu\text{l}$  20% amylose resin in buffer A (10 mM sodium

phosphate, pH 7.2, 0.5 M NaCl, 10 mM  $\beta$ -mercaptoethanol, 1 mM EGTA and 1 mM sodium azide) on ice for 15 min. MBP alone was used as a control. The resin was washed in 1 ml buffer B (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% Nonidet P-40) and resuspended in 1 ml buffer B.  $^{35}\text{S}$ -Labelled TCF11 translated *in vitro* (2.5  $\mu\text{l}$ ) was added and the dimerization reaction incubated by gently mixing overnight at 4°C. Unbound protein was removed by washing three times in 750  $\mu\text{l}$  cold buffer B and bound protein was eluted by denaturation in SDS-PAGE sample buffer (2% SDS, 0.1 M dithiothreitol, 60 mM Tris-HCl, pH 6.8, 10% glycerol and 0.001% bromophenol blue). Released proteins were analyzed by SDS-PAGE/Coomassie staining and autoradiography.

### Electrophoretic mobility shift assay (EMSA)

Bacterially expressed fusion proteins were centrifuged at 30 000 *g* for 60 min to remove large particular material. Ten nanograms of each purified fusion protein in 2  $\mu\text{l}$  pMAL column buffer (10 mM Tris-HCl, pH 7.4, 200 mM NaCl and 1 mM EDTA):ethylene glycol (1:1) or K562 nuclear extract (3  $\mu\text{l}$ ) was incubated in EMSA reaction buffer [5 mM Tris-HCl, pH 8.0, 25 mM NaCl, 2.5 mM  $\text{MgCl}_2$ , 0.5 mM EDTA, 5% glycerol and 100  $\mu\text{g}/\text{ml}$  poly(dI-dC)] in a total volume of 10  $\mu\text{l}$  for 15 min at 20°C. When fusion proteins were used, 1% Tween-20 was added to the binding reaction. For heterodimer formation of MBP-TCF11-A or -H and MBP-MafG, fusion proteins were preincubated at 37°C for 30 min. In preliminary experiments these conditions were shown to enhance heterodimerization. Where indicated, cold competitor DNA (50 or 100 nM) was added to the reaction buffer.  $^{32}\text{P}$ -Labelled DNA probe (20 000–30 000 c.p.m.) was added to a final concentration of 1 nM and the reaction mixture incubated for 15 min at 20°C. Complexes were resolved by electrophoresis on a pre-run 5% polyacrylamide gel (acrylamide:bisacrylamide 36:1) in 0.4 $\times$  TBE buffer (1 $\times$  TBE is 89 mM Tris-borate, pH 8.3, and 2 mM EDTA) for 10 min at 10 V/cm. Electrophoresis was continued for 60 min at 6 V/cm. The EMSA reaction was applied to the gel while the electrophoresis was running. Subsequently, the gels were dried for autoradiography. The concentration of unlabelled single-stranded oligonucleotides was determined using OD<sub>260</sub> or gel electrophoresis and ethidium bromide staining. Double-stranded fragments were generated by hybridizing equimolar amounts of complementary single-stranded DNA oligonucleotides. Probes were end-labelled using [ $\gamma$ - $^{32}\text{P}$ ]dATP and polynucleotide kinase and purified by polyacrylamide gel electrophoresis.

### Binding assay using paramagnetic beads

All incubations and washings were performed at room temperature using slow vertical rotation to keep the beads in suspension. Dynabeads M280-streptavidin (Dyna) (2 mg) were washed three times in 700  $\mu\text{l}$  of each of the following solutions: PBS/1% BSA, PBS/1 M NaCl, EMSA buffer containing 1% Tween-20 and 0.3% Nonidet P-40 and 10 mM Tris (pH 8.0) 1 mM EDTA (TE) and finally resuspended in 1 ml TE. Biotinylated (5') double-stranded DNA was added to 0.72  $\mu\text{M}$ , which is approximately three times the binding capacity of the beads, and the suspension incubated for 45 min to promote biotin-streptavidin binding. The beads were washed twice in TE, suspended in 1 ml EMSA buffer and stored at 4°C.

MBP fusion proteins were preincubated for 30 min at 37°C in EMSA buffer to promote heterodimerization. Proteins (50 ng each) were then mixed with 50  $\mu\text{g}$  DNA beads resuspended in EMSA buffer containing 2  $\mu\text{g}$  poly(dI-dC), 1% Tween-20 and (where indicated) 180 pmol unbiotinylated competitor DNA in a total volume of 50  $\mu\text{l}$ . After 45 min slow rotation, unbound proteins were washed off by three 5 s whirlmix washes in 500  $\mu\text{l}$  EMSA buffer containing 1% Tween-20 and 0.3% Nonidet P-40. For proteins obtained by translation *in vitro*, the reticulocyte lysate containing translated proteins was mixed with an equal volume of 50% ethylene glycol at room temperature and stored on ice immediately following the translation reaction. Ten microlitres of this solution was added to 50  $\mu\text{g}$  DNA beads to reach a total volume of 50  $\mu\text{l}$  EMSA buffer containing 2  $\mu\text{g}$  poly(dI-dC), 0.5% Tween-20 and 0.15% Nonidet P-40. Protein-DNA binding and washing were performed as described, except that detergent concentrations were as in the binding reaction and incubation was performed for 2 h. Extraction of nuclear proteins from exponentially growing K562 cells was performed essentially according to Schreiber *et al.* (28). Approximately 40  $\mu\text{g}$  protein in 20  $\mu\text{l}$  nuclear extract was mixed with 50  $\mu\text{g}$  DNA beads in EMSA buffer containing 0.25% Tween-20, 0.08% Nonidet P-40 and 4  $\mu\text{g}$  poly(dI-dC) in a final volume of 100  $\mu\text{l}$ . Incubation and washing were performed as described except detergent concentrations were kept at the levels used in the binding reactions.

Retained proteins were eluted in SDS-PAGE sample buffer at 95°C and separated by SDS-PAGE. Gels were either fixed in 10% acetic acid/10% methanol for 30 min, incubated in Amplify (Amersham) for 30 min and dried for enhanced chemiluminescence detection ( $^{35}\text{S}$ -labelled proteins translated *in vitro*) or proteins were electroblotted from the gel onto membranes for detection with antibodies as described below.

### Production and purification of anti-TCF11 polyclonal antibodies

MBP-TCF11-E was purified by preparative SDS-PAGE. The appropriate gel slices were homogenized and antisera were raised in rabbits repeatedly injected subcutaneously with ~10  $\mu\text{g}$  recombinant MBP-TCF11-E fusion protein. High titre anti-TCF11 antibodies were obtained. Immunoassays established that these polyclonal antibodies against the C-terminus of TCF11 contained a high titre of both anti-TCF11 and anti-MBP antibodies. The anti-MBP antibody was removed by absorption to MBP immobilized on amylose resin. The IgG fraction was then isolated from the anti-MBP-depleted antiserum using protein A-Sepharose (Pharmacia) and further affinity purified on a MBP-TCF11-E-Sepharose column. The purified antibody was eluted with 1 M citric acid, 0.2 M  $\text{K}_2\text{HPO}_4$ , pH 2.2, and immediately dialysed against PBS.

### SDS-PAGE, immunoprecipitation and immunoblotting

Protein samples were prepared for SDS-PAGE by reduction and denaturation at 99°C for 3 min in SDS-PAGE sample buffer. Minigel electrophoresis (BioRad) was performed for 50 min, 10 V/cm, in Tris-glycine SDS running buffer (25 mM Tris base, 192 mM glycine, 3.5 mM SDS, pH 8.3). The gels were either stained with Coomassie brilliant blue or prepared for blotting by incubation in protein transfer buffer (25 mM Tris base, 192 mM glycine and 20% methanol, pH 8.3) for 5–10 min. Proteins were electroblotted onto polyvinylidene difluoride membranes (Immobilon P;

Millipore) in cold protein transfer buffer for 60 min at 20 V/cm. Membranes were either stained with amido black or incubated in TNT (150 mM NaCl, 0.1% Tween-20, 10 mM Tris-HCl, pH 8.0) containing 5% sterile BSA (Gibco BRL). Antibodies (0.1–1 µg/ml) were added and the protein blots incubated in TNT overnight at 4°C with gentle agitation. Unbound antibodies were removed by three 15 min washes in TNT and bound antibodies detected using horseradish peroxidase (HRP)-conjugated donkey anti-rabbit antibodies and the enhanced chemiluminescence (ECL) detection system (Amersham). Immunoprecipitations were performed on K562 nuclear extracts using affinity-purified anti-TCF11 antibodies or preimmune antibodies coupled to protein A-Sepharose using dimethyl pimelimidate (Sigma). Immunocomplexes were resolved by SDS-PAGE, electroblotted and detected with biotinylated anti-TCF11 antibodies, streptavidin-HRP and ECL (Amersham).

## RESULTS

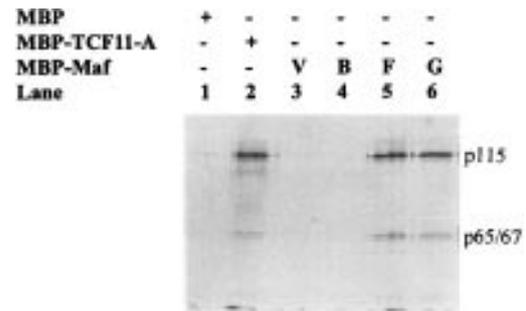
### TCF11 heterodimers with small Maf proteins

TCF11, the erythroid-specific p45 NF-E2, Nrf2 and ECH have highly similar amino acid sequences in the 63 amino acid CNC domain. Since this domain includes the basic DNA binding region, it is not surprising that these factors recognize very similar DNA sequences. The leucine zipper, located immediately C-terminal of the CNC domain, controls dimerization and hence determines which proteins can interact to form active, DNA binding homo- or heterodimers. This interaction surface is not particularly well conserved among CNC domain family members, however, both p45 NF-E2 and ECH have been shown to interact with small Maf proteins (7,8,13,14) and will probably compete for binding to Maf proteins when co-expressed.

To test if TCF11 could interact with the Maf protein family members, we performed a protein interaction assay. Freshly prepared <sup>35</sup>S-labelled *in vitro* translated TCF11 was incubated with bacterially expressed TCF11 (MBP-TCF11-A) (Fig. 1) or MBP-Maf proteins. Co-precipitated [<sup>35</sup>S]TCF11 and MBP fusion proteins were collected using amylose resin, eluted and analysed by SDS-PAGE and autoradiography (Fig. 2). Only the small Maf fusion proteins MBP-MafF and MBP-MafG and the MBP-TCF11-A fusion protein bound [<sup>35</sup>S]TCF11. No binding of [<sup>35</sup>S]TCF11 was observed with the large Maf fusion proteins MBP-v-Maf and MBP-MafB or MBP alone, indicating that the observed interaction between TCF11 and small Maf proteins was highly specific. Both TCF11 isoforms (p115 and p65/67) formed homodimers and also heterodimers with small Maf proteins. A likely conclusion based on these observations is that the TCF11 isoforms may compete with p45 NF-E2 and ECH in binding to small Maf proteins in erythroid tissue.

### Homo- and heterodimer binding to the 5'HS2 tandem NF-E2 site

Interestingly, LCR-F1 alone binds the β-globin 5'HS2 tandem NF-E2 site in an electrophoretic mobility shift assay (22), suggesting that TCF11, like the small Maf proteins, binds to this site as a homodimer. To test if TCF11-small Maf heterodimerization would enhance protein binding to the tandem NF-E2 site, we performed EMSA on bacterially synthesized TCF11 and MafG fusion proteins. Prior to DNA binding, the proteins were preincubated for 30 min at 37°C to promote heterodimerization.

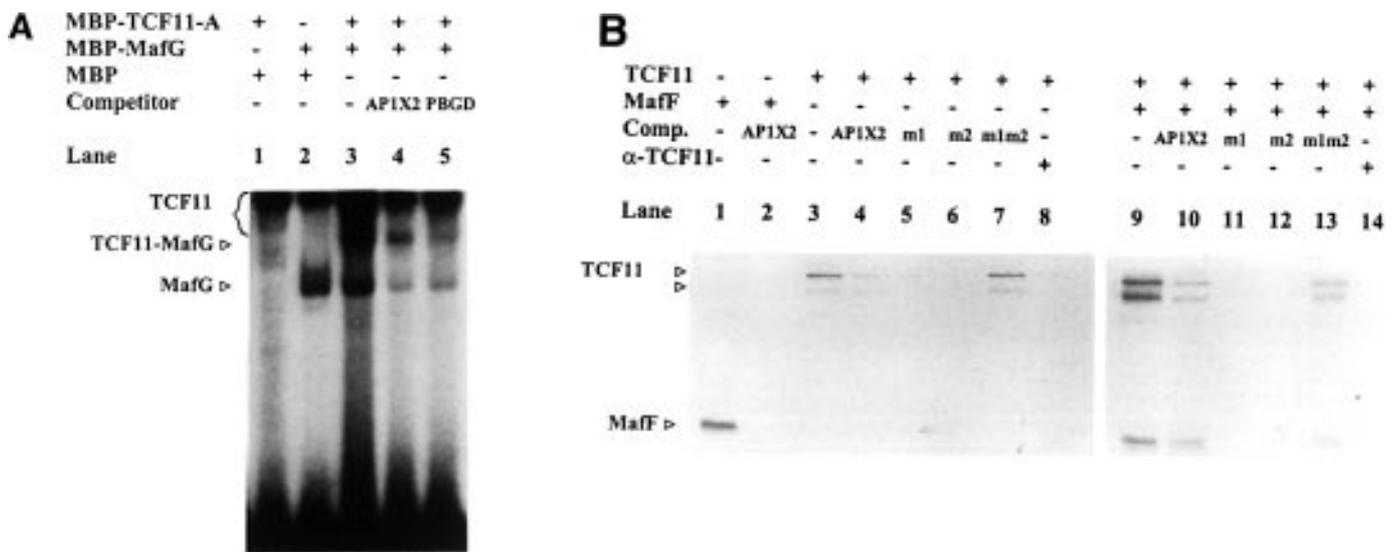


**Figure 2.** Protein interaction assay using <sup>35</sup>S-labelled TCF11 (translated *in vitro*) and various MBP fusion proteins immobilized on amylose resin. [<sup>35</sup>S]TCF11 incubated with: lane 1, MBP; lane 2, MBP-TCF11-A; lane 3, MBP-v-Maf; lane 4, MBP-MafB; lane 5, MBP-MafF; lane 6, MBP-MafG. After binding and washing, retained [<sup>35</sup>S]TCF11 and the MBP derivatives were eluted and subjected to SDS-PAGE on 8% gels and autoradiography. P115 and p65/67 represent full-length and internally translation initiated TCF11 respectively.

The TCF11 complex probably formed a homopolymer in the presence of the 5'HS2 probe. We were unable to resolve this by varying binding and electrophoresis conditions or by using the truncated TCF11 construct MBP-TCF11-H. Most of the complex remained in the well (Fig. 3A, lane 1) under non-denaturing conditions and when *in vitro* translated TCF11 was used (not shown). The presence of both TCF11 and MafG in the binding reaction resulted in enhanced binding to the 5'HS2 probe (Fig. 3A, lane 3), indicating modulation of binding following heterodimerization. Oligonucleotides containing either the 5'HS2 tandem NF-E2 site or the single PBGD site both competed significantly but not completely with the labelled 5'HS2 probe (lanes 4 and 5).

To circumvent the problem with resolution of TCF11 complexes, we coupled the tandem NF-E2 site to paramagnetic beads, which were incubated with the proteins to be tested. Retained proteins were eluted from the DNA beads and analysed by SDS-PAGE and autoradiography of <sup>35</sup>S-labelled proteins translated *in vitro*. Both TCF11 and MafF separately bound specifically to the tandem NF-E2 site in the bead assay (Fig. 3B, lanes 1–4). As a homodimer, both long and short TCF11 isoforms bound to the site; the long isoform (p115) was most active in DNA binding (lane 3). By including competitor oligonucleotides with four mutated positions (TGCTGAGTCAT→TGCACGTTTCAT) in the first (m1), second (m2) or both (m1m2) NF-E2 sites (Table 1), we observed that both m1 and m2 but not m1m2 inhibited DNA binding by TCF11 (Fig. 3B lanes 5–7). This indicates that both NF-E2 sites in the tandem repeat are used by TCF11 *in vitro* and that this four base substitution at one AP1 half-site inhibits binding by TCF11.

We next co-translated TCF11 and MafF and assayed for binding to the tandem NF-E2 site on beads (Fig. 3B). As shown, co-translation of the proteins increased the DNA binding ~5-fold, as determined by <sup>35</sup>S scanning (Fig. 3B, lane 3 versus lane 9). This increase in binding upon co-translation probably reflected an increased overall affinity of the heterodimer for the tandem NF-E2 site, consistent with our observations using bacterially synthesized proteins (Fig. 3A). By using the DNA bead assay we could observe the relative contribution of long and short TCF11 isoforms to the DNA-associated heterodimer compared with the TCF11 homodimer. The short TCF11 isoform appeared to be the preferred heterodimeric partner of MafF in binding to DNA



**Figure 3.** Binding of TCF11 and small Maf proteins to the 5'HS2 tandem NF-E2 site in the β-globin LCR. (A) EMSA of MBP-TCF11-A and MBP-MafG bound to the tandem NF-E2 site (APIX2). Lanes 4 and 5, cold competitor DNA was added to 50-fold molar excess. The proteins indicated were preincubated to promote heterodimerization as described in Materials and Methods. (B) Binding of *in vitro* translated TCF11 and MafF to the APIX2 probe immobilized to paramagnetic beads in the presence or absence of various oligonucleotide competitors. The <sup>35</sup>S-labelled proteins were eluted and analyzed by SDS-PAGE and autoradiography. Comp. indicates the cold competitor oligonucleotide used (Table 1). α-TCF11 indicates the presence or absence of affinity-purified anti-TCF11 antibody in the reaction.

**Table 1.** Oligonucleotides used in protein binding assays

APIX2:	5' CAAGCACAGCAATGCTGAGTCA TGATGAGTCA TGCTGAGGC 3'
m1:	5' CAAGCACAGCAATGCTGAGTCA TGATGAGTCA TGCTGAGGC 3'
m2:	5' CAAGCACAGCAATGCTGAGTCA TGAGCGTCA TGCTGAGGC 3'
m1m2:	5' CAAGCACAGCAATGCTGAGTCA TGAGCGTCA TGCTGAGGC 3'
PBGD:	5' TGGGGAACCTGTGCTGAGTCA CTGGAG 3'
mPBGD:	5' TGGGGAACCTGTGCTGAGTCA CTGGAG 3'
P1:	5' TGGGGAACCTGTGCTGAGTCA CTGGAG 3'
P2:	5' TGGGGAACCTGTGCTGAGTCA CTGGAG 3'
PCG:	5' TGGGGAACCTGTGCTGAGTCA CTGGAG 3'
ΔAPIX2	5' CAATGCTGAGTCA TGATGAGTCA TGCT 3'
CRE:	5' AGAGATTGCTGAGTCA GAGAGCTAG 3'
API:	5' CGCTTGATGAGTCA GCCGGAA 3'

Only the upper strand is shown. APIX2 contains a tandem NF-E2 site and flanking sequence from the human 5'HS2 in β-globin LCR. PBGD contains a single NF-E2 site and flanking sequence from the human erythroid-specific PBGD promoter. API/CRE binding sites are underlined, NF-E2 consensus bases (apart from the API/CRE core) are highlighted and mutated bases are shown in italics.

(Fig. 3B, lane 3 versus lane 9). As was observed for the TCF11 homodimer, both m1 and m2 completely inhibited DNA binding by the heterodimer (lanes 11 and 12), indicating that both NF-E2 sites in the tandem repeat are occupied by the heterodimer as well. The double mutant m1m2 also competed with the heterodimer (lane 13), suggesting that the heterodimer tolerates the TGAG→ACGT substitution in the AP1 half-site. Addition of polyclonal antibody directed against MBP-TCF11-E (Fig. 1) and depleted of anti-MBP inhibited DNA binding by the TCF11

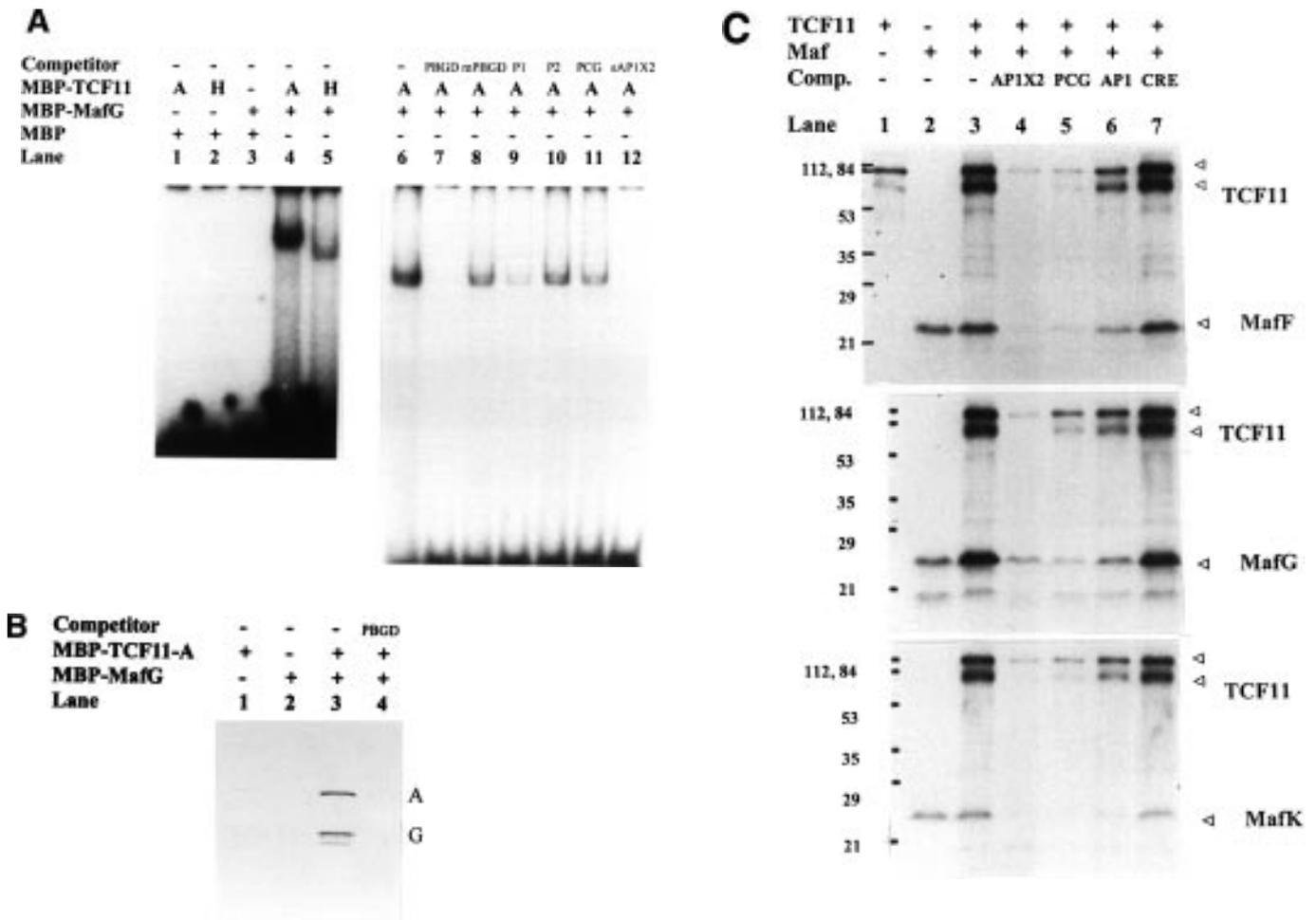
homodimer and the TCF11-MafF heterodimer (lanes 8 and 14 respectively). The lack of binding by MafF in the presence of the anti-TCF11 antibody suggests that most MafF is in a complex with TCF11 when the proteins are co-translated. Our antibody can thus be used as a specific inhibitor of DNA binding by TCF11 and TCF11-containing heterodimers.

These findings suggest that by interacting with a small Maf family member, TCF11 binding to the 5'HS2 tandem NF-E2 site is increased. The TCF11 homodimer and the TCF11-MafF heterodimer bind to both NF-E2 sites in the tandem repeat, the long TCF11 isoform (p115) apparently being most active as a homodimer while the short isoform (p65/67) appears to be most active as a heterodimer in the bead assay.

### Homo- and heterodimer binding to the PBGD single NF-E2 site

Another NF-E2 site is the single site present in the PBGD erythroid promoter (1,6). A G→T mutation at -2 relative to the AP1 core in this NF-E2 site results in loss of inducibility of the PBGD promoter in murine erythroleukemia (MEL) cells (6), indicating that the NF-E2 site is essential for erythroid-specific activation of the PBGD promoter. Interestingly, mice which do not express p45 NF-E2 revealed that this factor is not essential for expression of erythroid-specific PBGD (29). Thus, a factor(s) which regulates erythroid PBGD through this site remains to be identified. None of the CNC domain family members show strong binding to this site alone, but following heterodimerizing with small Maf proteins binding is enhanced (7,14).

In EMSA, no binding by the TCF11 homodimers MBP-TCF11-A or MBP-TCF11-H was detected to the single NF-E2 site derived from the erythroid PBGD promoter (Fig. 4A, lanes 1 and 2). MBP-MafG homodimer binding (lane 3) was weak but detectable using long exposure (data not shown). Mixing the MBP-TCF11 and MBP-MafG fusion proteins resulted in complexes very active in binding to the probe (lanes 4 and 5). This

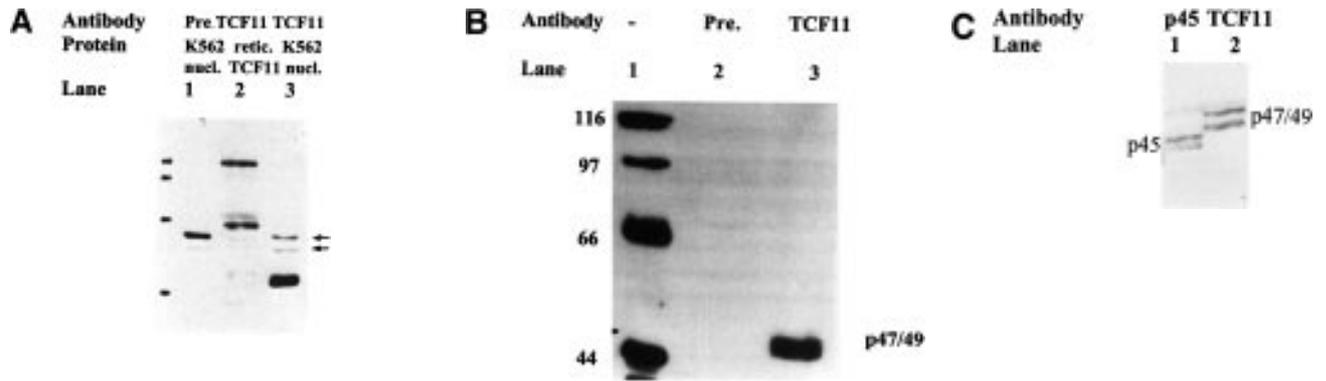


**Figure 4.** Binding of TCF11 and small Maf proteins to the single NF-E2 site from the PBGD erythroid-specific promoter. When purified MBP fusion proteins were used, a preincubation step to promote heterodimerization was performed as described in Materials and Methods. All translations *in vitro* were performed in rabbit reticulocyte lysate at the same time using the same reagents. Results were visualized by autoradiography, except where indicated. **(A)** EMSA of the purified MBP fusion proteins MBP-TCF11-A or -H, MBP-MafG and the PBGD probe. The various unlabelled competitor oligonucleotides are listed in Table 1. **(B)** Binding of the purified MBP fusion proteins MBP-TCF11-A and MBP-MafG to the PBGD sequence immobilized on paramagnetic beads. A indicates MBP-TCF11-A and G indicates MBP-MafG. Bound proteins were separated by SDS-PAGE, electroblotted and detected with an anti-MBP antibody. **(C)** Binding of the co-translated proteins to the PBGD sequence immobilized to paramagnetic beads. Reticulocyte lysate translations were used as protein input in DNA binding reactions including 50µg PBGD beads as described in Materials and Methods. Specifically bound proteins were eluted and separated by SDS-PAGE on 14% gels. At the top of the figures, RNA input into the translation reaction and unbiotinylated competitor DNA are indicated. Retained small Maf proteins are indicated on the right of the figure.

indicates that heterodimer formation between TCF11 and MafG is necessary for efficient binding to this site. We used the two MBP-TCF11 fusion proteins -A and -H (Fig. 1), differing by 15 kDa, as MBP-MafG partners. Heterodimers between each of these TCF11 fusion proteins and MBP-MafG migrated differently in EMSA, showing the presence of MBP-TCF11 in the DNA-bound complex (lanes 4 and 5). A second control for the presence of both proteins in the DNA binding complex was performed. By DNA affinity isolation of the complex formed between the fusion proteins and the single NF-E2 site using DNA beads, both MBP-TCF11-A and MBP-MafG were detected when mixed, but not when assayed separately (Fig. 4B). By including an excess of unlabeled PBGD probe containing an intact NF-E2 site (TGCTGAGTCAC, Table 1), the shift was competed out (Fig. 4A, lane 7, and B, lane 4). As expected, an oligonucleotide containing a G→T mutation in position -2 relative to the AP1 core (mPBGD, TTCTGAGTCAC) did not compete effectively with the probe (Fig. 4A, lane 8). Competition using mutant

oligonucleotides P1 (TGCTGAGTCAG) and P2 (TGATGAGTCAG) showed that P1 competed fully (lane 9), whereas P2 had an effect similar to that of mPBGD. The C→G mutation alone was therefore not sufficient to abolish competition. PCG, a PBGD NF-E2 site with CRE-like half-site spacing (TGCTGACGTCAC), was a weak competitor (lane 11). Such TRE-like half-site spacing preference has also been observed for native NF-E2 (7). As expected, the 5'HS2 tandem NF-E2 site effectively competed with the PBGD single NF-E2 site (lane 12). Neither in EMSA nor in the bead assay did MBP-TCF11-A or -H bind to the PBGD probe. This was most likely due to their N-terminal deletions causing reduced homodimer formation or DNA binding. These species of TCF11 can form homodimers with *in vitro* translated TCF11 (Fig. 2), but may be unable to form AA or HH homodimers.

The small Maf proteins have virtually identical DNA binding specificity (13). We therefore examined if *in vitro* co-translated TCF11 and small Maf proteins MafF, MafG and MafK displayed



**Figure 5.** Proteins detected by anti-TCF11 antibody. (A) SDS-PAGE and immunoblotting of 10  $\mu$ g K562 nuclear proteins (lanes 1 and 3) and 1  $\mu$ l reticulocyte lysate containing TCF11 translated *in vitro* (lane 2). Proteins were detected using the polyclonal anti-TCF11 antibody (lanes 2 and 3) or preimmune rabbit immunoglobulins (lane 1). Arrows indicate two proteins also recognized by the preimmune antibody. Molecular weight markers are (top to bottom): 116, 97, 66 and 45 kDa. (B) Immunoprecipitation of TCF11 from K562 nuclear extract using affinity-purified polyclonal anti-TCF11 antibodies (lane 3) or preimmune antibodies (lane 2). Lane 1 is a biotinylated protein standard of 116, 97, 66 and 44 kDa (top to bottom). Precipitated TCF11 proteins were resolved by SDS-PAGE (8% gel), electroblotted and detected using biotinylated anti-TCF11 antibody. (C) Detection of proteins in 10  $\mu$ g K562 nuclear protein extract resolved by SDS-PAGE (8% gel), electroblotted and incubated with anti-p45 NF-E2 (lane 1) or anti-TCF11 (lane 2) antibodies as described in Materials and Methods. The weakly stained upper band recognized by both antibodies is also recognized by several other antibodies (data not shown).

the same binding pattern to the single NF-E2 site. When DNA binding reactions using *in vitro* co-translated proteins and the single NF-E2 site coupled to beads were performed, an ~30-fold increase (determined by  $^{35}$ S scanning) in TCF11 binding was observed by co-translation of TCF11 and each of the small Mafs when compared with the homodimer (Fig. 4C, lanes 1 and 3). Again, this indicates modulation of substrate affinity by heterodimer formation. Using *in vitro* synthesized TCF11 and PBGD beads, specific homodimer binding was clearly detected, in contrast to the results using bacterially synthesized TCF11. As observed using the 5'HS2 tandem NF-E2 site, the long TCF11 isoform (p115) bound more efficiently as a homodimer and the short isoform (p65/67) was most active in binding as a heterodimer with the small Maf protein family members (Fig. 4C, lanes 1 and 3). The various TCF11–small Maf heterodimers displayed similar DNA recognition profiles using NF-E2-like sites (lanes 4 and 5), an AP1 site (lane 6) or a CRE site (lane 7) oligonucleotide as competitors in the binding reactions. This experiment also suggested that the AP1 (TGAGTCA) or CRE (TGACGTC) core sequences were not sufficient for strong binding by the TCF11–small Maf heterodimer.

#### Antibody detection of TCF11

The ubiquitous expression of TCF11 mRNA (19,21) led us to investigate if the TCF11 protein could be detected in different human tissues and cell lines. The polyclonal anti-TCF11 antibody used to inhibit DNA binding (Fig. 3B, lanes 8 and 14) recognized both TCF11 isoforms translated *in vitro* in immunoblotting experiments (Fig. 5A, lane 2) and could thus be expected to do so in other crude extracts as well. In K562 cell extracts, two proteins of 47 and 49 kDa were detected in both cytoplasmic (data not shown) and nuclear fractions (Fig. 5A, lane 3). This doublet was recognized in a native state as well, as shown by immunoprecipitation of p47/49 from K562 nuclear extracts (Fig. 5B). The possibility of cross-reacting with the other CNC domain family members, p45 NF-E2 and Nrf2, was eliminated, since the p47/49 doublet did not co-migrate with the proteins specifically detected by anti-p45 (Fig. 5C) or anti-Nrf2 (data not shown). The observed

protein sizes are in good agreement with the expected molecular weight of internally initiated TCF11 (49 kDa), but differ from the apparent size (21) of *in vitro* translated internally initiated TCF11 (65/67 kDa).

To specifically block DNA binding of TCF11 in K562 nuclear extracts, we included purified anti-TCF11 or preimmune antibodies in a preincubation step before DNA binding. This antibody does not react with p45 NF-E2 (Fig. 5C). Addition of anti-TCF11 antibody specifically reduced one protein complex (Fig. 6A, arrows) using both the single and the tandem NF-E2 sites as probes (Fig. 6A, lanes 3 and 8 respectively). The preimmune antibody did not affect binding (lanes 4 and 9) and addition of unlabelled competitor DNA verified that the shifts were specific (lanes 5 and 10). The other shifts are expected to be caused by a number of other proteins binding to AP1 sites.

We next examined if the p47/49 doublet in K562 cell extract could specifically bind to the PBGD single NF-E2 site immobilized on beads. K562 nuclear proteins were incubated with the DNA beads and retained proteins were separated by SDS-PAGE, electroblotted and incubated with the anti-TCF11 antibody. P47/49 bound to the NF-E2 site (Fig. 6B) and reacted with the anti-TCF11 antibody. No protein was detected when naked beads were used. These results show that the p47/49 doublet is TCF11 isoforms. They are also consistent with the finding that Nrf1/LCR-F1 binds to this site in K562 cells (21,22). P47/49 requires the conserved G at position –2 relative to the AP1 core for binding (data not shown). Unbiotinylated PBGD site and PCG oligonucleotides competed for binding by the p47/49 doublet, showing that the binding was specific (Fig. 6B, lanes 4 and 5). Thus, the TRE-like half-site spacing preference observed using bacterially expressed and *in vitro* translated proteins is less pronounced for the p47/49 doublet. In this bead assay, the uncertainty of cross-reactivity in antibody blocking of DNA binding is eliminated.

#### DISCUSSION

Increasing knowledge of the regulation of gene expression has revealed the combinatorial nature of regulatory complexes. The



product. The size discrepancy between p47/49 and TCF11 translated *in vitro* could be due to protein modifications in the reticulocyte lysate or the endogenous p47/49 doublet could be the result of post-translational modification/proteolysis of p65/67 *in vivo*. The p47/49 doublet was also observed in HL60 cells (not shown). Chan and co-workers observed a 65 kDa protein in K562 extracts using an anti-Nrf1 antibody directed against a synthetic peptide derived from the leucine zipper region (21). This 65 kDa protein could be a preprocessed form of internally initiated TCF11, although, using several nuclear protein preparations and a protease inhibitor cocktail for the extractions, we have never observed this species in nuclear extracts.

## ACKNOWLEDGEMENTS

This work was supported by grants to professor Anne-Brit Kolstø from the Research Council of Norway, the University of Oslo and the Norwegian Cancer Society. Øyvind Johnsen is a Research Fellow of the Research Council of Norway and University of Oslo. Nina Skammelsrud is a Research Fellow of the Norwegian Cancer Society.

## REFERENCES

- Mignotte, V., Wall, L., deBoer, E., Grosveld, F. and Romeo, P.H. (1989) *Nucleic Acids Res.*, **17**, 37–54.
- Andrews, N.C., Erdjument-Bromage, H., Davidson, M.B., Tempst, P. and Orkin, S.H. (1993) *Nature*, **362**, 722–728.
- Talbot, D., Philipsen, S., Fraser, P. and Grosveld, F. (1990) *EMBO J.*, **9**, 2169–2178.
- Ney, P.A., Sorrentino, P., McDonagh, K.T. and Nienhuis, A.W. (1990) *Genes Dev.*, **4**, 993–1006.
- Reitman, M. and Felsenfeld, G. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 6267–6271.
- Mignotte, V., Eleouet, J.F., Raich, N. and Romeo, P.-H. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6548–6552.
- Igarashi, K., Kataoka, K., Itoh, K., Hayashi, N., Nishizawa, M. and Yamamoto, M. (1994) *Nature*, **367**, 568–572.
- Andrews, N.C., Kotkow, K.J., Ney, P.A., Erdjument-Bromage, H., Tempst, P. and Orkin, S.H. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 11488–11492.
- Igarashi, K., Itoh, K., Motohashi, H., Hayashi, N., Matuzaki, Y., Nakauchi, H., Nishizawa, M. and Yamamoto, M. (1995) *J. Biol. Chem.*, **270**, 7615–7624.
- Nishizawa, M., Kataoka, K., Goto, N., Fujiwara, K.T. and Kawai, S. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 7711–7715.
- Fujiwara, K.T., Kataoka, K. and Nishizawa, M. (1993) *Oncogene*, **8**, 2371–2380.
- Kataoka, K., Fujiwara, K.T., Noda, M. and Nishizawa, M. (1994) *Mol. Cell Biol.*, **14**, 7581–7591.
- Kataoka, K., Igarashi, K., Itoh, K., Fujiwara, K.T., Noda, M., Yamamoto, M. and Nishizawa, M. (1995) *Mol. Cell Biol.*, **15**, 2180–2190.
- Itoh, K., Igarashi, K., Hayashi, N., Nishizawa, M. and Yamamoto, M. (1995) *Mol. Cell Biol.*, **15**, 4184–4193.
- Kataoka, K., Noda, M. and Nishizawa, M. (1994) *Mol. Cell Biol.*, **14**, 700–712.
- Kerppola, T.K. and Curran, T. (1994) *Oncogene*, **9**, 3149–3158.
- Kataoka, K., Noda, M. and Nishizawa, M. (1996) *Oncogene*, **12**, 53–62.
- Walters, M., Andrews, N.C., Magis, W. and Martin, D.I.K. (1996) *Exp. Hematol.*, **24**, 445–452.
- Luna, L., Johnsen, Ø., Skartlien, A.H., Pedetour, F., Turc-Carel, C., Prydz, H. and Kolstø, A.B. (1994) *Genomics*, **22**, 553–562.
- Luna, L., Skammelsrud, N., Johnsen, Ø., Abel, K.J., Weber, B.L., Prydz, H. and Kolstø, A.B. (1995) *Genomics*, **27**, 237–244.
- Chan, J.Y., Han, X.-L. and Kan, Y.W. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 11371–11375.
- Caterina, J.J., Donze, D., Sun, C.-W., Ciavatta, D.J. and Townes, T.M. (1994) *Nucleic Acids Res.*, **22**, 2383–2391.
- Mohler, J., Vani, K., Leung, S. and Epstein, A. (1991) *Mech. Dev.*, **34**, 3–9.
- 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.
- Moi, P., Chan, K., Asunis, I., Cao, A. and Kan, Y.W. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 9926–9930.
- 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.
- Markwell, M.A.K., Haas, S.M., Bieber, L.L. and Tolbert, N.E. (1978) *Anal. Biochem.*, **87**, 206–210.
- Schreiber, E., Matthias, P., Müller, M.M. and Schaffner, W. (1989) *Nucleic Acids Res.*, **17**, 6419.
- Shivdasani, R.A. and Orkin, S.H. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 8690–8694.
- Kotkow, K.J. and Orkin, S.H. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 3514–3518.