© 1997 Oxford University Press
Nucleic Acids Research, 1997, Vol. 25, No. 15 2953–2959

The world according to Maf

Hozumi Motohashi, Jordan A. Shavit[1], Kazuhiko Igarashi, Masayuki Yamamoto and James Douglas Engel[1,*]

Institute of Basic Medical Sciences and Center for TARA, University of Tsukuba, Tsukuba 305, Japan and [1]Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, IL 60208-3500, USA

Received April 11, 1997; Revised and Accepted June 11, 1997

ABSTRACT

Maf family proteins are so named because of their structural similarity to the founding member, the oncoprotein v-Maf. The small Maf proteins (MafF, MafG and MafK), as do all family members, include a characteristic basic region linked to a leucine zipper (b-Zip) domain which mediate DNA binding and subunit dimerization respectively. The small Maf proteins form homodimers or heterodimers with other b-Zip proteins present in the cell and bind to Maf recognition elements (MARE) in DNA. Since they lack known transcriptional activation domains, the small Maf proteins function either as obligatory heterodimeric partner molecules with numerous large subunits, discussed below, or alternatively as homo- or heterodimeric transcriptional repressors. The three small Maf proteins are expressed in a number of overlapping tissues, but their expression profiles nonetheless appear to be under meticulous tissue- and developmental stage-specific control. The MARE bears a striking resemblance to the NF-E2 binding sequence. NF-E2 binding sites in the human β-globin locus control region have been directly implicated as integral components in the circuitry required for eliciting changes in chromatin structure that precede globin gene activation. While the NF-E2 DNA sequence has been shown to be important for erythroid-specific gene regulation, a growing list of other genes may also be regulated through the same, or very similar, *cis* elements in non-erythroid cells. Taken together, these observations argue that comprehensive analysis of the activities of the small Maf proteins may provide a unique perspective for expanding our understanding of transcriptional regulation that can be elicited through interacting transcription factor networks.

INTRODUCTION

Mammalian embryogenesis is a complex process during which initially naive, often multipotent, cells proliferate, migrate and differentiate in response to inductive cues to form the tissues that will eventually comprise an independent organism. During this time, numerous signaling and response genes are turned on and off in different developing and migrating cells, tissues and organs. This process is regulated by a variety of extrinsic and self-signaling events which ultimately lead to the nucleus. There the transduced signals exert regulatory control over transcription factors that bind to specific *cis*-regulatory elements which then activate or repress expression of specific sets of genes. In this regard, numerous oncoproteins and related cellular factors have become conspicuous as critical players in this intricate regulatory pavane (1).

The founding member of the Maf protein family (v-Maf) was originally discovered as the transduced transforming component of avian musculoaponeurotic fibrosarcoma virus, AS42 (2). Subsequent studies identified the cellular homolog of this gene, c-*maf*, from which the v-*maf* oncogene was originally transduced (3), but in addition, c-*maf* was found to be but one member of an extended multigene family. Products of the *maf* proto-oncogene and related family members (the Maf family proteins) share a common, relatively well-conserved basic region and leucine zipper (b-Zip) motif which mediate DNA binding and dimer formation. Members of the Maf family are divided into two subgroups: the large Maf proteins, c-Maf (3), MafB (52) and NRL (5), all of which contain a distinctive acidic domain that probably enables transcriptional activation, and the small Maf proteins, MafK (6), MafF (6) and MafG (7), all of which lack activation domains.

A direct physiological role(s) for Maf family proteins remained elusive from 1989, when c- and v-*maf* were originally described, until ~4 years later, when the small Maf proteins were first shown to function as one subunit of nuclear factor-erythroid 2 (NF-E2), an erythroid-specific transcription factor (8,9). NF-E2 was shown to be comprised of 45 and 18 kDa subunits, the former being p45 NF-E2 (the founding member of the vertebrate CNC transcription factor family; below) and the latter was shown to be a small Maf protein. Since then, studies detailing the functional activity of various large and small Maf proteins have appeared and reports describing their important contributions to development and differentiation have become more widespread (10–13). In this review we summarize recent progress in analysis of the small Maf proteins (and, more briefly, describe the activity of the numerous and sundry partners of the small Mafs) and present speculation and evidence supporting the contention that both homodimers and heterodimers of the small Maf proteins exert finely articulated transcriptional control over gene expression during development, differentiation and oncogenesis.

*To whom correspondence should be addressed. Tel: +1 847 491 5139; Fax: +1 847 467 2152; Email: d-engel@nwu.edu

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

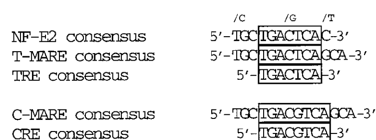

Figure 1. Consensus sequences for binding of Maf homodimers and heterodimers of Maf and related b-Zip proteins. The 11 bp consensus sequence for NF-E2/MARE binding is shown at the top. The MARE motif shows a close similarity to the TRE type-Maf recognition element (T-MARE; 4). Maf homodimers bind to a 13 bp palindromic sequence T-MARE or a 14 bp palindromic sequence CRE-type MARE (C-MARE; 4). Note that the NF-E2 consensus sequence and T-MARE contain a TRE and C-MARE contains a CRE. Both TRE and CRE sequences are boxed.

THE *CIS*-REGULATORY TARGETS OF MAF PROTEINS

Since the β-globin gene clusters in vertebrate organisms undergo a highly orchestrated pattern of gene 'switching' during development (where different genes are turned on and off at precise times during embryogenesis), they have proven to be an excellent model system for analyzing both tissue-specific and temporal aspects of gene regulation (14). The β-globin LCR (locus control region) plays a vitally important role in regulation of β-globin gene expression. The LCR is composed of five 5′ DNase I hypersensitive sites, which together act both as a cooperative activity stimulating globin gene expression (15–17) and to dominate the required architectural changes in the locus which result in chromatin 'opening' (18).

The human β-globin LCR contains consensus binding motifs for a variety of erythroid-restricted as well as (apparently) ubiquitous transcription factors. Of these motifs, the GATA and AP-1-related NF-E2 sequences are among the most intriguing, since both GATA and NF-E2 motifs had been identified in the 'core' element domains of HS2, HS3 and HS4 (the three HS elements that individually display the most significant enhancer activity). These binding sites were shown to be important for the function of the individual HS elements (19–22,24). Thus, identifying the LCR NF-E2 motif binding protein(s) became a critical hurdle in further elucidating LCR function.

The NF-E2 binding site was originally identified in the promoter of the porphobilinogen deaminase gene (25) and was subsequently identified in the regulatory regions of other erythroid-specific genes, such as ferrochelatase (26), as well as in the LCR. The consensus sequence for binding of NF-E2 was found to be TGCTGA(C/G)TCA(T/C) (Fig. 1). We observed that the NF-E2 motif bore a striking resemblance to the TRE [phorbol-12-*O*-tetradecanoate-13-acetate (TPA)-responsive element]-type Maf recognition element (T-MARE). Maf homodimers were shown to bind either to a 13 bp palindromic sequence T-MARE, TGCTGA-CTCAGCA, or a 14 bp palindromic sequence CRE (cAMP-responsive element)-type MARE (C-MARE), TGCTGACGTC-AGCA (4). The resemblance between the NF-E2 binding site and the T-MARE sequence provided our initial insight into the possible role that the small Maf proteins might play as an essential subunit of transcription factor NF-E2 (9), while other investigators arrived at many of the same conclusions through transcription factor purification and cloning by reverse genetics (8,23). Since NF-E2 has been used to describe both a *cis*-regulatory element and the transcription factor that binds to it, we use MARE to describe the DNA binding site and NF-E2 to indicate the protein in the rest of this review.

Another crucial observation in furthering our understanding of the role of Maf proteins in development was that sequences resembling the MARE motif were also identified in the regulatory regions of several non-erythroid genes (4). This was especially intriguing when considering the fact that many such non-erythroid MARE motifs contain TREs and that factors binding to these elements had been previously presumed to be heterodimers of Jun and Fos family members (i.e. subunits of transcription factor AP-1). For example, the type IV collagenase gene is a cellular TPA-inducible gene containing a TRE within a MARE sequence and thus the DNA motif can interact not only with AP-1, but also with NF-E2. Similarly, the antioxidant response element (ARE) in the glutathione S-transferase and NAD(P)H:quinone reductase genes shows marked similarity to the MARE sequence, and the opsin gene, a potential target of the large Maf protein NRL (specifically expressed in neural retina), harbors a TRE/MARE binding site. Recently the rhodopsin gene has been shown to be a specific target for NRL (27,28). These observations all suggest that transcription factors binding to the MARE *cis*-regulatory motif may play important roles in a variety of vertebrate cell types and that only very few of the actual target genes of these factors have been identified. In the light of these observations, the small Maf family proteins must be considered to be likely participants in regulation of gene expression in multiple tissues through MARE-related DNA regulatory elements. To restate this conjecture in a more provocative way: of the numerous TPA response elements that have been identified in the literature, how many (and which) of these TREs are actually *in vivo* targets for Maf transcription factor activity?

EXPRESSION PROFILES OF THE SMALL MAF PROTEINS

Three different small Maf proteins (MafK, MafF and MafG) have been identified and characterized in chickens (6,7) and the expression profiles of *mafK* and *mafF* mRNAs have also been examined (6). The expression of both mRNAs was detected in many tissues and organs, but their abundance was found to differ between tissues. Recently cDNA and genomic clones encoding human *mafK* and *mafG* have also been identified and the expression profiles of *mafK* and *mafG* mRNAs in human tissues were also examined (29). While both were detected in all tissues examined, *mafK* mRNA was found to be relatively more abundant in (for example) the placenta, while *mafG* mRNA was more abundant in the brain.

We previously cloned cDNAs encoding mouse MafK (also known as p18 NF-E2) and analyzed the expression of *mafK* mRNA in various murine tissues and cell lines (30). While mouse *mafK* mRNA was detected in numerous tissues by RNA blot analysis, the level of RNA was found to vary, often significantly, between tissues, as well as temporally during a particular developmental process. For example, the amount of *mafK* mRNA increases in parallel with hematopoietic activity in the mouse fetal liver (30). These results showed that the expression level of the *mafK* gene is regulated temporally and spatially in a dynamic manner and raised the possibility that quantitative control over MafK protein levels might significantly affect (for example) hematopoietic differentiation. The mouse *mafK* gene is the only

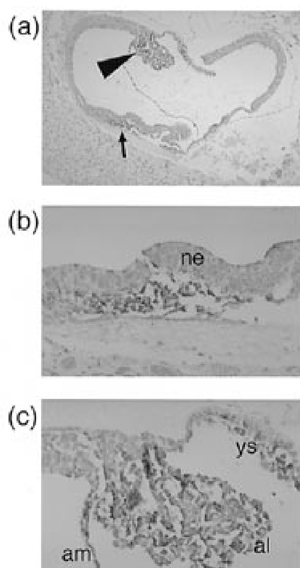

Figure 2. Expression of *mafK* mRNA in early mesodermal cells. *In situ* hybridization analysis was performed in murine embryos at 7.5 dpc (31). (**b** and **c**) Higher magnifications of the areas indicated by the arrow and the arrowhead respectively in (**a**). Intra-embryonic mesoderm and extra-embryonic mesoderm (allantois, amnion and yolk sac) display signals indicating abundant expression of *mafK* mRNA. Some of the blood cells in the yolk sac display a similarly intense signal. al, allantois; am, amnion; ne, neuroectoderm; ys, yolk sac.

small Maf family member whose developmental expression profile has been examined in detail (30,31).

The earliest that *mafK* mRNA has been detected in mouse embryos is at 7.5 days post-coitus (d.p.c.) (31). At this stage it is broadly expressed in mesodermal tissues, including intra-embryonic mesoderm, allantois, yolk sac and amnion, but not in other tissues of the embryo (Fig. 2). In early embryonic stages preceding 10.5 d.p.c. the major expression site of *mafK* mRNA is in mesenchymal cells. After that time, mesenchymal expression weakens reciprocally with an intensifying signal in hematopoietic cells of the fetal liver. Whereas *mafK* mRNA was almost undetectable in the neural tube prior to 10.5 d.p.c., ventrolateral cells and ventral midline cells in the spinal cord as well as dorsal root ganglia began to display prominent *mafK* mRNA expression by 13 d.p.c. (31). These results thus demonstrated that expression of *mafK* is under quite stringent cell lineage- and stage-specific transcriptional control during murine embryogenesis and further suggested that MafK could be important for differentiation of early mesodermal, mesenchymal, hematopoietic and neuronal cells.

In addition to this developmentally discrete temporal profile observed for *mafK* transcription, we also found that two distinct *mafK* mRNAs were generated in the mouse. cDNA and genomic structural analysis revealed that the mouse *mafK* gene is directed by two independent promoters, which encode separate 5′-untranslated regions. The remaining two exons, containing the entire coding portion of *mafK*, are used by both mature mRNAs. The upstream promoter (driving expression via exon IM) is used for mesenchymal and hematopoietic expression, while the other promoter (located 6 kbp 3′ of exon IM), which we called exon IN, is used for neuronal *mafK* transcription (31).

PARTNER MOLECULES OF THE SMALL MAF PROTEINS

The small Maf proteins form homodimers (7), heterodimers with one another (7) and also heterodimers with an ever-increasing constellation of additional b-Zip proteins (9,23,29,32,33). The first identified partner molecule of the small Mafs was p45, which was initially isolated as the large subunit of transcription factor NF-E2 (23,34,35). We and others then showed that the small Maf proteins serve as the lower molecular weight subunit (also referred to as the ubiquitous p18 subunit) of NF-E2 (8,9). Recently several p45-related molecules that participate in heterodimeric complex formation with the small Maf proteins have also been identified, named Nrf1, LCRF1 or TCF11 (36,37,60) and Nrf2 or ECH (32,38,39). Nrf1 and Nrf2 cDNAs were originally cloned from human erythroleukemia cells, whereas ECH was cloned from a chicken erythroid cell cDNA library. The structure of ECH is very similar to that of Nrf2, suggesting that the two are functional homologs. These three molecules (referred to as p45 NF-E2, Nrf1 and Nrf2) then form the ‘CNC family’ proteins (36), since they share the conserved b-Zip structure originally identified in the *Drosophila* CNC (cap’n’collar) protein (40).

Expression of p45 is restricted within hematopoietic cells and intestinal epithelia, while the expression profiles of both Nrf1 and Nrf2 show somewhat broader distribution. Nrf1 is strongly expressed in heart and skeletal muscle, kidney, lung and ovary (37), whereas Nrf2 is most prominently expressed in kidney, lung, fetal liver and fetal as well as mature muscle (38). ECH mRNA was most abundant in peripheral blood and was induced during differentiation of chicken erythroid cell lines (32). As is also the case for p45, neither Nrf1 (29) nor Nrf2 (29,32) form homodimers that effectively bind the T-MARE DNA sequence. Instead, all three CNC family proteins appear to form obligate heterodimers with one or another of the small Maf proteins, which then allows them to bind to a MARE to activate transcription (9,29,30,41).

What are the possible advantages in this scheme of compulsory heterodimeric interactions between these varied (CNC and small Maf family members) b-Zip factors that only together constitute the final activator proteins? One possible answer is that subtle variations in DNA binding specificity or in *trans*-activation or *trans*-repression potential (below) may well be generated through dimer formation of only a limited number of b-Zip transcription factor partners. This would then theoretically result in combinatorially complex, but exquisitely sensitive control over gene expression. In this manner, the small Maf proteins enable CNC, as well as other b-Zip proteins, to bind to DNA and exert their function. Thus different biological activities, all elicited through sometimes slightly different MARE binding sites, may come into play depending on the partner molecules with which the small Maf proteins heterodimerize. The available CNC and small Maf proteins that are present in any given tissue or at different stages of differentiation in that tissue, as well as the relative affinities between them, probably determine the final quality and quantity of transcriptional effects exerted at individual MARE sites.

In this regard, it should be noted that the small Maf proteins do not possess a canonical *trans*-activation domain and homodimers of the small Maf proteins have been shown to act as direct transcriptional repressors. It has been demonstrated that regulation from MARE sites can be turned on and off in living cells by experimentally manipulating the balance between the small Maf

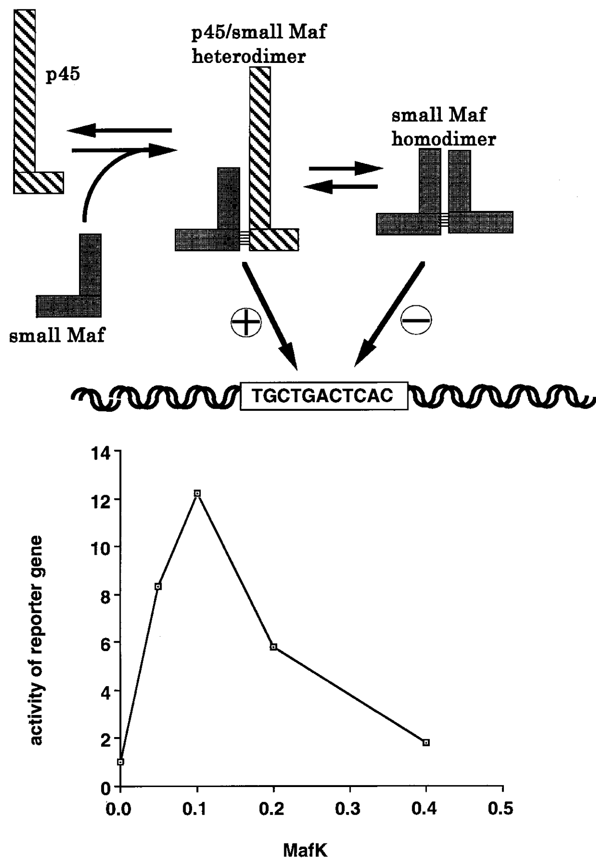

Figure 3. The quantitative balance between the small Maf proteins and p45 is essential for transcriptional activation or repression via MARE sites. (Upper) A model for binary transcriptional regulation transduced through a MARE site (9). Since p45 cannot form functional homodimers, it requires the small Mafs as obligatory partner molecules for binding to the MARE site. If there exists a lower abundance of the small Mafs relative to the amount of p45, heterodimers of the small Maf(s) and p45 are the dominant species in the cells, which then activate transcription from the MARE site. If the abundance of the small Mafs exceeds that of p45, the small Mafs form homodimers or heterodimers among one another and then suppress transcription by occupying the MARE sites. (Lower) This relationship was demonstrated experimentally. Increasing amounts of MafK were expressed in fibroblasts which had been transfected with a fixed amount of p45 (Nagai *et al.*, in preparation). Transfected MafK exerts both positive and negative transcriptional effects on a plasmid possessing three MARE sites directing expression of a reporter gene. Note that only a 4-fold difference in the amount of transfected *mafK* plasmid causes the change from activation to repression.

proteins and partner molecules that contain *trans*-activation domains (Fig. 3). If the abundance of CNC partner molecules is inadequate to 'titrate' all of the small Maf proteins produced in a cell, homodimers of the small Mafs would then be predicted to exert a dominant effect, leading to silencing through the MARE sites (9). We suspect that this mechanism could be operative *in vivo* as well as in cell culture, since the switch from activation to repression was found to occur within only a 4-fold difference in abundance of the small Maf protein (Nagai *et al.,* in preparation).

If this activator to repressor switch (or the reverse) is operative *in vivo*, it is intriguing to imagine the scenarios during cell differentiation where it would be advantageous for genes to be quiescent until they are required to execute a terminally differentiated function. For instance, the small Maf proteins might repress the expression of erythroid-specific genes in hematopoietic progenitor cells, but after commitment to differentiation of a mature erythrocyte, one of the large CNC family members could be activated (or the expression of a small Maf protein inactivated), thereby shifting the balance from repression toward activation of erythroid MARE target genes (42). As an additional level of complexity, p45–MafK heterodimers were reportedly able to bind to chromatin *in vitro* and disrupt nucleosomal structure (43), possibly reflecting a dual role for this heterodimer as both a transcriptional activator and as an architectural remodeling component of the LCR.

In addition to CNC family members, new partner molecules of the small Mafs have been described only very recently which have been implicated as playing an even more prominent role in modulating chromatin activity through the MARE element. The Bach family proteins (33) harbor a protein–protein interaction domain (called a BTB domain, for broad complex tramtrack bric a brac, also known as a POZ domain, for pox and zinc finger) in addition to the characteristic CNC-type b-Zip structural elements. Both Bach1 and Bach2 bind to the MARE consensus sequence as heterodimers with MafK (33). Since homologous proteins bearing a BTB domain are known to directly modulate chromatin structure (44–46), the Bach family proteins, together with MafK, may make it possible for MARE sites to exert transcriptional regulatory effects while eliciting structural changes in chromatin. Such bifunctional regulation could be advantageous for loci in which chromatin structural alterations either precede or develop in concert with acquisition of transcriptional activity, for example as is known to take place in the β-globin LCR in differentiating hematopoietic cells (18).

A final partner molecule with which the small Maf proteins are known to associate through traditional protein–protein interactions is c-Fos, which has been reported to heterodimerize with small Mafs and thence bind to the MARE consensus sequence, however, the c-Fos–small Maf heterodimeric complex is unable to activate transcription through that site (7). It is well known that the addition of TPA inhibits DMSO-induced terminal differentiation of mouse erythroleukemia (MEL) cells (47) and the above observation provides one interesting possible explanation for this inhibition, i.e. since TPA is a potent inducer of AP-1, either transcription factor AP-1 itself or the c-Fos protein alone could act to reduce transcriptional activity mediated through MARE sites. Since all MARE elements are also AP-1 binding sites, an increase in AP-1 activity could result in a direct competition between AP-1 (Jun/Fos) and CNC/Maf for the site. Alternatively, induction of c-Fos alone could lead to sequestration of small Mafs away from their presumptive CNC partner molecules, thus providing another avenue to inhibit small Maf/CNC heterodimer formation and activity (7).

In summary, the small Maf proteins, acting in concert with multiple larger partner molecules, each of which displays its own unique developmental and temporal expression profile, form a complex network of interacting b-Zip transcription factors (Fig. 4). In addition, the small Maf proteins form homodimers, which share DNA binding site specificity (unlike CNC family homodimers, which may not bind or bind only weakly to DNA). Some of the large Maf proteins are also capable of heterodimerizing with components of AP-1 and these heterodimers then preferentially bind to MARE *cis* elements. One should probably conclude at the present time that the b-Zip protein network that appears to impinge on transcription and chromatin structure includes the large and small Maf proteins, c-Fos (and, by extension, members

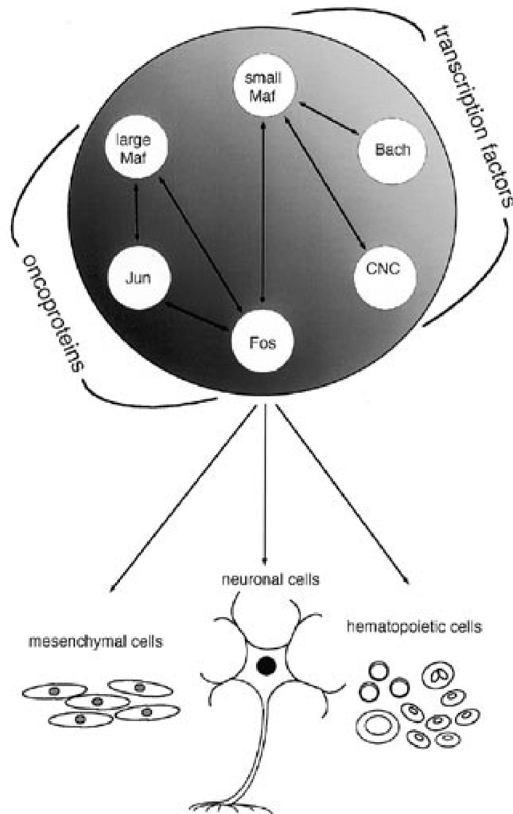

Figure 4. A network of transcription factors and oncoproteins capable of interacting through the b-Zip structure. Binary arrows indicate possible interactions among the families of b-Zip transcription factors. These network interactions appear to be important for the Maf family proteins to regulate differentiation of mesenchymal, hematopoietic and neuronal cells.

of the Fos and Jun families) as well as the Bach and CNC proteins (Fig. 4).

Barring the future cloning of a few more outlying members of these extended families, are these the likely limits of the interactions the small Maf proteins might have? Very recent observations portend that, rather than having reached an end to further interactions, there may be future collisions between two rather extended galaxies of transcription factor families. This speculation rests on the recent demonstration that the CNC protein p45 can also interact with the thyroid hormone (T3R) and retinoic acid receptors (RAR) to serve as a potent co-activator of these nuclear hormone receptors (48); the small Maf proteins interact with p45 to antagonize this potentiation. Thus the small Mafs might also function as a switch to turn off nuclear hormone signaling responses. The possibility exists that this report heralds the revelation of an increasing number of transcription factors that may participate in the Maf network by interacting with these proteins through functional domains lying outside the b-Zip structure.

CONTRIBUTION OF THE MAF NETWORK TO DEVELOPMENT

Recent data from gene targeting experiments provide some insight to help us understand the roles the small Maf proteins might play in development. *mafK* null mutant mice were found to be viable, fertile and apparently normal and healthy (49). Their peripheral blood cell counts, red cell parameters and blood smears were also within the normal range. NF-E2 binding activity in fetal livers of *mafK* null mutants was indistinguishable from that of the wild-type animals (49), suggesting the presence of a fully complementing activity. Since the same three small Maf proteins found in chickens (50) are also expressed in the mouse (unpublished observations), the most likely explanation is that MafF or MafG replace the ablated *mafK* gene function, since the chicken small Maf proteins are able to functionally substitute for one another (9).

Additional key insights into small Maf function were also provided from forced expression experiments (50,51). When *mafK* was stably transformed into murine erythroleukemia (MEL) cells under the control of a conditionally inducible (metallothionein gene) promoter, the MEL cells were induced to differentiate simply by the addition of zinc to the culture medium (i.e. in the absence of any other inducers; 50). This result suggested that quantitative control over *mafK* expression was important for the erythroid differentiation process (see Fig. 3). The result could be interpreted in two ways. The ectopic increase in MafK could have resulted in recruitment of a heterodimeric partner molecule which could then activate transcription of terminal erythroid genes through MARE sites. Alternatively, increased levels of MafK could have led to formation of homodimers which then repressed transcription from target genes required for proliferation and which normally inhibit differentiation of MEL cells. A likely resolution to these alternative explanations for the role of small Maf proteins in erythroid differentiation was suggested by preparing a dominant negative mutant of MafK (dnMafK), which was able to form homo- and heterodimers but rendered any such complex unable to bind to a MARE site (51). Expression of dnMafK in MEL cells lowered the overall binding activity to MARE sites in cell extracts, as anticipated, and led to decreased expression of the globin genes (51). These results suggested that under normal conditions MafK might be limiting in uninduced MEL cells and that its forced expression allowed complex formation with an activating partner molecule that was already present in the cell.

Recently several publications have appeared which underscore the developmental significance of the large Maf family proteins during neurogenesis. c-Maf was shown to be an important regulator of neuron-specific *L7* gene expression in Purkinje cells (13). MafB plays an apparently critical role in segmentation of the hindbrain (11), as concluded from the discovery that the mouse *Kreisler* (*kr*) phenotype is due to mutation of the *mafB* gene. The neural retina is another site where large Maf family proteins play prominent roles. Both the opsin and rhodopsin genes were shown to be targets for the activity of NRL (27,28) and the retina-specific *QR1* gene has recently been shown to be activated by both c-Maf and MafB (53). The expression profile of the small Maf protein MafK largely overlaps that of the large Maf proteins in neural tissues (31).

While hematopoiesis has been a focal topic for CNC family activity for some time, large Maf molecules are also emerging as equally important players in hematopoietic differentiation. c-Maf was found to be a key regulator of Th2 (one subset of $CD4^+$ T helper cells)-specific expression of IL-4 (12). The *cis* elements targeted by c-Maf in the IL-4 promoter differ substantially from the MARE consensus sequence, which may explain the reason

why synergy with NF-AT is critical for c-Maf to function on the IL-4 promoter. MafB was shown to interact with Ets-1 to inhibit *trans*-activation of the transferrin receptor gene and additionally was found to exert an inhibitory effect on differentiation of chicken erythroblast cell line HD3 (10). MafB is highly expressed in myelomonocytic cells, but not in erythroid cells. Thus MafB is a good candidate to play an important role in myelomonocytic cell differentiation, possibly by inhibiting expression of erythroid genes.

Gene disruption of the CNC family members has also been reported. A *p45* null mutant mouse showed unexpectedly severe hemorrhage due to a lack of platelets, indicating that p45 is essential for megakaryopoiesis (54). Surprisingly, erythropoiesis was not affected significantly by the mutation, suggesting that activation signals for erythroid transcription can be transduced through MARE sites even in the absence of p45. Another member of the CNC family, Nrf2, has also been disrupted (55) and *Nrf2* null mutant mice developed normally and were fertile. *Nrf2* null mutant mice were not anemic, again indicating that erythropoiesis was not significantly affected by loss of Nrf2. Very recently the third member of the CNC family, Nrf1, was disrupted (56). While *Nrf1* null mutant mice die prior to 7.5 d.p.c. from a failure to complete gastrulation, *Nrf1* null ES cells injected into wild-type blastocysts contributed to all mesodermal lineages tested, including blood, indicating that the *Nrf1* homozygous null mutation results in a defect in hematopoiesis that is not cell autonomous. These results provoke the question: which is the genuine partner molecule of the small Maf proteins with which it binds to and activates erythroid MARE sites? Is there a new molecule involved, or are p45, Nrf1 and Nrf2 mutually compensating for one anothers' loss in erythroid cells?

The β-globin LCR is generally agreed to confer the ability to regulate transcription by first initiating structural alterations in chromatin (18,57); what it does thereafter is the subject of considerable current debate. In this regard, it is interesting to note that the Bach proteins contain a so-called BTB domain, which has been implicated in remodeling of chromatin structure (44–46). We speculate that the Bach proteins are one of the important potential candidates for heterodimerizing partners of the small Mafs in erythroid cells, since the very real possibility exists that a heterodimer formed between a small Maf and a Bach protein could be a key regulator of LCR chromatin structure. Disruption of the Bach genes may provide further clues to this puzzle, since several of the predicted phenotypes are so clear.

Given this vast repertoire of potential partners for the small Mafs, studies are now underway to determine which molecule is the physiological partner in each of the exemplary physiological situations. Still unidentified partners may yet play a significant role in this evolving discovery process. Characterization of heterodimers of the small Mafs and partner molecules, especially with regard to the differences they might evoke in transcriptional activity of specific target sequences, in their responses to signals from outside the cell and their influences on chromatin structure, will help to decipher the mechanisms of gene regulation during these varied cellular processes.

CONCLUSIONS

Just as one gently folds back the petals of an intricately layered flower to finally reveal the intrinsic beauty of the whole, detailed examination at each level of discovery has led to an ever-increasing comprehension and appreciation of the complexity of regulatory control over vertebrate gene expression elicited by Maf family transcription factors. The number of proteins that participate in the Maf network continues to grow, both through the discovery of new partner molecules that heterodimerize with the Maf leucine zipper motif and through apparently non-canonical protein interaction domains. Where this will all eventually lead is unknown, but the recent intersection of the Maf network with the AP-1 and subsequently with the T3R/RAR networks leads to anticipation that the story will become even more intriguing before it is complete.

Analysis of small Maf proteins, along with analysis of NF-E2, have led to significant new insights into the mechanisms that mediate lineage determination and differentiation. We envisage that establishing or elaborating a cell lineage occurs through combinatorial interactions that precisely balance the activities of all the factors that participate in this elaborate protein interaction network. This hypothesis suggests that it is the sum of a combination of interactive transcription factor affinities (both for one another and for the target sites to which each binds in DNA) that finally dictates cellular responses which either prohibit or initiate differentiation and growth programs. This concept thus advocates the inverse view of the 'master regulator' hypothesis, where it is thought that a single determinative protein dictates the fate of multipotent cells to a particular differentiated tissue or cell type.

Finally, it is also interesting to speculate how this finely balanced intracellular equilibrium might become unbalanced during oncogenesis. One simple hypothesis suggests that cell transformation may be the manifestation of perturbation of this b-Zip network (7,42). It is well known that forced expression of components of transcription factor AP-1 induces cell transformation (58,59). As discussed in this review, the possibility certainly exists that inappropriately expressed AP-1 components may sequester small Mafs or other members of the b-Zip network from their partner molecules or target sequences, so that these AP-1 components can inhibit the cellular differentiation process and provoke cell transformation. To elucidate how this b-Zip protein network contributes to cell differentiation processes and how its perturbation leads to cell transformation, it is important to identify and comprehensively categorize the factors interacting with the Maf family proteins, to examine their modes of DNA binding and to analyze the interactions among them.

ACKNOWLEDGEMENTS

We thank Drs Kim-Chew Lim and Makoto Nishizawa for illuminating discussions. This work was supported by research grants from the Ministry of Education, Science and Culture (H.M., K.I. and M.Y.), the Japanese Society for Promotion of Sciences (M.Y.), the Uehara Memorial Foundation (M.Y.), the Mochida Memorial Foundation for Medical and Pharmaceutical Research (K.I.), the Ciba-Geigy Foundation for the Promotion of Science (K.I.), an MSTP training grant to Northwestern University (GM 08152, J.A.S.) and the NIH (GM 28896, J.D.E.).

REFERENCES

1 Hunter,T. (1997) *Cell*, **88**, 333–346.
2 Nishizawa,M., Kataoka,K., Goto,N., Fujiwara,K. and Kawai,S. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 7711–7715.
3 Kataoka,K., Nishizawa,M. and Kawai,S. (1993) *J. Virol.*, **67**, 2133–2141.

4 Kataoka,K., Noda,M. and Nishizawa,M. (1994) *Mol. Cell. Biol*., **14**, 700–712.
5 Swaroop,A., Xu,J., Pawar,H., Jackson,A., Skolnick,C. and Agarwal,N. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 266–270.
6 Fujiwara,K., Kataoka,T.K. and Nishizawa,M. (1993) *Oncogene*, **8**, 2371–2380.
7 Kataoka,K., Igarashi,K., Itoh,K., Fujiwara,K.T., Noda,M., Yamamoto,M. and Nishizawa,M. (1995) *Mol. Cell. Biol*., **15**, 2180–2190.
8 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.
9 Igarashi,K., Kataoka,K., Itoh,K., Hayashi,N., Nishizawa,M. and Yamamoto,M. (1994) *Nature*, **367**, 568–572.
10 Sieweke,M.H., Tekotte,H., Frampton,J. and Graf,T. (1996) *Cell*, **85**, 49–60.
11 Cordes,S.P. and Barsh,G.S. (1994) *Cell*, **79**, 1025–1034.
12 Ho,I.-C., Hodge,M.R., Rooney,J.W. and Glimcher,L.H. (1996) *Cell*, **85**, 973–983.
13 Kurschner,C. and Morgan,J.I. (1995) *Mol. Cell. Biol*., **15**, 246–254.
14 Stamatoyannopoulos,G. and Neinhuis,A.W. (1994) In Stamatoyannopoulos,G., Nienhuis,A.W., Majerus,P. and Varmus,H. (eds) *The Molecular Basis of Blood Diseases.* (2nd edn) W.B.Saunders Co., Philadelphia, PA, USA. Chapter 4, pp. 107–155.
15 Fraser,P., Pruzina,S. and Grosveld,F. (1993) *Genes Dev*., **7**, 106–113.
16 Bungert,J., Dave,U., Lieuw,K.-H., Lim,K.-C., Shavit,J.A., Liu,Q. and Engel,J.D. (1995) *Genes Dev*., **9**, 3083–3096.
17 Wijgerde,M., Grosveld,F. and Fraser,P. (1995) *Nature*, **377**, 209–213.
18 Jimenez,G., Griffiths,S.D., Ford,A.M., Greaves,M.F. and Enver,T. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 10618–10622.
19 Ney,P.A., Sorrentino,B.P., McDonagh,K.T. and Nienhuis,A.W. (1990) *Genes Dev*., **4**, 993–1006.
20 Ney,P.A., Sorrentino,B.P., Lowrey,C.H. and Nienhuis,A.W. (1990) *Nucleic Acids Res*., **18**, 6011–6017.
21 Talbot,D. and Grosveld,F. (1991) *EMBO J*., **10**, 1391–1398.
22 Strauss,E.C. and Orkin,S.H. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 5809–5813.
23 Andrews,N.C., Erjument-Bromage,H., Davidson,M.B., Tempst,P. and Orkin,S.H. (1993) *Nature*, **362**, 722–728.
24 Stamatoyannopoulos,J.A., Goodwin,A., Joyce,T. and Lowrey,C.H. (1995) *EMBO J*., **14**, 106–116.
25 Mignotte,V., Wall,L., deBoer,E., Grosveld,F. and Romeo,P.-H. (1989) *Nucleic Acids Res*., **17**, 37–54.
26 Taketani,S., Inazawa,J., Nakanishi,Y., Abe,T. and Tokunaga,R. (1992) *Eur. J. Biochem*., **205**, 217–222.
27 Kumar,R., Chen,S., Scheurer,D., Wang,Q.L., Duh,E., Sung,C.H., Rehemtulla,A., Swaroop,A., Adler,R. and Zack,D.J. (1996) *J. Biol. Chem*., **271**, 29612–29618.
28 Rehemtulla,A., Warwar,R., Kumar,R., Ji,X., Zack,D.J. and Swaroop,A. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 191–195.
29 Toki,T., Itoh,J., Kitazawa,J., Arai,K., Hatakeyama,K., Adasaka,J., Igarashi,K., Nomura,N., Yokoyama,M., Yamamoto,M. and Ito,E. (1997) *Oncogene*, **14**, 1901–1910.
30 Igarashi,K., Itoh,K., Motohashi,H., Hayashi,N., Matuzaki,Y., Nakauchi,H., Nishizawa,M. and Yamamoto,M. (1995) *J. Biol. Chem*., **270**, 7615–7624.
31 Motohashi,H., Igarashi,K., Ohtani,H., Nishizawa,M., Engel,J.D. and Yamamoto,M. (1996) *Genes Cells*, **1**, 223–238.
32 Itoh,K., Igarashi,K., Hayashi,N., Nishizawa,M. and Yamamoto,M. (1995) *Mol. Cell. Biol*., **15**, 4184–4193.
33 Oyake,T., Itoh,K., Motohashi,H., Hayashi,N., Hoshino,H., Nishizawa,M., Yamamoto,M. and Igarashi,K. (1996) *Mol. Cell. Biol*., **16**, 6083–6095.
34 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.
35 Chan,J.Y., Han,X.-L. and Kan,Y.W. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 11366–11370.
36 Caterina,J.J., Donze,D., Sun,C.-W., Ciavatta,D.J. and Townes,T.M. (1994) *Nucleic Acids. Res*., **22**, 2383–2391.
37 Chan,J.Y., Han,X.-L. and Kan,Y.W. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 11371–11375.
38 Moi,P., Chan,K., Asunis,I., Cao,A. and Kan,Y.W. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 9926–9930.
39 Chui,D.H.K., Tang,W. and Orkin,S.H. (1995) *Biochem. Biophys. Res. Commun*., **209**, 40–46.
40 Mohler,J., Vani,K., Leung,S. and Epstein,A. (1991) *Mech. Dev*., **34**, 3–9.
41 Johnsen,O., Skammelsrud,N., Luna,L., Nishizawa,M., Prydz,H. and Kolsto,A.B. (1996) *Nucleic Acids Res*., **24**, 4289–4297.
42 Engel,J.D. (1994) *Nature*, **367**, 516–517.
43 Armstrong,J.A. and Emerson,B.M. (1996) *Mol. Cell. Biol*., **16**, 5634–5644.
44 Dorn,R., Krauss,V., Reuter,G. and Saumweber,H. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 11376–11380.
45 Farkas,G., Gausz,J., Galloni,M., Reuter,G., Gyurkovics,H. and Karch,F. (1994) *Nature*, **371**, 806–808.
46 Albagli,O., Dhordain,P., Deweindt,C., Lecocq,G. and Leprince,D. (1995) *Cell Growth Differentiat*., **6**, 1193–1198.
47 Yamasaki,H., Fibach,E., Nudel,U., Weinstein,I.B., Rifkind,R.A. and Marks,P.A. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 3451–3455.
48 Cheng,X., Reginato,M.J., Andrews,N.C. and Lazar,M.A. (1997) *Mol. Cell. Biol*., **17**, 1407–1416.
49 Kotkow,K.J. and Orkin,S.H. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 3514–3518.
50 Igarashi,K., Itoh,K., Hayashi,N., Nishizawa,M. and Yamamoto,M. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 7445–7449.
51 Kotkow,K.J. and Orkin,S.H. (1995) *Mol. Cell. Biol*., **15**, 4640–4647.
52 Kataoka,K., Fujiwara,K.T., Noda,M. and Nishizawa,M. (1994) *Mol. Cell. Biol*., **14**, 7581–7591.
53 Pouponnot,C., Nishizawa,M., Calothy,G. and Pierani,A. (1995) *Mol. Cell. Biol*., **15**, 5563–5575.
54 Shivdasani,R.A., Rosenblatt,M.F., Zucker-Franklon,D.C., Jackson,W., Hunt,P., Saris,C.J. and Orkin,S.H. (1995) *Cell*, **81**, 695–701.
55 Chan,K., Lu,R., Chan,J.C. and Kan,Y.W. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 13943–13948.
56 Farmer,S.C., Sun,C.W., Winnier,G.E., Hogan,B.L. and Townes,T.M. (1997) *Genes Dev*., **11**, 786–798.
57 Martin,D.I.K., Fiering,S. and Groudine,M. (1996) *Curr. Opin. Genet. Dev*., **6**, 488–495.
58 Distel,R. and Spiegelman,B.M. (1990) *Adv. Cancer Res*., **55**, 37–55.
59 Vogt,P.K. and Bos,T.J. (1990) *Adv. Cancer Res*., **55**, 1–35.
60 Luna,L., Johnsen,O., Skartlien,A.H., Pedeutoru,F., Turc-Carel,C., Prydz,H. and Kolsto,A.-B. (1994) *Genomics*, **22**, 553–562.